

Т.Фудзисава
Т.Касами

**МАТЕМАТИКА
ДЛЯ
РАДИОИНЖЕНЕРОВ**

**Теория
дискретных
структур**

Т.Фудзисава
Т.Касами

МАТЕМАТИКА ДЛЯ РАДИОИНЖЕНЕРОВ

Теория дискретных структур

Перевод с японского
А. В. Кузнецова



Москва
«Радио и связь»
1984

ББК 22.12
Ф94
УДК 5+621.37/39

Фудзисава Т., Касами Т.

Ф94 Математика для радиоинженеров: Теория дискретных структур: Пер. с япон. — М.: Радио и связь, 1984. — 240 с., ил.

В пер.: 1 р. 50 к.

Излагаются основы специальных разделов современной дискретной математики, широко используемые при разработке цифровых систем связи, сетей ЭВМ и вычислительной техники. Рассматриваются теория чисел, комбинаторика, теория графов, теория конечных автоматов, булева алгебра и логические функции. Значительное место уделено решению задач, примеров.

Для инженерно-технических работников и студентов, специализирующихся в области систем связи, электроники и вычислительной техники.

Ф $\frac{1702070000-041}{046(01)-84}$ 102-83

ББК 22.12

6Ф0.3

Редакция литературы по электронной технике

Fujisawa T., Kasami T.

MATHEMATICS FOR ELECTRONICS AND COMMUNICATIONS ENGINEERS, THEORY OF DISCRETE STRUCTURE

Тосио Фудзисава, Тадао Касами

МАТЕМАТИКА ДЛЯ РАДИОИНЖЕНЕРОВ:
ТЕОРИЯ ДИСКРЕТНЫХ СТРУКТУР

Редактор С. И. Гельфанд

Редактор издательства Т. В. Жукова

Художественный редактор Г. Н. Кованов

Технические редакторы Л. А. Горшкова, Г. И. Колосова

Корректор Л. С. Глаголева

ИБ № 495

Сдано в набор 27.04.83 г.

Подписано в печать 13.09.83 г.

Формат 60×90/16

Бумага тип. № 2

Гарнитура литературная

Печать высокая

Усл. печ. л. 15,0

Усл. кр.-отт. 15,0

Уч.-изд. л. 15,81

Тираж 10 000 экз.

Изд. № 20008

Зак. № 62

Цена 1 р. 50 к.

Издательство «Радио и связь». 101000 Москва, Почтамт, а/я 693

Московская типография № 5 ВГО «Союзучетиздат»
101000 Москва, ул. Кирова, д. 40

© Научное общество электроники и связи.

Фудзисава Т., Касами Т., 1977

© Перевод на русский язык, предисловие, примечание,
издательство «Радио и связь», 1984

Предисловие к русскому изданию

Теория чисел и комбинаторика, теория графов и теория автоматов, булевы алгебры и основы теории логических схем — вот основные темы этой книги. Все эти разделы современной дискретной математики абсолютно необходимы инженерам, специализирующимся в области систем связи, электроники и современной вычислительной техники. Однако до сих пор знакомиться с ними приходилось по специализированным университетским курсам, каждый из которых посвящен, как правило, только одному из разделов, и зачастую является весьма солидным изданием. Это представляет значительные неудобства, преодолеть которые иногда оказывается очень трудным.

Книга Т. Фудзисавы и Т. Касами «Математика для радиоинженеров: Теория дискретных структур» позволяет значительно облегчить процесс овладения необходимыми элементами современной математики. Предлагаемая книга возникла из лекций, которые авторы в течение ряда лет читали будущим инженерам в университете г. Осака. Поэтому выбор материала и его изложение тщательно продуманы, что облегчает читателю понимание новых для него (и зачастую, достаточно непривычных) вопросов.

В своем предисловии авторы книги приводят достаточно подробное описание содержания книги, и дают указания, в каком порядке можно изучать различные ее разделы. Следует только сказать, что большую пользу принесет решение предлагаемых задач.

При переводе был исправлен ряд замеченных опечаток и неточностей. Мы старались также переводить специальные термины так, чтобы они возможно ближе соответствовали терминологии, принятой в советской литературе.

В заключение хотелось бы надеяться, что предлагаемая книга принесет пользу возможно более широкому кругу читателей.

С. И. Гельфанд

Предисловие к японскому изданию

Среди технических наук электроника и техника связи в наибольшей степени используют математику. Действительно, если взглянуть на программы обучения студентов в университетах как в Японии, так и в других странах, то легко обнаружить, что математические дисциплины занимают в них существенное место. В последние годы особую важность приобрели те разделы математики, которые имеют отношение к развитию цифровых устройств, цифровой связи и цифровых вычислительных машин. Базой для преподавания этих дисциплин наряду с классическими методами анализа непрерывных физических моделей, составляющих до последнего времени основной предмет радиотехники и электроники, стали алгебраические, логические и комбинаторные методы исследования различных моделей дискретной математики. К сожалению, в средней школе и институтах этим предметам не уделяется достаточного внимания. Поэтому особо важное значение приобретают специальные курсы. Настоящая книга была задумана именно как специальный курс по теории дискретных структур. С момента возникновения идеи написать такую книгу до ее выхода в свет прошло несколько лет. За это время было опубликовано несколько монографий на близкие темы. Появление этих книг, а также данной книги значительно расширяет возможность выбора материала для чтения или преподавания с учетом предварительной подготовки и интересов той или иной аудитории. Можно сказать, что пришло время, когда стало возможным удовлетворить запросы «разнотипных» читателей.

Книга предназначена как для студентов начальных курсов, изучающих основы электроники, техники связи и вычислительной техники, так и для студентов старших курсов, приступающих к самостоятельным исследованиям. Материал книги и способ его изложения подбирались таким образом, чтобы ее можно было использовать и как учебник, и как специальную монографию. С теми вопросами, которые лишь слегка затронуты в книге, можно ознакомиться по другим указанным в книге и вполне доступным источникам. Выбрав главу, которую вы будете изучать, следует провести классификацию и поиск в избранной области, ознакомиться с подробностями теории групп, колец, полей и других алгебраических конструкций, конечных полей и их приложениями. Эти вопросы рассмотрены в учебниках по алгебре или теории кодирования, указанных в библиографии.

Главы 1—5 написаны Т. Фудзисава, гл. 6—10 — Т. Касами. В гл. 1 приводятся элементарные сведения из теории целых чисел. Главы 2, 3 и 4, 5 можно читать почти независимо. Первые две из них дают представление об основных комбинаторных методах, а гл. 4, 5 представляют собой введение в теорию графов. Главы 6—10 можно читать почти независимо от глав 1—5. В этих главах используются элементарные понятия из гл. 1, а также определения и простейшие свойства графов из гл. 4, 5. В гл. 7—10 материал гл. 6 непосредственно не используется (за исключением задач в гл. 7, помеченных знаком γ). Прочитав, например, гл. 7, § 8.1, 8.3, гл. 9 и § 10.3, можно ознакомиться с основами теории проектирования цифровых устройств. Параграфы 10.1 и 10.2 можно читать независимо от гл. 6—9; их можно рассматривать как введение в теорию конечных автоматов.

Поскольку в последние годы знакомство с теорией множеств происходит на самых ранних этапах обучения, основные термины и понятия, относящиеся к множествам, используются в книге без каких-либо пояснений. В остальном, для чтения книги достаточно знакомства с обычными курсами математики для студентов технических специальностей институтов. Для понимания § 4.5, 5.5 и второй половины § 8.4 необходимы некоторые знания из линейной алгебры. Параграфы 4.5 и 5.5 посвящены теории электрических цепей, § 8.4 — теория кодирования и их можно опустить без ущерба для понимания остального материала.

В основу этой книги положен материал, использованный авторами при чтении лекций по курсам анализа информационных систем, логического проектирования и конечных автоматов студентам отделений электротехники, систем управления и информационной техники факультета инженерных наук университета г. Осака. Мы хотели бы выразить искреннюю признательность за помощь в написании этой книги профессору Н. Токура, доцентам Н. Тамота и К. Танигути. Особо хотелось бы поблагодарить профессора Н. Токура за подготовку упражнений и задач к гл. 8—10. Мы также признательны членам редакционной коллегии монографий для университетов Научного общества электроники связи Японии за предоставленную нам возможность написать эту книгу, а также сотрудникам издательства «Корона» за большую помощь при подготовке книги к изданию.

Авторы

Глава 1

Свойства целых чисел

1.1. Упорядоченность

Обозначим через Z множество всех целых чисел. Целые числа имеют естественный порядок

..., -3 , -2 , -1 , 0 , 1 , 2 , 3 , ...

Это значит, что они связаны отношением больше — меньше, которое устанавливает, что находящееся справа число всегда больше числа, находящегося слева от него. Множество всех положительных целых чисел обозначим через Z^+ . При этом запись $a < b$, означающая, что в вышеприведенной последовательности число a находится слева от числа b , эквивалентна тому, что $b - a \in Z^+$. Очевидно, бинарное отношение \leq обладает следующими тремя свойствами.

A1. (Рефлексивность): $a \leq a$.

A2. (Антисимметричность): если $a \leq b$ и $b \leq a$, то $a = b$.

A3. (Транзитивность): если $a \leq b$ и $b \leq c$, то $a \leq c$.

Эти три свойства называются аксиомами порядка. Кроме отношения сравнимости целых чисел аксиомам порядка удовлетворяет также отношение включения \subset между множествами. Если в некотором множестве G определено бинарное отношение, \leq , удовлетворяющее аксиомам A1—A3, то G называется упорядоченным множеством. Если для произвольных двух элементов, $x, y \in G$, обязательно либо $x \leq y$, либо $y \leq x$, то такую упорядоченность называют полной; в противном случае говорят о частичной упорядоченности. Отношение больше — меньше в Z является отношением полной упорядоченности.

Минимальным элементом непустого подмножества S упорядоченного подмножества G называется элемент $s \in S$ такой, что $s \leq x$ для любого $x \in S$.

Упражнение 1.1. Покажите, что если непустое подмножество S упорядоченного множества G имеет минимальный элемент, то он единствен.

Решение. Предположим, что s и s' — два минимальных элемента подмножества S . Тогда, по определению, $s \leq s'$ и $s' \leq s$. Следовательно, в силу аксиомы антисимметричности $s = s'$.

Если произвольное непустое подмножество упорядоченного множества G имеет минимальный элемент, то G называется вполне упорядоченным множеством.

Изложение общей теории целых чисел не является целью данной книги, и поэтому следующую теорему приведем здесь без доказательства, полагая, что читатель знаком с основными свойствами четырех арифметических операций и при необходимости может обратиться к более полным монографиям [1, 2].

Теорема 1.1. Множество всех целых положительных чисел является вполне упорядоченным множеством относительно отношения больше — меньше, и его минимальным элементом является число 1.

Например, минимальный элемент множества положительных четных чисел число 2. Множество всех целых чисел Z минимального элемента не имеет и поэтому не является вполне упорядоченным множеством.

Следствие из теоремы 1.1. Пусть некоторое множество целых положительных чисел S удовлетворяет следующим условиям: 1) $1 \in S$; 2) если $n \in S$, то $n+1 \in S$. Тогда $S = Z^+$.

Доказательство. Пусть S' — множество целых положительных чисел, не входящих в S . Покажем, что $S' = \emptyset$. Если S' не пусто, то оно имеет минимальный элемент; обозначим его через s' . Так как $s' \in S'$, то $s' \neq 1$ и, следовательно, $s'-1$ является целым положительным числом и, значит, принадлежит S . В таком случае, по второму условию $s' \in S$. Имеет место противоречие.

Следствие дает следующий закон индукции.

Закон индукции. Пусть каждому целому положительному числу n соответствует некоторое истинное или ложное суждение $P(n)$. Если 1) $P(1)$ истинно и 2) для всех $k \in Z^+$ можно доказать, что из истинности $P(k)$ следует истинность $P(k+1)$, то суждение $P(n)$ будет истинным для всех целых положительных чисел $n \in Z^+$.

Закон индукции можно модифицировать, если заменить второе условие следующим: если для всех $k \in Z^+$ можно доказать, что из истинности $P(1), \dots, P(k)$ следует истинность $P(k+1)$. Доказательство этого факта можно провести по аналогии с доказательством следствия из теоремы 1.1; читателям предлагается провести его самостоятельно.

Если для чисел $a, b, q \in Z$ выполняется соотношение $a = bq$, то говорят, что a является кратным b , а b — делителем a . При этом также говорят, что b делит a , и записывают это следующим образом: $b|a$. В противном случае пишут $b \nmid a$.

Очевидно, выполняются следующие свойства:

Закон рефлексивности:

$$a|a. \quad (1.1)$$

Закон транзитивности: если

$$a|b \text{ и } b|c, \text{ то } a|c. \quad (1.2)$$

Закон антисимметричности, вообще говоря, не выполняется. Так, например, числа 2 и -2 делят друг друга, но равными не являются. Основные свойства четырех арифметических операций над целыми числами предполагаются читателям хорошо известными. Поэтому следующая теорема, непосредственно из них вытекающая, здесь приводится без доказательства.

Теорема 1.2. Делителями числа 1 являются только $+1$ и -1 .

Следствие 1 из теоремы 1.2. Если $a|b$ и $b|a$, то $a=b$ или $a=-b$.

Доказательство. Так как $a=bd_1$ и $b=ad_2$, то $a=ad_1d_2$. Поэтому $1=d_1d_2$ и, согласно теореме, $d_1=\pm 1$.

Это следствие показывает, что в множестве целых положительных чисел закон антисимметричности выполняется.

Следствие 2 из теоремы 1.2. Если $a, b \in Z^+$, $a|b$ и $b|a$, то $a=b$.

Как следует из свойств (1.1), (1.2) и следствия 2, бинарное отношение $|$ (называемое делимостью) в множестве натуральных чисел Z^+ является отношением порядка, удовлетворяющим аксиомам А.1—А.3. Таким образом одно и то же множество может быть упорядочено по-разному, а следовательно, всякий раз, когда в этом есть необходимость, необходимо ясно указывать, с помощью какого бинарного отношения, $\langle Z^+, \leq \rangle$ или $\langle Z^+, | \rangle$, осуществляется упорядочивание. Так как $2 \nmid 3$ и $3 \nmid 2$, то в Z^+ отношение $|$ является отношением частичной упорядоченности. Далее, так как множество $\{n \in Z^+ \mid n > 1\}$ не имеет минимального элемента, то, очевидно, что Z^+ не является вполне упорядоченным множеством относительно отношения порядка $|$.

Наиболее часто используемые свойства делимости содержатся в следующих двух теоремах.

Теорема 1.3. Если все целые числа, входящие в равенство $k+l+\dots+m=p+q+\dots+s$, за исключением одного, являются кратными числу b , то и это оставшееся число является кратным b .

Теорема 1.4. Для произвольного целого числа $a \in Z$ и целого положительного числа $b \in Z^+$ однозначно определяются частное q и остаток r , такие, что

$$a=bq+r, \quad 0 \leq r < b. \quad (1.3)$$

Предположим, что непустое множество целых чисел S замкнуто относительно операций сложения и вычитания. Другими словами, будем считать, что для произвольных чисел $a, b \in S$ их сумма $a+b$ и разность $a-b$ также являются элементами S . Предположим, что S содержит элемент $a \neq 0$. Тогда $a-a=0$ и, следовательно, $0-a=-a \in S$. Поскольку либо a , либо $-a$ — целое положительное число, S содержит, по крайней мере, одно целое положительное число. Следовательно, в S имеется минимальное положительное целое число b .

По индукции можно доказать, что в S содержатся и все числа, кратные b . Действительно, $2b=b+b$ входит в S , $3b=2b+b$ также входит в S и т. п. Так как $-b \in S$, то же верно для чисел $-2b, -3b, \dots$

Если разделить произвольное число $a \in S$ на b , то можно найти частное q и остаток r , удовлетворяющие (1.3). Так как $a \in S$ и $bq \in S$, то $r = a - bq \in S$. Таким образом, если $r > 0$, то неравенство $r < b$ противоречит тому, что b — это минимальное целое число в S . Следовательно, $r = 0$, и a кратно b . Отсюда получаем следующую теорему.

Теорема 1.5. Непустое множество целых чисел, замкнутое относительно операций сложения и вычитания, состоит либо только из одного числа 0, либо в нем имеется минимальное положительное целое число b и множество состоит из всех чисел, кратных b .

1.2. Наибольший общий делитель

Число p называется общим кратным чисел a, b, \dots, l , если оно кратно каждому из них. Очевидно, что множество S всех общих кратных чисел a, b, \dots, l замкнуто относительно операций сложения и вычитания. За исключением случая, когда одно из чисел a, b, \dots, l равно 0, когда S состоит только из одного числа 0, в S всегда имеется минимальное положительное целое число (теорема 1.5). Это число называется наименьшим общим кратным. В данной главе оно обозначается через $[a, b, \dots, l]$. Другими словами, наименьшим общим кратным некоторой совокупности отличных от 0 целых чисел называется положительное целое число, являющееся наименьшим в совокупности всех общих кратных этих чисел. Согласно теореме 1.5, множество всех общих кратных S совпадает с множеством всех чисел, кратных наименьшему общему кратному.

Число p называется общим делителем чисел a, b, \dots, l , если оно является делителем каждого из них. За исключением крайнего случая, когда все числа a, b, \dots, l равны 0, среди общих делителей имеется наибольший, поскольку все они по абсолютной величине не превосходят наименьшего из $|a|, |b|, \dots, |l|$. Этот делитель называется наибольшим общим делителем, в данной главе он обозначается через (a, b, \dots, l) . Если $(a, b, \dots, l) = 1$, то числа a, b, \dots, l называют взаимно простыми.

Упражнение 1.2. Если $a \in \mathbb{Z}^+$, то $(a, 0) = a$; если $-a \in \mathbb{Z}^+$, то $(a, 0) = -a$.

Упражнение 1.3. Покажите, что если $b|a$ и $b \in \mathbb{Z}^+$, то множество общих делителей чисел a и b и множество делителей b совпадают и, в частности, $(a, b) = b$.

Решение. Ясно, что общий делитель a и b является делителем b . Наоборот, если $x|b$, то $x|a$ (так как $b|a$), т. е. x является общим делителем a и b . Следовательно, совокупность общих делителей чисел a и b совпадает с совокупностью делителей b . Так как $b \in \mathbb{Z}^+$, то среди делителей числа b наибольшим является оно само, так что $(a, b) = b$.

Упражнение 1.4. Покажите, что если $a = bq + r$, то совокупность общих делителей чисел a и b совпадает с совокупностью общих делителей чисел b и r и, в частности, $(a, b) = (b, r)$.

Решение. Следует показать, что общий делитель чисел a и b является делителем b , а следовательно, и общим делителем b и r и что аналогично общий делитель b и r является общим делителем чисел a и b .

Упражнения 1.3 и 1.4 приводят к следующему алгоритму Евклида для определения наибольшего общего делителя (a, b) чисел $a, b \in \mathbb{Z}^+$:

$$\left. \begin{aligned} a &= bq_1 + r_2, & 0 < r_2 < b; \\ b &= r_2 q_2 + r_3, & 0 < r_3 < r_2; \\ r_2 &= r_3 q_3 + r_4, & 0 < r_4 < r_3; \\ . & \\ r_{n-2} &= r_{n-1} q_{n-1} + r_n, & 0 < r_n < r_{n-1}; \\ r_{n-1} &= r_n q_n. \end{aligned} \right\} \quad (1.4)$$

Заметим, что остаток в этих равенствах с каждым шагом уменьшается, а поэтому процесс деления через некоторое число шагов заканчивается. Так как $(a, b) = (b, r_2)$, $(b, r_2) = (r_2, r_3)$, ..., то

$$(a, b) = (r_{n-1}, r_n). \quad (1.5)$$

Например,

$$\begin{array}{r}
 231 \overline{) 525} \\
 \underline{2 } \\
 63 \\
 \underline{3 } \\
 189 \\
 \underline{42 } \\
 1 \\
 \underline{1 } \\
 0
 \end{array}$$

Далее, из (1.4) получаем, что

$$r_2 = a + (-q_1)b,$$

$$r_3 = (-q_2)a + (1 + q_1q_2)b.$$

Продолжая этот процесс, находим такие $s, t \in \mathbb{Z}$, что $r_n = sa + tb$. Таким образом, получим следующую лемму.

Лемма 1.1. Для любых двух целых не равных одновременно нулю чисел a и b существуют числа $s, t \in \mathbb{Z}$ такие, что

$$(a, b) = sa + tb. \quad (1.6)$$

Пусть S — множество всех целых чисел, получающихся при сложении кратных числа a с кратными числа b , т. е.

$$S = \{n \in \mathbb{Z} \mid n = sa + tb, \ st \in \mathbb{Z}\}.$$

Очевидно, что S замкнуто относительно операций сложения и вычитания. В самом деле,

$$(s_1 a + t_1 b) \pm (s_2 a + t_2 b) = (s_1 \pm s_2) a + (t_1 \pm t_2) b.$$

Так как a и b одновременно не равны нулю, то в S содержатся числа, отличные от 0. Следовательно, согласно теореме 1.5, в S имеется минимальное целое положительное число

$$k = s' a + t' b, \quad (1.7)$$

и S совпадает с совокупностью кратных этого числа k . Поэтому

$$k | (a, b). \quad (1.8)$$

Однако, $(a, b) | a$ и $(a, b) | b$, так что в силу (1.7)

$$(a, b) | k. \quad (1.9)$$

Из (1.8), (1.9) и следствия 2 из теоремы 1.2 вытекает, что $(a, b) = k$.

Таким образом, доказана следующая теорема.

Теорема 1.6. Пусть a и b — два целых числа, одновременно не равных 0. Минимальное целое положительное число в множестве всех чисел вида $sa + tb$, где s и t — произвольные целые числа, является наибольшим общим делителем (a, b) чисел a и b .

Заметим, что если $c | a$ и $c | b$, то в силу (1.6) $c | (a, b)$, т. е. что множество общих делителей чисел a и b совпадает с множеством делителей числа (a, b) .

Теорема 1.7. Если $(a, b) = 1$, $b | ac$, то $b | c$.

Доказательство. Как следует из леммы 1.1, существуют такие числа $s, t \in \mathbb{Z}$, что $sa + tb = 1$. Если это равенство умножить на c и подставить $ac = bq$, то получим, что $sbq + tbc = c$, т. е. $b(sq + tc) = c$.

Упражнение 1.5. Покажите, что если $m \in \mathbb{Z}^+$, то $(am, bm) = m(a, b)$.

Последнее утверждение можно доказать различными способами. Например, можно умножить на m каждое из равенств в (1.4). Так как на m умножаются все числа a, b, r_2, \dots, r_n , то то же самое происходит и с числами q, \dots, q_n . Также просто можно доказать и следующее утверждение.

Упражнение 1.6. Если $m \in \mathbb{Z}^+$, $m | a$, $m | b$, то $(a/m, b/m) = (a, b)/m$ и, в частности, $(a/(a, b), b/(a, b)) = 1$.

Для того чтобы исследовать связь между наименьшим общим кратным и наибольшим общим делителем, можно ограничиться рассмотрением целых положительных чисел. Пусть $a, b \in \mathbb{Z}^+$ и $(a, b) = d$. Если $a = a_1 d$ и $b = b_1 d$, то из леммы 1.6 следует, что $(a_1, b_1) = 1$. Пусть m — общее кратное чисел a и b . Поскольку $m = ak = a_1 dk$ и $b = b_1 d | m$, то $b_1 | a_1 k$.

Замечая, что $(a_1, b_1) = 1$, отсюда и из теоремы 1.7 получаем, что $b_1 | k$. Полагая $k = b_1 t$, получаем

$$m = (a_1 b_1 d) t = \frac{ab}{d} t. \quad (1.10)$$

Следовательно, в таком виде можно представить произвольное общее кратное чисел a и b . Обратное, любое число такого вида является общим кратным a и

b. Следовательно, наименьшее общее кратное получается при подстановке $t=1$, так что

$$[a, b] = \frac{ab}{(a, b)}. \quad (1.11)$$

В заключение рассмотрим способ определения наибольшего общего делителя и наименьшего общего кратного совокупности a_1, a_2, \dots, a_n более двух целых чисел. Имеем

$$(a_1, a_2) = d_2, (d_2, a_3) = d_3, \dots, (d_{n-1}, a_n) = d_n;$$

$$[a_1, a_2] = m_2, [m_2, a_3] = m_3, \dots, [m_{n-1}, a_n] = m_n.$$

Отсюда видно, что d_n и m_n являются соответственно наибольшим общим делителем и наименьшим общим кратным чисел a_1, \dots, a_n .

1.3. Разложение на простые сомножители

Произвольное целое положительное число a всегда имеет делители $\pm a$ и ± 1 . Если целое положительное число $p > 1$ не имеет других делителей кроме $\pm p$ и ± 1 , то оно называется простым. Положительные целые числа, не являющиеся простыми и большими 1, называют составленными.

Для перечисления всех простых чисел, не превосходящих заданное положительное целое число N , существует так называемый метод «решета Эратосфена», заключающийся в следующем. Выписываются числа

$$1, 2, \dots, N. \quad (1.12)$$

Число 1 не является простым, а поэтому из этой совокупности вычеркивается. Число 2 является простым и поэтому оставляется. Далее из совокупности (1.12) вычеркиваются все кратные числу 2. Наименьшее из оставшихся — число 3, оно простое, а поэтому также оставляется. Все кратные числу 3 из совокупности (1.12) вычеркиваются. Повторяя эти операции, можно получить все простые числа, содержащиеся в совокупности (1.12).

Теорема 1.8. Если a — целое положительное число, большее 1, то минимальный отличный от 1 положительный делитель a является простым числом.

Доказательство. Пусть $q > 1$ — минимальный отличный от 1 положительный делитель числа $a > 1$. Если q — составное число, что $q = q_1 q_2$, $q > q_1 > 1$, $q > q_2 > 1$, а следовательно, q_1 и q_2 являются делителями a , меньшими чем q . Получили противоречие.

Простые числа, являющиеся делителями целого числа a , называются простыми сомножителями a .

Теорема 1.9. Если p — простое число, то для произвольного целого числа a либо $p|a$, либо $(p, a) = 1$.

Доказательство. Так как положительными делителями p являются лишь числа p и 1, то либо $(p, a) = p$, либо $(p, a) = 1$. В первом случае $p|a$.

Теорема 1.10. Если p — простое число и $p|ab$, то либо $p|a$, либо $p|b$.

Доказательство. Если $p \nmid a$, то по предыдущей теореме $(p, a) = 1$. Так как $p \mid ab$, то из теоремы 1.7 следует, что $p \mid b$.

Следствие. Если p — простое число и $p \mid a_1 a_2 \dots a_n$, то $p \mid a_i$ при некотором $1 \leq i \leq n$.

Теорема 1.11. (Основная теорема элементарной теории чисел). Любое целое число, большее 1, однозначно разлагается в произведение простых чисел.

Доказательство. Рассмотрим произвольное целое число $a > 1$. Если p_1 — минимальный простой сомножитель a , то $a = p_1 a_1$. Если $a_1 > 1$ и p_2 — минимальный простой сомножитель a_1 , то $a_1 = p_2 a_2$. Этот процесс можно продолжить. Так как $a > a_1 > a_2 > \dots$, то в некоторый момент $a_n = 1$, и процесс закончится. При этом число $a_{n-1} = p_n$ будет простым.

Таким образом,

$$a = p_1 p_2 \dots p_n, \quad (1.13)$$

т. е. a разлагается в произведение простых чисел.

Пусть a другим способом разлагается в произведение простых чисел: $a = q_1 q_2 \dots q_m$. Тогда

$$p_1 p_2 \dots p_n = q_1 q_2 \dots q_m. \quad (1.14)$$

Так как простое число p_1 делит правую часть последнего равенства, то, согласно следствию из теоремы 1.10, p_1 также делит одно из чисел q_1, \dots, q_m . Без ограничения общности можно считать, что $p_1 \mid q_1$. Так как $(p_1, q_1) = p_1 > 1$, то, согласно теореме 1.10, $q_1 \mid p_1$. Отсюда и из следствия 2 теоремы 1.2 получаем, что $p_1 = q_1$. Разделив обе части равенства (1.14) на $p_1 = q_1$, приходим к равенству

$$p_2 \dots p_n = q_2 \dots q_m. \quad (1.15)$$

Повторяя эту процедуру достаточное число раз, получаем

$$n = m, \quad p_i = q_i \quad (i = 1, 2, \dots, n). \quad (1.16)$$

При разложении на простые сомножители (1.13) некоторые простые сомножители могут совпадать. Поэтому часто используется стандартное разложение

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}, \quad (1.17)$$

где $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{Z}^+$ и p_1, \dots, p_k — различные простые числа.

Следствие из теоремы 1.11. Если (1.17) — стандартное разложение числа $a > 1$ на простые сомножители, то стандартное разложение числа $d \in \mathbb{Z}^+$, $d \mid a$ имеет вид

$$d = p_1^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k}, \quad 0 \leq \beta_i \leq \alpha_i \quad (i = 1, \dots, k). \quad (1.18)$$

У п р а ж н е н и е 1.7. $588\,800 = 2^5 \cdot 3 \cdot 5^3 \cdot 7^2$.

У п р а ж н е н и е 1.8. Пусть (1.17) — стандартное разложение положительного целого числа $a > 1$ на простые множители. Тогда общее число всех различных положительных делителей числа a задается равенством

$$\tau(a) = (\alpha_1 + 1)(\alpha_2 + 1) \dots (\alpha_k + 1). \quad (1.19)$$

Решение. Положительные делители имеют вид (1.18), где в качестве β_i можно взять любое значение от 0 до α_i . Следовательно, β_1, \dots, β_k можно выбрать $(\alpha_1+1)(\alpha_2+1) \dots (\alpha_k+1)$ различными способами. Так как разложение на простые множители единственно, то все получающиеся делители будут различными.

Предположим, что задана некоторая функция $g(n)$, определенная для всех целых положительных чисел. Для произвольного положительного целого числа a сумма значений функции $g(d)$ во всех целых числах b , которые являются положительными делителями a , обозначается через

$$\sum_{d|a} g(d). \quad (1.20)$$

Например, если $a=6$, то указанная сумма равна $g(1)+g(2)+g(3)+g(6)$. Заметим, что для $\tau(a)$

$$\tau(a) = \sum_{d|a} 1, \quad (1.21)$$

что соответствует функции $g(n) \equiv 1$.

1.4. Сравнимость целых чисел

Рассмотрим остатки, которые получаются при делении целых чисел на число $m \in \mathbb{Z}^+$, называемое модулем. Целые числа a и b называются сравнимыми по модулю m , если $m|(a-b)$. Символически это записывается следующим образом:

$$a \equiv b \pmod{m}. \quad (1.22)$$

Деля числа a и b на m , получаем

$$a = mq_1 + r_1, \quad 0 \leq r_1 < m;$$

$$b = mq_2 + r_2, \quad 0 \leq r_2 < m.$$

Ясно, что a и b сравнимы по модулю m тогда и только тогда, когда $r_1 = r_2$. Отношение сравнимости удовлетворяет следующим законам.

A.1. (Рефлексивность): $a \equiv a$.

A.2. (Симметричность): если $a \equiv b$, то $b \equiv a$.

A.3. (Транзитивность): если $a \equiv b$ и $b \equiv c$, то $a \equiv c$.

С помощью определения сравнимости можно легко убедиться в справедливости следующих утверждений.

Теорема 1.12. К любой из частей сравнения можно прибавить число, кратное модулю. Обе части сравнения можно сложить с одним и тем же целым числом. Кроме того, обе части сравнения можно умножить на одно и то же целое число. Другими словами, если $a \equiv b \pmod{m}$ и $k \in \mathbb{Z}$, то

$$a + mk \equiv b, \quad a + k \equiv b + k, \quad ak \equiv bk \pmod{m}.$$

Теорема 1.13. Сравнения можно почленно складывать и умножать. А именно, если $a_1 \equiv b_1$, $a_2 \equiv b_2 \pmod{m}$, то

$$a_1 + a_2 \equiv b_1 + b_2, \quad a_1 a_2 \equiv b_1 b_2 \pmod{m}.$$

Сравнение остается справедливым при умножении обеих его частей на одно и то же число, однако при делении обеих частей

В общем случае, класс вычетов, содержащий произвольное целое число a , условимся обозначать через $(a)_m$.

Если взять по одному числу в каждом из m классов вычетов $(0)_m, (1)_m, \dots, (m-1)_m$, то полученная совокупность целых чисел будет называться полной системой вычетов. Например, полной системой вычетов является совокупность $\{0, 1, \dots, m-1\}$, которая получается, если выбрать в каждом из m смежных классов наименьший неотрицательный вычет.

Теорема 1.15. Если никакие два числа из заданного множества целых чисел a_1, \dots, a_m не сравнимы между собой по модулю m , то это множество является полной системой вычетов по модулю m .

Так как никакие два числа не принадлежат одному и тому же классу вычетов, то ясно, что эта совокупность из m целых чисел получена выбором ровно одного представителя из каждого из m классов вычетов.

Следствие из теоремы 1.15. Пусть $(c, m)=1$. Тогда, если $\{a_1, \dots, a_m\}$ — полная система вычетов по модулю m и b — целое число, то совокупность $\{ca_1+b, \dots, ca_m+b\}$ также будет полной системой вычетов по модулю m .

Доказательство. Пусть $i \neq j$. Если предположить, что $ca_i+b \equiv ca_j+b \pmod{m}$, то нетрудно прийти к противоречию. Действительно, прибавляя к обеим частям последнего сравнения $-b$, получаем $ca_i \equiv ca_j \pmod{m}$. Так как $(c, m)=1$, то, согласно теореме 1.14, $a_i \equiv a_j \pmod{m}$. Это противоречит тому, что $\{a_1, \dots, a_m\}$ является полной системой вычетов.

Рассмотрим сравнение первого порядка, содержащее неизвестное число x :

$$ax \equiv b \pmod{m}. \quad (1.27)$$

Это сравнение, вообще говоря, может не иметь решения; например, $2x \equiv 3$ или $2x-3 \equiv 0 \pmod{2}$. Так как $2x$ является четным числом при любом x , то $2x-3$ будет нечетным и, следовательно, никогда не кратно двум. Однако если сравнение имеет хотя бы одно решение x , то легко видеть, что решениями сравнения (1.27) будут все целые числа, сравнимые с x по модулю m . Из существования решения не следует, что совокупность всех решений совпадает в точности с одним классом вычетов. Например, рассмотрим сравнение $2x \equiv 0 \pmod{4}$, решения которого $x=0$ и $x=2$. В этом случае оба класса вычетов $\{\dots, -8, -4, 0, 4, 8, \dots\}$, $\{\dots, -6, -2, 2, 6, \dots\}$ дают решения рассматриваемого сравнения.

Следующая теорема относится к случаю, когда все решения сравнения составляют ровно один класс вычетов.

Теорема 1.16. Если $(a, m)=1$, то сравнение $ax \equiv b \pmod{m}$ имеет решение в целых числах; совокупность всех его решений составляет один класс вычетов.

Доказательство. Согласно лемме 1.1, существуют такие числа s и t , что $sa+tm=1$. Умножая это равенство на b , получаем $a(sb) + (bt)m = b$, т. е. $a(sb) \equiv b \pmod{m}$. Поэтому $x=sb$ является

решением рассматриваемого сравнения, при этом решениями оказываются также все целые числа, сравнимые с sb по модулю m .

Если числа x_1 и x_2 являются решениями, то $ax_1 \equiv ax_2 \pmod{m}$. Отсюда из равенства $(a, b) = 1$ и теоремы 1.14 следует, что $x_1 \equiv x_2 \pmod{m}$.

У п р а ж н е н и е 1.10. Решите сравнение $6x \equiv 1 \pmod{7}$.

Решение. Подставим вместо x в это сравнение числа $0, 1, \dots, 6$. Получим: $6 \cdot 0 = 0$, $6 \cdot 1 = 6$, $6 \cdot 2 = 12 \equiv 5$, $6 \cdot 3 = 18 \equiv 4$, $6 \cdot 4 = 24 \equiv 3$, $6 \cdot 5 = 30 \equiv 2$, $6 \cdot 6 = 36 \equiv 1 \pmod{7}$. Отсюда следует, что совокупностью решений является класс вычетов $(6)_7 = \{\dots, -8, -1, 6, 13, \dots\}$.

1.5. Области целостности и поля

До сих пор мы пользовались четырьмя арифметическими операциями над целыми числами, считая их свойства известными. В этом параграфе будут рассмотрены фундаментальные свойства арифметических операций.

В множестве целых чисел Z определены две операции: сложение « $+$ » и умножение « \cdot ». Вначале перечислим их основные свойства.

A.1. Ассоциативность сложения: $(x+y)+z=x+(y+z)$.

A.2. Коммутативность сложения: $x+y=y+x$.

Из ассоциативности сложения следует, что при сложении скобки можно расставлять произвольным образом или вообще их не писать, т. е. вместо записи $(x+y)+z=x+(y+z)$ пользоваться более краткой записью $x+y+z$. Кроме того, из коммутативности следует, что при сложении слагаемые можно менять местами.

A.3. Существование единичного по сложению элемента. Существует такое число θ , что $x+\theta=x$ для всех x .

В множестве целых чисел роль единичного по сложению элемента играет 0 . В единственности такого элемента можно убедиться следующим образом. Если числа θ, θ' являются единичными по сложению элементами, то $\theta=\theta+\theta'=\theta'+\theta=\theta'$. Это означает, что роль элемента θ в множестве целых чисел выполняет только число 0 . Поэтому далее вместо θ мы будем писать 0 .

A.4. Существование обратного по сложению элемента. Для каждого x существует элемент x' такой, что $x+x'=0$.

В множестве целых чисел роль x' играет $-x$. Можно показать, что обратный элемент определяется единственным образом. Действительно, если x', x'' — обратные элементы для x , то, прибавляя к обеим частям равенства $0=x+x'$ число x'' , получаем $x''=0+x''=(x'+x)+x''=x'+(x+x'')=x'+0=x'$.

Здесь мы воспользовались ассоциативностью, коммутативностью и свойствами единичного по сложению элемента 0 . Далее элемент, обратный к x , обозначается через $-x$.

Разность $x-y$ определяется формулой $x+(-y)$. Обратным к обратному элементу является исходный элемент, т. е. $-(-x)=x$. Описанными выше свойствами сложение обладает не только

в множестве целых чисел, но, и, например, в множестве всех рациональных чисел, в множестве всех действительных чисел, в множестве всех комплексных чисел. Вообще, множество G с определенной на нем операцией сложения «+», удовлетворяющей законам А.1—А.4, называется коммутативной группой (по сложению).

Рассмотрим множество классов вычетов по модулю

$$Z_m = \{(0)_m, (1)_m, \dots, (m-1)_m\}. \quad (1.28)$$

Класс вычетов $(i)_m$ представляет собой некоторое множество целых чисел. Однако при изучении множеств Z_m он рассматривается как один элемент. Если $a_1, a_2 \in (a)_m$, $b_1, b_2 \in (b)_m$, то согласно теореме 1.13, $a_1 + b_1 \equiv a_2 + b_2 \pmod{m}$, т. е. $(a_1 + b_1)_m = (a_2 + b_2)_m$. Следовательно, сложение

$$(a)_m + (b)_m = (a + b)_m \quad (1.29)$$

однозначно определяет операцию сложения в множестве Z_m .

Упражнение 1.11. Постройте таблицу сложения в Z_2 .

Решение.

	$(0)_2$	$(1)_2$
$(0)_2$	$(0)_2$	$(1)_2$
$(1)_2$	$(1)_2$	$(0)_2$

Упражнение 1.12. Постройте таблицу сложения в Z_3 .

Решение.

	$(0)_3$	$(1)_3$	$(2)_3$
$(0)_3$	$(0)_3$	$(1)_3$	$(2)_3$
$(1)_3$	$(1)_3$	$(2)_3$	$(0)_3$
$(2)_3$	$(2)_3$	$(0)_3$	$(1)_3$

Теорема 1.17. Z_m является группой по сложению.

Доказательство. Ассоциативность. Равенство

$$((a)_m + (b)_m) + (c)_m = (a)_m + ((b)_m + (c)_m)$$

является следствием того, что в силу ассоциативности сложения чисел как левая, так и правая его части равны $(a + b + c)_m$.

Коммутативность. Равенство $(a)_m + (b)_m = (b)_m + (a)_m$ следует из того, что обе его части в силу коммутативности сложения чисел совпадают с $(a + b)_m$. Очевидно, что единственным по сложению элементом в данном случае является класс вычетов $(0)_m$, а обратным элементом для класса $(a)_m$ — класс $(-a)_m$.

Умножение целых чисел обладает следующими свойствами.

А.5. Ассоциативность умножения: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

А.6. Коммутативность умножения: $a \cdot b = b \cdot a$.

А.7. Существование единичного элемента по умножению. Существует элемент $e \neq 0$ такой, что $xe = x$ для всех x .

В множестве целых чисел единичным по умножению элементом является число 1. Как и для сложения, существует только

один единичный элемент по умножению. В множестве целых чисел обратный элемент по умножению, как правило, отсутствует. Далее единичный элемент по умножению будет обозначаться через 1 и точка \cdot , обозначающая умножение, будет опускаться.

А.8. Дистрибутивность. Для производных x, y, z , $x(y+z) = xy+xz$.

Из дистрибутивности следуют важные соотношения:

$$x0=0 \text{ для произвольного } x, \quad (1.30)$$

$$x(-y) = -(xy), \quad (1.31)$$

$$x(y-z) = xy-xz. \quad (1.32)$$

Чтобы получить (1.30), к обеим частям равенства $x0 = x(0+0) = =x0+x0$ следует прибавить $-(x0)$. При этом

$$0 = x0 + (-(x0)) = x0 + (x0 + (-(x0))) = x0 + 0 = x0.$$

Далее из равенства $xy+x(-y) = x(y+(-y)) = x0=0$ следует справедливость (1.31). Наконец, из равенств $x(y-z) = x(y+ +(-z)) = xy+x(-z) = xy-xz$ вытекает (1.32).

Так же, как и в случае сложения, можно показать, что соотношение

$$(a)_m(b)_m = (ab)_m \quad (1.33)$$

однозначно определяет умножение в Z_m .

Упражнение 1.13. Постройте таблицы умножения в Z_2 и Z_3 .

Решение.

	$(0)_2$	$(1)_2$		$(0)_3$	$(1)_3$	$(2)_3$
$(0)_2$	$(0)_2$	$(0)_2$	$(0)_3$	$(0)_3$	$(0)_3$	$(0)_3$
$(1)_2$	$(0)_2$	$(1)_2$	$(1)_3$	$(0)_3$	$(1)_3$	$(2)_3$
			$(2)_3$	$(0)_3$	$(2)_3$	$(1)_3$

Упражнение 1.14. Покажите, что умножение в $Z_m(m>1)$ удовлетворяет законам А.5—А.8.

Решение аналогично проверке соответствующих свойств для сложения. Единичным по умножению элементом является класс вычетов $(1)_m \neq (0)_m$.

Умножение в множестве целых чисел Z обладает, кроме того, следующими свойствами.

А.9. Правило сокращения. Если $x \neq 0$ и $xy=xz$, то $y=z$.

Мы уже неоднократно пользовались тем, что обе части равенства можно разделить (сократить) на их отличный от нуля общий множитель. Правило сокращения А.9 эквивалентно следующему.

А.9'. Правило сокращения. Если $x \neq 0$ и $y \neq 0$, то $xy \neq 0$.

Действительно, если допустить, что $x \neq 0$, $y \neq 0$, но $xy=0$, то $xy = x0$.

Отсюда и из А.9 следует, что $y=0$. Получили противоречие. Таким образом, из А.9 следует А.9'. Наоборот, если выполняется А.9', то из равенства $xy=xz$ следует, что $0=xy-xz=x(y-z)$. Таким образом, если $x \neq 0$, то $y-z=0$ и $y=z$.

Если на множестве G определены операции сложения и умножения, удовлетворяющие законам А.1—А.9, то G называется областью целостности.

Если m — составное число, то его можно записать, в виде $m=ab$, причем $m > a > 1$ и $m > b > 1$. Следовательно, $(a)_m \neq (0)_m$, $(b)_m \neq (0)_m$, но

$$(a)_m (b)_m = (m)_m = (0)_m.$$

т. е. произведение ненулевых элементов равно нулевому. Это означает, что в данном случае Z_m для составных m не является областью целостности.

Упражнение 1.15. Покажите, что если p — простое число, то Z_p — область целостности.

Решение. Пусть $(a)_p \neq (0)_p$ и $(b)_p \neq (0)_p$. Предположим, что ни a , ни b не являются делителями p , а $(a)_p (b)_p = (ab)_p = (0)_p$. Тогда $p|ab$ и, согласно теореме 1.10, либо $p|a$, либо $p|b$. Получили противоречие.

В множестве целых чисел Z обратный по умножению элемент может отсутствовать. Однако он всегда существует как в множестве рациональных чисел, так и в множестве действительных чисел.

А.10. Существование обратного по умножению элемента. Для каждого $x \neq 0$ существует элемент x' такой, что $xx'=1$.

Обратный элемент x' определяется однозначно; будем обозначать его через x^{-1} . Из А.10 автоматически следует правило сокращения А.9. Действительно, если $x \neq 0$ и $xy=xz$, то, умножая последнее равенство на x^{-1} , получаем

$$x^{-1}(xy) = (x^{-1}x)y = 1y = y = x^{-1}(xz) = (x^{-1}x)z = 1z = z.$$

Если на множестве G определены операции сложения и умножения, удовлетворяющие законам А.1—А.8 и А.10, то G называется полем. В поле всегда возможно деление. А именно, если $y \neq 0$, то результат деления x на y определяется как xy^{-1} . Кроме поля рациональных чисел и поля действительных чисел существуют также и другие поля.

Теорема 1.18. Если p — простое число, то Z_p является полем.

Доказательство. Если $(a)_p \neq (0)_p$, то $p \nmid a$. Отсюда и из теоремы 1.9 следует, что $(p, a)=1$. Согласно теореме 1.12, сравнение $ax \equiv 1 \pmod{p}$ имеет решение $x=b$. Так как $(a)_p (b)_p = (ab)_p = (1)_p$, то $(b)_p$ является обратным по умножению элементом для $(a)_p$.

Область целостности и поле, содержащие конечное число элементов, называются соответственно конечной областью целостности и конечным полем. В действительности, можно доказать, что конечная область целостности всегда поле [1, 2]. Заметим также, что Z_p является конечной областью целостности из p элемен-

тов. Если G — поле, то для каждого элемента совокупности $G - \{0\}$, получающейся из G удалением единичного по сложению элемента 0, существует обратный элемент по умножению, а следовательно, $G - \{0\}$ является группой по умножению.

Для обозначения элементов конечного поля Z_p вместо $(0)_p$, $(1)_p$, $(2)_p$, ..., $(p-1)_p$ можно просто использовать символы 0, 1, ..., $p-1$. Наиболее широко используемое множество Z_2 является полем, состоящим из единичного по сложению элемента 0 и единичного по умножению элемента 1. Согласно упражнениям 1.11 и 1.13, сложение и умножение в этом поле задаются следующими формулами: $0+0=0$, $0+1=1+0=1$, $1+1=0$, $0 \cdot 0=1 \cdot 0=0$, $0 \cdot 1=1$, $1 \cdot 1=1$. Все эти равенства, за исключением равенства $1+1=0$, выполняются для единичного по сложению элемента 0 и единичного по умножению элемента 1 произвольного поля. Если допустить, что в поле, которое не содержит никаких других элементов, кроме 0 и 1, сложение таково, что $1+1=1$, то $1+0=1$ и, следовательно, для элемента 1 не существует обратного по сложению элемента. Таким образом, в поле из двух элементов должно выполняться равенство $1+1=0$. Это означает, что Z_2 — единственное поле, состоящее только из единичного по сложению элемента 0 и единичного по умножению элемента 1.

Задачи

1.1. Найдите следующие наибольшие общие делители:
а) (6188, 4709); б) (81719, 52003, 30107, 33649).

1.2. Докажите, что если $(a, b)=1$, то $(ac, b)=(c, b)$.

1.3. Покажите, что если $(m_1, m_2)=1$, то система сравнений $x \equiv b_1 \pmod{m_1}$, $x \equiv b_2 \pmod{m_2}$ имеет решение и, более того, множество всех ее решений совпадает с одним из классов вычетов по модулю $m_1 m_2$.

1.4. Покажите, что если некоторое множество целых чисел S не пусто и замкнуто относительно вычитания, то оно будет замкнутым и относительно сложения.

1.5. Постройте таблицы сложения и умножения для конечного поля Z_5 .

Глава 2

Размещения, сочетания, принцип включения — исключения

2.1. Размещения без повторений

Способ расположения в определенном порядке некоторого числа элементов из заданного множества S или, что то же самое, способ выбора этих элементов из S , когда существенна последо-

вательность выбора элементов, называется *размещением*. Если же последовательность выбора элементов несущественна, то способ выбора называют *сочетанием*. При этом может допускаться или не допускаться повторный выбор одних и тех же элементов. Например, когда повторение элементов не допускается, из трех элементов, a, b, c , два элемента могут быть выбраны шестью различными способами, если учитывать последовательность выбора элементов ab, ac, ba, bc, ca, cb , и тремя различными способами, если не учитывать последовательность выбора элементов $a, b; a, c; b, c$. Если же повторение элементов допустимо, то имеется девять размещений $aa, ab, ac, ba, bb, bc, ca, cb, cc$ и шесть сочетаний $a, a; a, b; a, c; b, b; b, c; c, c$.

Общее число размещений без повторения из n различных элементов по r обозначается через ${}_nP_r$. Нетрудно видеть, что

$${}_nP_r = n(n-1)\dots(n-r+1), \quad n \geq r \geq 1. \quad (2.1)$$

Так как повторение элементов не допускается, то всегда $n \geq r$. Будем считать, что при $r=0$ имеется одно размещение (элементы вообще не выбираются), т. е. положим

$${}_nP_0 = 1. \quad (2.1)'$$

Размещение r элементов можно представлять себе как заполнение некоторых r позиций элементами заданного множества. При этом первую позицию можно заполнить n различными способами. После того, как 1-я позиция заполнена, элемент для заполнения 2-й позиции можно выбрать $(n-1)$ способами. Если этот процесс продолжить, то после заполнения позиций с 1-й по $(r-1)$ -ю будет иметься $(n-r+1)$ способов заполнения последней, r -й позиции. Перемножая эти числа, мы получаем формулу (2.1). Произведение, стоящее в правой части (2.1), в этом и последующих параграфах обозначается через $n^{(r)}$. В частном случае, когда $n=r$, имеем

$${}_nP_n = n(n-1)\dots 3 \cdot 2 \cdot 1 = n! \quad (2.2)$$

Число различных сочетаний без повторения n различных предметов будем обозначать через ${}_nC_r$. Заметим, что каждому сочетанию соответствует $r!$ различных способов упорядочения входящих в него элементов. Если такое упорядочение провести для всех сочетаний, то мы получим все размещения без повторения из n различных элементов по r и, следовательно,

$${}_nC_r = \frac{{}_nP_r}{r!} = \frac{n(n-1)\dots(n-r+1)}{r!}, \quad n \geq r \geq 1. \quad (2.3)$$

Будем считать, что при $r=0$ имеется одна возможность для выбора элементов — вообще их не выбирать, так что

$${}_nC_0 = 1. \quad (2.3)'$$

Выражения, стоящие в правой части (2.3), называют биномиаль-

ными коэффициентами и обозначают часто через $\binom{n}{r}$. При этом

$$\binom{n}{r} = \frac{n(n-1)\dots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!}, \quad n \geq r \geq 0, \quad (2.4)$$

$$\text{где } 0! = 1. \quad (2.5)$$

Из формулы (2.4) видно, что

$$\binom{n}{r} = \binom{n}{n-r}, \quad 0 \leq r \leq n. \quad (2.6)$$

Равенство (2.4) можно получить, не пользуясь формулой (2.4), с помощью следующего рассуждения. Так как каждому сочетанию из n различных элементов по r соответствует сочетание, состоящее из $n-r$ оставшихся элементов, то ясно, что число сочетаний из n элементов по r равно числу сочетаний из n элементов по $(n-r)$. С помощью аналогичных рассуждений можно убедиться в справедливости следующей формулы:

$$\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}, \quad 1 \leq r \leq n. \quad (2.7)$$

Действительно, отметим один из $n+1$ различных элементов. Выбор r элементов, включающий отмеченный элемент, эквивалентен выбору $(r-1)$ элементов из n неотмеченных элементов. Число таких способов выбора соответствует первому слагаемому в левой части (2.7). В то же время число способов выбора r элементов, не включая в их число отмеченный элемент, равно числу сочетаний из n различных не отмеченных элементов по r ; это дает второе слагаемое в правой части (2.7).

Общий член разложения произведения

$$(x+y)^n = \underbrace{(x+y)(x+y)\dots(x+y)}_n$$

имеет вид $c_r x^r y^{n-r}$. Заметим, что в правой части этого равенства произведение $x^r y^{n-r}$ можно получить, взяв элемент x из r сомножителей $x+y$ (которые можно выбрать произвольным образом из n сомножителей, входящих в произведение) и элемент y из оставшихся $(n-r)$ сомножителей. Следовательно, $c_r = {}_n C_r$, и мы получаем следующую важную теорему о разложении бинома:

$$(x+y)^n = \sum_{r=0}^n \binom{n}{r} x^r y^{n-r}. \quad (2.8)$$

Полагая $x=y=1$ и $x=-1, y=1$, можно получить следующие формулы:

$$\sum_{i=0}^n \binom{n}{i} = 2^n, \quad \sum_{r=0}^n (-1)^r \binom{n}{r} = 0. \quad (2.9)$$

С помощью (2.7) легко построить таблицу биномиальных коэффициентов. Эта таблица называется *треугольником Паскаля* и

может быть построена строка за строкой сверху вниз с помощью следующего правила: каждый элемент треугольника есть сумма двух его элементов, расположенных выше него слева и справа. Поперечным сечением пирамиды Паскаля является последовательность коэффициентов из разложения бинома (2.8), т. е. последовательность биномиальных коэффициентов.

1 1 по вертикали $n=1,2,3,\dots$
 1 2 1 по горизонтали $r=0,1,2,\dots$
 1 3 3 1
 1 4 6 4 1

У п р а ж н е н и е 2.1. Докажите равенство $\sum_{r=1}^n r(-1)^r \binom{n}{r} = 0$, $n > 1$.

Решение 1. Из (2.4) следует, что $r \binom{n}{r} = n \binom{n-1}{r-1}$. Полагая $r-1=k$, получаем

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

Согласно (2.9), последнее выражение равно нулю.

Решение 2. Полагая в (2.8) $y=1$ и $x=-w$, получим

$$(1-w)^n = \sum_{r=0}^n (-1)^r \binom{n}{r} w^r.$$

Дифференцируя это равенство по w , имеем

$$-n(1-w)^{n-1} = \sum_{r=1}^n r(-1)^r \binom{n}{r} w^{r-1}$$

и, полагая далее $w=1$, получаем нужное равенство.

У п р а ж н е н и е 2.2. Найдите число способов, которыми можно расставить n человек по кругу, т. е. число кольцевых последовательностей. При расположении по кругу абсолютное положение каждого человека не важно; во внимание следует принимать только их относительное расположение друг относительно друга. Например, три способа расстановки, показанные на рис. 2.1, приводят к одинаковым кольцевым последовательностям.

Решение. Попытаемся разместить n человек на прямой линии и далее замкнуть концы этой последовательности в кольцо. Число способов размещения n

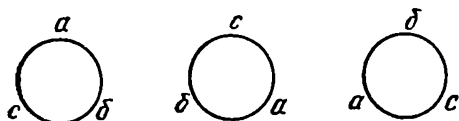


Рис. 2.1. Кольцевые последовательности и их представление на окружности

человек на прямой равно $n!$. Однако при этом следующие n последовательностей:

$$a_1 a_2 \dots a_{n-1} a_n, a_2 a_3 \dots a_n a_1, \dots, a_n a_1 \dots a_{n-2} a_{n-1},$$

каждая из которых получается в результате сдвига влево предыдущей последовательности и переноса крайнего левого элемента на освободившееся место крайнего правого элемента, приводят к одинаковым кольцевым последовательностям. Следовательно, общее число кольцевых последовательностей равно

$$n!/n = (n-1)!.$$

2.2. Размещения и сочетания с повторениями

Общее число размещений с повторениями r элементов, взятых из совокупности n различных элементов, равно

$$n^r, \quad r \geq 0. \quad (2.10)$$

В отличие от выражения (2.1) для размещения без повторений, здесь r может превосходить n . Будем считать, что при $r=0$ имеется один способ размещения элементов, т. е. $n^0=1$. При $r>0$ число способов заполнения 1-й позиции равно n , 2-ю позицию также можно заполнить n способами и т. д. Следовательно, общее число размещений равно n^r .

Буквы английского алфавита и другие символы, используемые при передаче информации, часто представляются в виде двоичных последовательностей определенной длины. Например, некоторый символ может быть представлен с помощью последовательности

$$a = (a_1, a_2, \dots, a_n), \quad (2.11)$$

компоненты которой равны 0 или 1. Максимальное число знаков, которые могут быть представлены с помощью l двоичных символов (l бит), равно числу размещений с повторениями l элементов из множества, содержащего два различных элемента (0 и 1), т. е.

$$2^l. \quad (2.12)$$

Найдем далее число сочетаний с повторениями из n различных элементов по r . Пусть n различных элементов — это символы a_1, a_2, \dots, a_n . Число символов a_1 , входящих в сочетание из r элементов с повторениями, обозначим через x_1 , аналогично число символов a_2 — через x_2, \dots , число символов a_n — через x_n . При этом

$$\left. \begin{aligned} x_1 + x_2 + \dots + x_n &= r; \\ x_1 \geq 0, x_2 \geq 0, \dots, x_n &\geq 0. \end{aligned} \right\} \quad (2.13)$$

Заметим, что каждый символ a_i , вообще говоря, не обязан входить в каждое сочетание. При этом мы можем к уже имеющимся в сочетании символам добавить один символ a_1 , один символ a_2, \dots , один символ a_n и, таким образом, получить новое

сочетание со следующим свойством. Число y_i символов a_i в новом сочетании равно $x_i + 1$ и

$$\left. \begin{aligned} y_1 + y_2 + \dots + y_n &= n + r; \\ y_1 > 0, y_2 > 0, \dots, y_n > 0. \end{aligned} \right\} \quad (2.14)$$

Если одинаковые элементы, входящие в сочетание, объединить в группы и расположить группы элементов друг за другом, помещая между группами разделительную вертикальную черту, получится следующая последовательность:

$$\underbrace{a_1 \dots a_1}_{y_1} | \underbrace{a_2 \dots a_2}_{y_2} | \dots | \underbrace{a_n \dots a_n}_{y_n}.$$

Так как каждый символ входит в эту последовательность, по крайней мере, 1 раз, то число разделителей в последовательности равно $n - 1$. Число позиций в последовательности, в которых может находиться разделитель, $n + r - 1$, так как последовательность содержит $n + r$ символов. Общее число различных способов, которыми можно разместить разделители, представляет собой искомое число сочетаний:

$$\binom{n + r - 1}{n - 1} = \binom{n + r - 1}{r}. \quad (2.15)$$

Таким образом, число различных целочисленных решений уравнения (2.13) определяется формулой (2.15); число различных целочисленных решений уравнения (2.14) также определяется формулой (2.15).

Упражнение 2.3. Найти число сочетаний из пяти различных элементов по 3.

Решение. Если повторение элементов не разрешается, то

$$\binom{5}{3} = \frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = 10.$$

Если разрешить повторение элементов, то

$$\binom{7}{3} = \frac{5 \cdot 6 \cdot 7}{3 \cdot 2 \cdot 1} = 35.$$

Упражнение 2.4. Определить число способов, которыми можно выбрать l различных целых чисел из совокупности $1, 2, \dots, k$, так, чтобы в выборке не было ни одной пары чисел, следующих друг за другом.

Решение. Пусть w_1, w_2, \dots, w_l — это l чисел, выбранных в соответствии с условиями задачи и расположенных в порядке их возрастания, т. е. таких, что

$$1 \leq w_1 < w_2 < \dots < w_l \leq k \text{ и } w_{i+1} - w_i \geq 2, \quad i = 1, \dots, l - 1.$$

Положим $z_1 = w_1 - 1$, $z_{l+1} = k - w_l$, $z_j = w_j - w_{j-1}$ ($j = 2, \dots, l$).

При этом $z_1 \geq 0$, $z_{l+1} \geq 0$, $z_j \geq 2$ ($j = 2, \dots, l$), $z_1 + z_2 + \dots + z_l + z_{l+1} = k - 1$.

Далее, полагая $x_1 = z_1$, $x_{l+1} = z_{l+1}$, $x_j = z_j - 2$ ($j = 2, \dots, l$),

получим $x_1 \geq 0, x_2 \geq 0, \dots, x_l \geq 0, x_{l+1} \geq 0$,

$$x_1 + x_2 + \dots + x_l + x_{l+1} = (k-1) - 2(l-1) = k - 2l + 1.$$

Так как искомое число сочетаний равно числу различных целочисленных (неотрицательных) решений последнего уравнения и все рассматриваемые сочетания имеют смысл при

$$k - 2l + 1 \geq 0, \text{ т.е. } k - l + 1 \geq l,$$

то, согласно (2.15), это число сочетаний

$$\binom{k-l+1}{l}.$$

Вернемся к задаче передачи информации и рассмотрим множество V из всех 2^l двоичных последовательностей (векторов), имеющих вид (2.11). Введем в множестве V операцию сложения $+$, определив для последовательностей

$$a = (a_1, a_2, \dots, a_l), \quad b = (b_1, b_2, \dots, b_l) \quad (2.16)$$

их сумму с помощью соотношения

$$a + b = (a_1 + b_1, a_2 + b_2, \dots, a_l + b_l), \quad (2.17)$$

где $a_i + b_i$ — это сумма в поле Z_2 , описанном в предыдущем параграфе.

Упражнение 2.5. Покажите, что V является группой относительно введенной операции сложения.

Решение. Нужно доказать, что выполняются аксиомы А.1 — А.4 из § 1.5. Ассоциативность (А.1) и коммутативность (А.2) сложения в V следует соответственно из ассоциативности и коммутативности сложения в Z_2 . Кроме того, легко видеть, что последовательность

$$0 = (0, 0, \dots, 0) \quad (2.18)$$

— единичный элемент по сложению. Столь же очевидно, что обратным элементом по сложению для элемента a будет он сам. Действительно,

$$a + a = 0. \quad (2.19)$$

Так как $-a = a$, то вычитание совпадает со сложением и

$$a - b = a + b. \quad (2.20)$$

Расстояние Хэмминга $d(a, b)$ между кодовыми словами определяется как число позиций i таких, что $a_i \neq b_i$. При построении вектора $a - b = a + b$ пары одинаковых двоичных символов (a_i, b_i) , $a_i = b_i$ будут давать 0 в сумме $a + b$, а пары различных двоичных символов (a_i, b_i) , $a_i \neq b_i$, будут давать 1. Например, $(0, 1, 0, 1) + (1, 1, 1, 0) = (1, 0, 1, 1)$. Следовательно, расстояние Хэмминга $d(a, b)$ можно определить как число компонент вектора $a - b = a + b$, которые равны 1.

Введенное расстояние Хэмминга удовлетворяет следующим условиям.

А.1. $d(a, b) \geq 0$; $d(a, b) = 0$ тогда и только тогда, когда $a = b$.

А.2. $d(a, b) = d(b, a)$.

А.3. $d(a, b) + d(b, c) \geq d(a, c)$.

Справедливость утверждений А.1 и А.2 для расстояния Хэмминга почти очевидна, а утверждение А.3, которое называют так-

же неравенством треугольника, можно доказать следующим образом. Если i -е двоичные символы векторов a и c равны, т. е. $a_i = c_i$, то их вклад в расстояние $d(a, c)$ равен нулю. Если при этом $b_i \neq a_i = c_i$, то в левой части вклад i -х компонент в расстояния $d(a, b)$ и $d(b, c)$ равен 1. Если $a_i = b_i = c_i$, то в левой части i -е компоненты также дадут нуль. Если $a_i \neq c_i$, вклад этих компонент в расстояние $d(a, c)$ равен 1. При этом i -е компоненты дают такой же единичный вклад в одно из расстояний $d(a, b)$ или $d(b, c)$ в зависимости от того, совпадает ли b_i с a_i или с c_i . Поэтому правая часть неравенства треугольника (A.3) не может быть больше левой.

В общем случае расстоянием называют функцию, удовлетворяющую аксиомам A.1—A.3. Этим аксиомам, естественно, удовлетворяет расстояние между двумя точками, (x_1, y_1) и (x_2, y_2) , на плоскости: $[(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}$. При этом неравенство треугольника соответствует известному утверждению о том, что сумма длин двух сторон треугольника не меньше длины третьей стороны.

При наличии l двоичных символов (бит) можно построить 2^l векторов (кодовых слов), однако часто нужны не все 2^l слов. Если использовать только часть кодовых слов, то с помощью проверок на четность и других приемов можно обнаружить ошибки и даже их исправлять. Предположим, что из 2^l слов используются только p слов:

$$A = \{a^{(1)}, a^{(2)}, \dots, a^{(p)}\}. \quad (2.21)$$

Если эти слова таковы, что при всех $i \neq j$

$$d(a^{(i)}, a^{(j)}) \geq 2r + 1, \quad r \geq 0, \quad (2.22)$$

то можно исправлять любые ошибки кратности r . Предположим, что в некоторых r компонентах слова при передаче возникли ошибки, в результате которых символы 0 были приняты как 1, а символы 1 — как 0. Получившееся в результате слов b находится на расстоянии r от слова a . Следовательно, если считать, что при передаче ошибки могут возникать не более чем в r компонентах, то совокупность всех слов, которые могут получиться из $a^{(i)}$, представляет собой множество

$$A_i = \{x \in V | d(a^{(i)}, x) \leq r\}. \quad (2.23)$$

Из неравенства треугольника и условия (2.22) следует, что

$$A_i \cap A_j = \emptyset, \quad (2.24)$$

если только $i \neq j$. Действительно, допустим, что существует слово $w \in A_i \cap A_j$. Тогда $d(a^{(i)}, w) \leq r$, $d(w, a^{(j)}) \leq r$ и, согласно неравенству треугольника,

$$d(a^{(i)}, a^{(j)}) \leq d(a^{(i)}, w) + d(w, a^{(j)}) \leq 2r,$$

что приводит к противоречию. Множество A_i называется областью декодирования для слова $a^{(i)}$. Так как области декодирования не пересекаются, то при приеме слова из области декодиро-

вания A_i можно считать, что передавалось слово $a^{(i)}$. Если при этом число ошибок не превосходит r , то при приеме всегда будет приниматься правильное решение.

Так как число элементов $|A_i|$ в множестве A_i равно числу способов выбора некоторых k ($0 \leq k \leq r$) компонент слова $a^{(i)}$, то

$$|A_i| = \binom{l}{0} + \binom{l}{1} + \dots + \binom{l}{r}. \quad (2.25)$$

Далее, так как $|V| = 2^l$, то

$$p \sum_{k=0}^r \binom{l}{k} \leq 2^l, \text{ т. е. } p \leq 2^l / \left[\sum_{k=0}^r \binom{l}{k} \right]. \quad (2.26)$$

Последнее неравенство называется границей Хэмминга и дает оценку сверху для числа p кодовых слов, которые можно использовать для передачи, если их длина равна l бит и нужно исправлять все r -кратные ошибки.

2.3. Формула обращения Мебиуса

Функция Мебиуса $\mu(n)$ определяется для целых положительных чисел n следующим образом. Если $n=1$, то $\mu(1)=1$. Пусть, далее, число $n>1$ разлагается в произведение степеней простых сомножителей:

$$n = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}. \quad (2.27)$$

Тогда $\mu(n)$ определяется с помощью соотношения

$$\mu(n) = \begin{cases} (-1)^k, & \text{если } e_1 = e_2 = \dots = e_k = 1; \\ 0 & \text{в противном случае.} \end{cases} \quad (2.28)$$

Упражнение 2.6.

n	1	2	3	4	5	6	7	8	9	10
$\mu(n)$	1	-1	-1	0	-1	1	-1	0	0	1

Теорема 2.1.

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{при } n=1; \\ 0 & \text{при } n>1. \end{cases} \quad (2.29)$$

Сумма здесь (см. обозначения в § 1.3) берется по всем положительным делителям числа n .

Доказательство. При $n=1$ справедливость равенства (2.29) очевидна. Пусть $n>1$ представлено в виде произведения степеней простых сомножителей (2.27). Положим $n' = p_1 p_2 \dots p_k$.

Так как для любого делителя

$$d = p_1^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k}$$

числа n , не являющегося делителем n' , одно из чисел β_i больше или равно 2, то по определению функции Мебиуса $\mu(d) = 0$.

Следовательно,

$$\sum_{d|n} \mu(d) = \sum_{d|n'} \mu(d).$$

Так как делители числа n' — произведения из r ($0 \leq r \leq k$) различных простых чисел, взятых из совокупности k различных простых чисел p_1, \dots, p_k , то число делителей, которые являются произведениями r попарно различных простых чисел, равно $\binom{k}{r}$ и значение функции Мебиуса для каждого из них равно $(-1)^r$. Отсюда и из равенства (2.9)

$$\sum_{d|n'} \mu(d) = \sum_{r=0}^k (-1)^r \binom{k}{r} = 0.$$

Теорема 2.2. (Теорема обращения Мебиуса). Пусть $f(n)$ и $g(n)$ — функции, определенные для всех положительных целых чисел n . Если при любом целом положительном n

$$f(n) = \sum_{d|n} g(d), \quad (2.30)$$

$$\text{то } g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right). \quad (2.31)$$

Верно и обратное утверждение, а именно, из равенства (2.31) следует равенство (2.30).

Доказательство. Для всех делителей d числа n имеет место равенство

$$f\left(\frac{n}{d}\right) = \sum_{d'|n/d} g(d').$$

Поэтому

$$\sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \sum_{d'|n/d} g(d').$$

Так как $n = d'dn_1$, то при фиксированном d' сомножитель d пробегает все делители числа n/d' . Меняя порядок суммирования, получаем

$$\sum_{d|n} \mu(d) \sum_{d'|n/d} g(d') = \sum_{d'|n} g(d') \sum_{d|n/d'} \mu(d).$$

Согласно теореме 2.1 при $n/d' > 1$, сумма значений $\mu(d)$ равна нулю, а поэтому правая часть последнего равенства равна $g(n)$.

Для того чтобы получить (2.30) из (2.31), следует положить $d = d'k$. При этом

$$\sum_{d|n} g(d) = \sum_{d|n} \sum_{d'|d} \mu(d') f\left(\frac{n}{d'}\right) = \sum_{k|n} f(k) \sum_{d'|n/k} \mu(d') = f(n).$$

Как мы видели при решении упражнения 2.2, число кольцевых последовательностей из n элементов, которые выбираются без повторения из совокупности n попарно различных элементов,

равно $(n-1)!$. Найдем общее число $T(n)$ кольцевых последовательностей из n элементов, выбираемых с повторениями, из совокупности r попарно различных элементов.

Расположим символы a_1, a_2, \dots, a_n на окружности по часовой стрелке, при этом за символом a_n будет следовать символ a_1 . Следующие последовательности из n символов, расположенных на прямой:

$$\begin{aligned} a_1 a_2 \dots a_{n-1} a_n, a_2 a_3 \dots a_n a_1, \dots, \\ a_n a_1 \dots a_{n-2} a_{n-1}, \end{aligned} \quad (2.32)$$

соответствуют одной и той же кольцевой последовательности. Однако, вообще говоря, они могут быть совпадающими последовательностями из n символов на прямой. Например, если один и тот же символ повторяется n раз, все приведенные последовательности состоят из одинаковых символов и, следовательно, представляют одну и ту же последовательность на прямой.

Назовем периодом кольцевой последовательности минимальное целое число d такое, что последовательность a_1, a_2, \dots, a_n может быть получена (n/d) -кратным повторением ее начала a_1, \dots, a_d . Например, последовательности $abababab$, $abbcabbc$ длины 8 имеют соответственно периоды 2 и 4. Ясно, что если период равен d , то из n последовательностей на прямой (2.32) только d являются различными. Пусть $M(d)$ — общее число кольцевых последовательностей длины d (т. е. наборов d элементов, расположенных на окружности) с периодом, также равным d . При этом число различных последовательностей на прямой, соответствующих кольцевым последовательностям длины n с периодом d , будет равно $dM(d)$. Общее число последовательностей на прямой определяется формулой (2.10).

Поэтому

$$r^n = \sum_{d|n} d M(d).$$

Это равенство можно записать в виде (2.30), если положить $f(n) = r^n$ и $g(n) = nM(n)$. Но тогда по теореме обращения Мебиуса

$$n M(n) = \sum_{d|n} \mu(d) r^{n/d}$$

и, следовательно,

$$M(n) = \frac{1}{n} \sum_{d|n} \mu(d) r^{n/d}. \quad (2.33)$$

Так как $M(n)$ — это число кольцевых последовательностей длины n с периодом n , то общее число кольцевых последовательностей длины n

$$T(n) = \sum_{d|n} M(d). \quad (2.34)$$

Упражнение 2.7. Покажите, что если $\tau(n)$ — общее число различных положительных делителей целого положительного числа n , то

$$\sum_{d|n} \mu(d) \tau\left(\frac{n}{d}\right) = 1.$$

Решение. Утверждение следует из (1.21) и формулы обращения.

2.4. Принцип включения — исключения

Рассмотрим свойства $p(1), \dots, p(n)$, которыми обладают элементы множества S , состоящего из N элементов. Предположим, что для каждого элемента можно однозначно определить, обладает он свойством $p(i)$ или нет. Число элементов, обладающих свойствами $p(i_1), \dots, p(i_r)$, обозначим через

$$N_{i_1 \dots i_r}. \quad (2.35)$$

Следует заметить, что здесь не рассматриваются другие свойства, т. е. допускается, что некоторые элементы, быть может, обладают и другими свойствами. Число элементов, не обладающих ни одним из рассматриваемых свойств, обозначим через $N(0)$. Число элементов, обладающих ровно r ($1 \leq r \leq n$) свойствами из n возможных, будем обозначать через $N(r)$.

В самом простом случае $n=2$ почти очевидно, что

$$N(0) = N - (N_1 + N_2) + N_{12}. \quad (2.36)$$

Действительно, каждый элемент, обладающий свойствами 1 и 2 одновременно, учитывается и в N_1 и в N_2 , а потому при вычислении разности $N - (N_1 + N_2)$ он вычитается из N дважды. Следовательно, если к этой разности добавить N_{12} , то получится $N(0)$ — число элементов, не обладающих ни одним из двух рассматриваемых свойств. Обобщением этого правила является следующий принцип включения — исключения.

Теорема 2.3. (Принцип включения — исключения).

$$N(0) = N - \sum_i N_i + \sum_{i_1 < i_2} N_{i_1 i_2} - \dots + (-1)^s \sum_{i_1 < \dots < i_s} N_{i_1 \dots i_s} + \dots + (-1)^n N_{12 \dots n}. \quad (2.37)$$

Доказательство. Элементы, не обладающие ни одним из свойств, входят в правую часть последнего равенства ровно по 1 разу, а именно в первое слагаемое. Рассмотрим элементы, обладающие ровно r свойствами $p(j_1), \dots, p(j_r)$ и суммы вида

$$(-1)^s \sum_{i_1 < \dots < i_s} N_{i_1 \dots i_s}, \quad s \leq r.$$

Так как совокупность индексов i_1, \dots, i_s из множества j_1, \dots, j_r можно выбрать $\binom{r}{s}$ способами, то вклад каждого из рассматриваемых элементов в правую часть (2.37) будет

$$1 - \binom{r}{1} + \binom{r}{2} - \dots + (-1)^s \binom{r}{s} + \dots + (-1)^r \binom{r}{r} = 0.$$

Здесь мы воспользовались равенством (2.9).

Теорема 2.4. При $n \geq r \geq 1$

$$N(r) = \sum_{i_1 < \dots < i_r} N_{i_1 \dots i_r} + \dots + (-1)^{s-r} \binom{s}{r} \sum_{i_1 \dots i_s} N_{i_1 \dots i_s} + \dots + (-1)^{n-r} \binom{n}{r} N_{12 \dots n}. \quad (2.38)$$

Доказательство. Элементы, обладающие менее чем r свойствами, в правой части вообще не учитываются. Элементы, имеющие ровно r свойств, учитываются ровно по 1 разу в первом члене.

В общем случае рассмотрим элементы, обладающие ровно $t > r$ свойствами $p(j), \dots, p(j_t)$. Каждый из этих элементов входит в сумму

$$(-1)^{s-r} \binom{s}{r} \sum_{i_1 < \dots < i_s} N_{i_1 \dots i_s}, \quad s \leq t$$

ровно $\binom{t}{s}$ раз. Следовательно, их вклад в правую часть (2.38) равен

$$\sum_{s=r}^t (-1)^{s-r} \binom{s}{r} \binom{t}{s}. \quad (2.39)$$

При этом

$$\binom{s}{r} \binom{t}{s} = \frac{s!}{r! (s-r)!} \frac{t!}{s (t-s)!} = \frac{t!}{r! (t-r)!} \frac{(t-r)!}{(t-s)! (s-r)!} = \binom{t}{r} \binom{t-r}{s-r}.$$

Следовательно, сумма (2.39) равна

$$\binom{t}{r} \sum_{s=r}^t (-1)^{s-r} \binom{t-r}{s-r} = \binom{t}{r} \sum_{i=0}^{t-r} (-1)^i \binom{t-r}{i} = 0$$

(здесь мы положили $s-r=i$ и воспользовались равенством (2.29)).

Упражнение 2.8. Найдём число размещений без повторения из букв a, b, c, d, e, f , в которые не входят следующие две комбинации: (a, c, e) и (f, d) .

Решение. Свойства, состоящие в том, что размещение включает комбинации ace и fd , обозначим соответственно через $p(1)$ и $p(2)$. Если комбинация ace входит в размещение, то она может занимать $6-2=4$ различных положения (элемент a может занимать любую из позиций с первой по четвертую). В каждом из этих случаев остается еще три позиции, которые могут быть заполнены элементами b, d, f , взятыми в произвольном порядке. Поэтому $N_1 = 4(3!) = 4!$. Аналогично $N_2 = 5(4!) = 5!$. Кроме того, очевидно, $N = 6!$.

Рассмотрим теперь размещения, в которые входят обе комбинации ace и fd . Такие размещения можно разбить на две группы в зависимости от того, какая из двух комбинаций расположена первой. В каждом из этих вариантов оставшийся элемент может быть помещен либо перед этими комбинация-

ми, либо между ними, либо за ними. Поэтому $N_{12}=3 \cdot 2=3!$, и, согласно принципу включения — исключения, $N(0)=6!-(4!+5!)+3!=582$.

Определим функцию $[x]$ для вещественного x , как наибольшее целое число, не превосходящее x . Для положительных целых чисел a и b значение этой функции

$$\left\lfloor \frac{b}{a} \right\rfloor \quad (2.40)$$

равно количеству чисел в совокупности $1, 2, \dots, b$, которые делятся на a , т. е. количеству чисел, кратных a .

У п р а ж н е н и е 2.9. Сколько целых положительных чисел от 1 до 500 делятся либо на 3, либо на 5?

Решение. Если свойства делимости на 3 и на 5 обозначить соответственно через $p(1)$ и $p(2)$, то для $N=500$

$$N_1 = \left\lfloor \frac{500}{3} \right\rfloor = 166, \quad N_2 = \left\lfloor \frac{500}{5} \right\rfloor = 100.$$

Так как N_{12} — число общих кратных чисел 3 и 5, наименьшее общее кратное которых равно 15, то N_{12} совпадает с числом чисел, которые делятся на 15, и, следовательно,

$$N_{12} = \left\lfloor \frac{500}{15} \right\rfloor = 33.$$

Согласно принципу включения — исключения, количество чисел, которые не делятся ни на 3, ни на 5, равно $N(0)=500-(166+100)+33=267$. Следовательно, число чисел, которые делятся либо на 3, либо на 5, равно $500-267=233$.

Размещения без повторения из n целых чисел $1, 2, \dots$, по n , в которых ни одно число не занимает своего естественного места, называются *беспорядками*. Например, при $n=3$ размещение 312 является беспорядком, а 321 не является, так как число 2 в нем занимает свое естественное место. Найдем с помощью принципа включения — исключения общее число беспорядков d_n .

Обозначим через $p(i)$ свойство размещения, состоящее в том, что i -я позиция в нем занимает число i . Так как всего имеется $n!$ размещений, то $N=n!$. Так как $N_{i_1} \dots i_r$ — число размещений, i_1 -я, \dots , i_r -я позиции в которых должны быть заняты соответственно числами i_1, \dots, i_r , а остальные $(n-r)$ позиций могут быть заполнены оставшимися $(n-r)$ числами произвольным образом, то $N_{i_1} \dots i_r = (n-r)!$.

Имеется $\binom{n}{r}$ способов выбрать r позиций i_1, \dots, i_r из возможных, так что

$$\sum_{i_1 < \dots < i_r} N_{i_1 \dots i_r} = \binom{n}{r} (n-r)! = \frac{n!}{r!}.$$

Согласно принципу включения — исключения,

$$d_n = n! \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^r \frac{1}{r!} + \dots + (-1)^n \frac{1}{n!} \right] =$$

$$= n! \sum_{r=0}^n \frac{(-1)^r}{r!}. \quad (2.41)$$

Заметим, что $d_n/n!$ равно сумме первых $(n+1)$ членов ряда Тейлора при разложении $e^{-1}=0,36806 \dots$, $d_6/6!=0,36788 \dots$, и, как мы видим, скорость сходимости к e^{-1} велика.

2.5. Размещения с запрещенными позициями

Здесь рассматривается задача о подсчете общего числа размещений без повторения n различных элементов по позициям, некоторые из которых могут быть запрещенными для определенных элементов.

Так как произвольные n различных элементов всегда можно отождествить с целыми числами $1, 2, \dots, n$, то в дальнейшем будут рассматриваться только размещения целых чисел $1, 2, \dots, n$. Для описания запрещенных позиций будем пользоваться таблицей размера $n \times n$. Элемент (i, j) таблицы заштриховывается, если целому числу i запрещено занимать j -ю позицию. По аналогии с матрицами совокупности клеток таких таблиц, расположенных горизонтально и вертикально, будем называть соответственно строками и столбцами. При этом элементом (i, j) таблицы называется ее клетка, расположенная в i -строке и j -м столбце. Например, таблица рассмотренных в предыдущем разделе беспорядков имеет вид, показанный на рис. 2.2. На этом рисунке кружочки указывают для каждого целого числа позицию, на которую оно помещается. Естественно, что заштрихованные элементы кружочков не содержат и что каждая строка имеет не более одного

Рис. 2.2. Запрещенные (заштрихованные) позиции для беспорядков

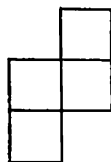
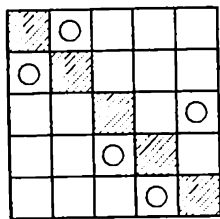


Рис. 2.3. Шахматная доска из упражнения 2.10

кружочка. Это связано с тем, что в случае беспорядков повторения элементов не допускаются и каждое число может занимать не более чем одну позицию. Каждый столбец также содержит не более одного кружочка, это следует из того, что каждую позицию может занимать только один элемент. Совокупность клеток описанной таблицы можно рассматривать как шахматную доску, а кружочки — как шашки, которые разрешается перемещать в горизонтальном и вертикальном направлениях. На рис. 2.2 приведен пример расположения шашек на шахматной доске, при котором на одной вертикали расположено не более одной шашки.

Рассмотрим шахматную доску произвольной формы и попробуем определить число способов, которыми можно разместить на ней k шашек так, чтобы никакие две из них не оказались в одном столбце. Обозначим это число через r_k . Очевидно,

$$r_0 = 1, \quad (2.42)$$

$$r_1 = \text{числу клеток на шахматной доске}. \quad (2.43)$$

Упражнение 2.10. Для шахматной доски, изображенной на рис. 2.3: $r_0 = 1$, $r_1 = 4$, $r_2 = 3$, $r_3 = r_4 = 0$.

Многочленом размещений шахматной доски C называют многочлен

$$R(x, C) = \sum_{k=0}^n r_k x^k, \quad (2.44)$$

где n — число столбцов шахматной доски. Так как число шашек, которые можно разместить так, чтобы никакие две из них не находились в одном столбце, не может быть больше, чем число столбцов шахматной доски, то

$$r_{n+1} = r_{n+2} = \dots = 0. \quad (2.45)$$

Так, для шахматной доски из примера 2.10 многочленом размещений является $R(x) = 1 + 4x + 3x^2$.

Если рассматривается одновременно несколько различных шахматных досок, то через $r_k(C_1)$ мы условимся обозначать число способов, которыми можно разместить k шашек на шахматной доске C_1 так, чтобы никакие две из них не находились в одном столбце.

Предположим, что на заданной шахматной доске C выделена одна определенная клетка. Обозначим через C_i шахматную доску, которая получается удалением из C строки и столбца, содержащих эту выделенную клетку, а через C_e — шахматную доску, получающуюся из C удалением только одной выделенной клетки. При размещении k шашек на шахматной доске C возможно два случая в зависимости от того, помещается или нет шашка в выделенную клетку. Если шашка помещается в выделенную клетку, то в столбец и строку, ее содержащие, другие шашки уже помещены быть не могут. Следовательно, число размещений k шашек, при которых выделенная клетка заполнена и никакие две шашки не входят в один столбец, равно $r_{k-1}(C_i)$. Число размещений k шашек, при которых выделенная клетка не заполнена и никакие две шашки не входят в один столбец, равно $r_k(C_e)$. Таким образом,

$$r_k(C) = r_{k-1}(C_i) + r_k(C_e). \quad (2.46)$$

Если многочлен размещений шахматной доски C обозначить через $R(x, C)$, то последнему равенству будет соответствовать следующее соотношение для многочленов размещений:

$$R(x, C) = xR(x, C_i) + R(x, C_e). \quad (2.47)$$

В этом равенстве для наглядности вместо многочленов размещений можно в круглых скобках записывать непосредственно соответствующие им шахматные доски. Например,

$$\left(\begin{array}{|c|c|} \hline * & \\ \hline \hline \end{array} \right) = x \left(\begin{array}{|c|} \hline \\ \hline \end{array} \right) + \left(\begin{array}{|c|} \hline \\ \hline \end{array} \right)$$

$$\left(\begin{array}{|c|c|} \hline & * \\ \hline \hline \end{array} \right) = x \left(\begin{array}{|c|} \hline \\ \hline \end{array} \right) + \left(\begin{array}{|c|} \hline \\ \hline \end{array} \right)$$

(здесь знаком * отмечены выделяемые клетки). Если в результате удаления строки и столбца, содержащих помеченную знаком * клетку, в шахматной доске не остается ни одной клетки, то некоторые скобки в указанной символической записи также могут оказаться пустыми.

У п р а ж н е н и е 2.11.

$$\left(\begin{array}{|c|} \hline \\ \hline \end{array} \right) = 1, \left(\begin{array}{|c|c|} \hline & \\ \hline \end{array} \right) = 1+x, \left(\begin{array}{|c|c|} \hline & \\ \hline \hline \end{array} \right) = 1+2x,$$

$$\left(\begin{array}{|c|c|} \hline & \\ \hline \hline \end{array} \right) = 1+2x+x^2$$

Из примера, приведенного выше, следует, что

$$\left(\begin{array}{|c|c|} \hline & \\ \hline \hline \end{array} \right) = x \cdot 1 + (1+2x+x^2) = 1+3x+x^2$$

Если шахматная доска C разбивается на две непересекающиеся доски C_1 и C_2 такие, что C_1 и C_2 не имеют ни одного общего столбца и ни одной общей строки, то

$$r_k(C) = \sum_{i=0}^k r_i(C_1) r_{k-i}(C_2). \quad (2.48)$$

Это равенство следует из того, что никакие шашки, помещаемые в C_1 , не могут входить в тот же столбец C , что и шашки, помещаемые в C_2 , а следовательно, для размещения с нужными свойствами k шашек в C можно i шашек разместить таким же образом в C_1 ; а оставшиеся $(k-i)$ — в C_2 . Из полученного равенства следует равенство для многочленов:

$$R(x, C) = R(x, C_1) R(x, C_2). \quad (2.49)$$

Например,

$$\left(\begin{array}{|c|c|} \hline & \\ \hline \hline \end{array} \right) = \left(\begin{array}{|c|} \hline \\ \hline \end{array} \right) \left(\begin{array}{|c|} \hline \\ \hline \end{array} \right) = (1+x)^2 = 1+2x+x^2.$$

Общее число беспорядков d_n равно числу способов размещения n шашек на шахматной доске, из которой исключены диаго-

нальные клетки. В общем случае для определения числа размещений $1, 2, \dots, n$ следует рассмотреть шахматную доску, получающуюся из шахматной доски размера $n \times n$ удалением клеток, в которые числа запрещено помещать, и определить для нее многочлен размещений. Коэффициент r_n этого многочлена будет искомым числом размещений. Однако если число запрещенных позиций мало, лучше рассмотреть шахматную доску, состоящую, наоборот, только из клеток, в которые элементы запрещается помещать, и определить многочлен размещений для такой доски. Этот многочлен обозначим через

$$R(x) = \sum_{k=0}^n r_k x^k. \quad (2.50)$$

Для того чтобы вычислить число размещений, можно воспользоваться принципом включения — исключения. При этом свойство $p(i)$ заключается в том, что целое число i не занимает запрещенных для него позиций. В случае когда целые числа i_1, \dots, i_s расположены на запрещенных позициях, а остальные размещаются произвольно,

$$N_{i_1 \dots i_s} = (n-s)!$$

Число размещений, при которых ровно s элементов занимают запрещенные позиции, совпадает с числом r_s таких размещений k шашек на рассматриваемой шахматной доске, при которых никакие две шашки не расположены в одном столбце. Следовательно,

$$\sum_{i_1 < \dots < i_s} N_{i_1 \dots i_s} = r_s (n-s)!$$

Согласно принципу включения — исключения, число размещений, при которых запрещенные позиции не занимаются,

$$\begin{aligned} N(0) &= n! - r_1 (n-1)! + r_2 (n-2)! + \dots + (-1)^{n-1} r_{n-1} 1! + \\ &+ (-1)^n r_n 0! = \sum_{j=0}^n (-1)^j r_j (n-j)!. \end{aligned} \quad (2.51)$$

Например, в случае беспорядков, шахматная доска, состоящая только из запрещенных клеток, имеет вид, показанный на рис. 2.4. Для такой шахматной доски многочлен размещений

$$(\square)^n = (1+x)^n = \sum_{j=0}^n \binom{n}{j} x^j \quad (2.52)$$

и, следовательно,

$$d_n = \sum_{j=0}^n (-1)^j \binom{n}{j} (n-j)! = n! \sum_{j=0}^n \frac{(-1)^j}{j!},$$

что, естественно, совпадает с (2.41).

У п р а ж н е н и е 2.12. Три дома, a , b , c , должны быть окрашены тремя различными цветами, например зеленым, синим и желтым. Дом a нельзя красить желтым цветом, дом b — синим и

желтым, дом c — зеленым. Спрашивается, сколькими способами можно окрасить дома?

Решение. Если дома и цвета перенумеровать целыми числами 1, 2, 3, то искомое число будет равно числу размещений целых чисел 1, 2, 3 на шахматной доске с запрещенными позициями, изображенной на рис. 2.5. Определим многочлен размещений для шахматной доски, состоящей только из запрещенных клеток. Имеем

$$\begin{pmatrix} \square & \square & \square \\ \square & & \end{pmatrix} = \begin{pmatrix} \square \end{pmatrix} \begin{pmatrix} \square & \square \\ \square & \end{pmatrix} = (1+x)(1+3x+x^2) \\ = 1+4x+4x^2+x^3$$

Отсюда и из (2.51) находим искомое число: $3! - 4(2!) + 4(1!) - 1(0!) = 6 - 8 + 4 - 1 = 1$.

В последнем упражнении очевидно, что дом b может быть окрашен только в зеленый цвет. При этом для дома a остается только синий цвет, а для дома c — только желтый, т. е. очевидно, что в данном случае дома можно раскрасить лишь единственным способом.

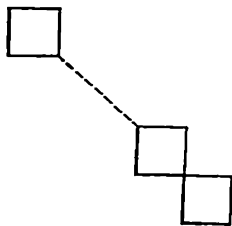


Рис. 2.4. Шахматная доска, соответствующая запрещенным позициям для беспорядков

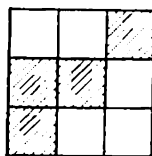


Рис. 2.5. Запрещенные (заштрихованные) позиции для упражнения 2.12

Задачи

2.1. Покажите, что показатель степени простого числа p в разложении n на простые сомножители равен

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots$$

2.2. Покажите, что для произвольных чисел n и k , $1 \leq k < n$, справедлива формула

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

2.3. Функция $\theta(n)$, определенная на множестве всех целых положительных чисел n , называется мультипликативной, если она обладает следующими двумя свойствами: 1) $\theta(1) = 1$, 2) $\theta(n_1 n_2) = \theta(n_1) \theta(n_2)$, если $(n_1, n_2) = 1$. Покажите, что функция Мебиуса $\mu(n)$ является мультипликативной.

2.4. Найдите число размещений с повторениями четырех букв, a, b, c, d , по r ($r \geq 3$) позициям, использующих каждую из букв a, b, c , по крайней мере, по одному разу.

2.5. Рассмотрим матрицу из $2n$ строк и n столбцов. Найдите многочлен размещений R_n шахматной доски C_n , состоящей из элементов $(1,1), (2,2), \dots, (n,n)$ и $(2n,1), (2n-1,2), \dots, (n+1,n)$ рассматриваемой матрицы.

Глава 3

Рекуррентные уравнения

3.1. Производящие функции

При подсчете числа предметов широко используются следующие две формулы:

$$1 + 2 + \dots + n = n(n+1)/2; \quad (3.1)$$

$$1^2 + 2^2 + \dots + n^2 = n(n+1)(2n+1)/6. \quad (3.2)$$

Если обозначить сумму левой части (3.1) через f_n , а сумму левой части (3.2) через g_n , то будут иметь место соотношения

$$f_{n+1} - f_n = n+1; \quad g_{n+1} - g_n = (n+1)^2. \quad (3.3)$$

Взяв в качестве начальных значений $f_1 = g_1 = 1$, с помощью (3.3) можно определить f_2 и g_2 ; также можно определить f_3 и g_3 и т. д. Таким образом, последовательности $\{f_n\}$ и $\{g_n\}$ однозначно определяются соотношением (3.3) и начальными условиями. Обобщением (3.3) является линейное рекуррентное уравнение с постоянными коэффициентами

$$f_{n+r} + a_1 f_{n+r-1} + a_2 f_{n+r-2} + \dots + a_r f_n = \Phi(n), \quad (3.4)$$

выполненное для всех неотрицательных целых чисел n . Коэффициенты a_1, a_2, \dots, a_r — фиксированные числа, причем

$$a_r \neq 0, \quad (3.5)$$

а $\Phi(n)$ — заданная функция n . Если зафиксировать значения

$$f_{r-1}, f_{r-2}, \dots, f_1, f_0 \quad (3.6)$$

и рассматривать их как начальные условия, то шаг за шагом можно однозначно определить значения f_r, f_{r+1}, \dots и таким образом с помощью линейного рекуррентного уравнения (3.4) определить всю последовательность $\{f_n\}$. Степенной ряд

$$F(z) = f_0 + f_1 z + f_2 z^2 + \dots, \quad (3.7)$$

коэффициентами которого являются элементы последовательности f_0, f_1, f_2, \dots , называется *производящей функцией*. Последовательность чисел однозначно определяет производящую функцию, но обратное утверждение верно не всегда. Если указанный степенной ряд сходится в круге радиуса $R > 0$, то, как следует из теории аналитических функций, коэффициенты f_0, f_1, \dots определяются по $F(z)$ однозначно.

Пусть

$$G(z) = g_0 + g_1 z + g_2 z^2 + \dots \quad (3.8)$$

— производящая функция последовательности чисел g_0, g_1, g_2, \dots , а c и d — произвольные фиксированные числа. Поскольку

$$cF(z) + dG(z) = (cf_0 + dg_0) + (cf_1 + dg_1)z + \dots, \quad (3.9)$$

то последовательности $\{cf_n + dg_n\}$ отвечает производящая функция $cF(z) + dG(z)$. Далее, если взять произведение производящих функций

$$H(z) = F(z)G(z) = h_0 + h_1 z + h_2 z^2 + \dots, \quad (3.10)$$

то последовательность чисел $\{h_n\}$ может быть получена из последовательностей $\{f_n\}$ и $\{g_n\}$ с помощью соотношений

$$h_n = f_0 g_n + f_1 g_{n-1} + \dots + f_{n-1} g_1 + f_n g_0. \quad (3.11)$$

Упражнение 3.1. Покажите, что производящей функцией последовательности чисел $\{1\}$ является $1/(1-z)$, а производящей функцией последовательности $\{n\}$ является $z/(1-z)^2$.

Решение. Производящая функция последовательности $\{1\}$: $1 + z + z^2 + \dots = 1/(1-z)$. Последовательности чисел $\{n\}$ соответствует ряд $z + 2z^2 + 3z^3 + \dots = z(1 + 2z + 3z^2 + \dots)$. Поскольку выражение в скобках в правой части последнего равенства получается дифференцированием ряда $1 + z + z^2 + \dots$, то оно равно производной от функции $1/(1-z)$. Следовательно, правая часть равна

$$z \frac{d}{dz} \left(\frac{1}{1-z} \right) = \frac{z}{(1-z)^2}.$$

Для решения этой задачи можно было бы воспользоваться соотношением (3.11) и показать, что последовательностью, соответствующей производящей функции $1/(1-z)^2$, является последовательность $1, 2, 3, \dots$. Отсюда также следовало бы, что производящей функцией последовательности чисел $0, 1, 2, \dots$ является $z/(1-z)^2$.

Следующая задача показывает, что производящие функции могут быть полезными при решении линейных рекуррентных уравнений типа (3.4).

Упражнение 3.2. Решите уравнение $f_{n+2} - f_{n+1} - f_n = 0$ с начальными условиями $f_0 = f_1 = 1$.

Решение. Обозначим через $F(z)$ производящую функцию последовательности чисел $\{f_n\}$. Умножая обе части рекуррентного уравнения на z^{n+2} , получаем

$$f_{n+2} z^{n+2} - z f_{n+1} z^{n+1} - z^2 f_n z^n = 0, \quad n = 0, 1, 2, \dots$$

Складывая эти равенства для всех n от 0 до ∞ , имеем

$$\sum_{n=0}^{\infty} f_{n+2} z^{n+2} - z \sum_{n=0}^{\infty} f_{n+1} z^{n+1} - z^2 \sum_{n=0}^{\infty} f_n z^n = 0.$$

Заметим, что первая сумма в левой части последнего равенства равна разности функции $F(z)$ и первых двух членов ее разложения $f_0 + f_1 z = 1 + z$, вторая сумма равна разности $F(z)$ и первого члена $f_0 = 1$, а третья сумма равна $F(z)$, поэтому можно записать следующее равенство:

$$[F(z) - (1 + z)] - z[F(z) - 1] - z^2 F(z) = 0.$$

Отсюда находим $F(z) = 1/(1 - z - z^2)$.

Полагая

$$\alpha_1 = (1 + \sqrt{5})/2; \quad \alpha_2 = (1 - \sqrt{5})/2, \quad (3.12)$$

получаем

$$F(z) = \frac{1}{\sqrt{5}} \left(\frac{\alpha_1}{1 - \alpha_1 z} - \frac{\alpha_2}{1 - \alpha_2 z} \right) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{5}} (\alpha_1^{n+1} - \alpha_2^{n+1}) z^n.$$

Таким образом,

$$f_n = \frac{1}{\sqrt{5}} (\alpha_1^{n+1} - \alpha_2^{n+1}), n = 0, 1, 2, \dots \quad (3.13)$$

Члены последовательности, полученной в этой задаче, известны как числа Фибоначчи. Несколько первых членов этой последовательности равны: 1, 1, 2, 3, 5, 8, 13, 21, 34, ... При больших n члены приближенно равны $\alpha_1^{n+1} / \sqrt{5}$. Коэффициент $\alpha_1 = 1,6180337 \dots$ известен в геометрии как «золотое сечение».

3.2. Решение однородного линейного рекуррентного уравнения

Найдем общее решение однородного уравнения, получающегося из (3.4) при $\varphi(n) = 0$. Метод решения является обобщением решения упражнения 3.2. При этом вначале производящая функция находится как рациональная функция, которая далее представляется в виде суммы частичных дробей и разлагается в степенной ряд.

Предположим, что последовательность чисел f_0, f_1, \dots удовлетворяет следующему однородному линейному рекуррентному уравнению:

$$f_{n+r} + a_1 f_{n+r-1} + \dots + a_r f_n = 0, \quad n = 0, 1, 2, \dots, \quad (3.14)$$

где a_1, \dots, a_r — заданные постоянные числа и

$$a_r \neq 0. \quad (3.15)$$

Для задания начальных условий фиксируем значения f_0, \dots, f_{r-1} .

Обозначим через $F(z)$ производящую функцию последовательности f_0, f_1, \dots . По заданным постоянным коэффициентам уравнения построим многочлен

$$K(z) = 1 + a_1 z + a_2 z^2 + \dots + a_r z^r. \quad (3.16)$$

Этот многочлен можно рассматривать как производящую функцию последовательности $1, a_1, \dots, a_r, 0, 0, \dots$. Коэффициент c_{n+r} при z^{n+r} и $r > 0$ в произведении производящих функций $F(z)K(z)$, как следует из (3.11), определяется соотношением

$$\begin{aligned} c_{n+r} &= f_0 \cdot 0 + \dots + f_{n-1} \cdot 0 + f_n \cdot a_r + \dots + f_{n+r} \cdot 1 = \\ &= f_{n+r} + a_1 f_{n+r-1} + \dots + a_r f_n \end{aligned}$$

и, согласно равенству (3.14), равен нулю. Это означает, что многочлен

$$F(z)K(z) = c_0 + c_1 z + \dots + c_{r-1} z^{r-1} = C(z) \quad (3.17)$$

имеет степень самое большее $(r-1)$, а следовательно, степень числителя рациональной функции

$$F(z) = C(z)/K(z) \quad (3.18)$$

меньше степени знаменателя.

Характеристическим многочленом линейного рекуррентного уравнения (3.14) называется многочлен

$$g(z) = z^r + a_1 z^{r-1} + \dots + a_r, \quad (3.19)$$

имеющий степень r ; корни этого многочлена называются *характеристическими*. Если различные характеристические корни (среди которых могут быть мнимые) обозначить через $\alpha_1, \dots, \alpha_s$, а их кратности обозначить соответственно через e_1, \dots, e_s , то можно записать следующие равенства:

$$g(z) = (z - \alpha_1)^{e_1} (z - \alpha_2)^{e_2} \dots (z - \alpha_s)^{e_s}; \quad (3.20)$$

$$e_1 + e_2 + \dots + e_s = r. \quad (3.21)$$

Характеристический многочлен $g(z)$ и многочлен $K(z)$ связаны между собой соотношением

$$K(z) = z^r g\left(\frac{1}{z}\right). \quad (3.22)$$

Отсюда и из (3.20) следует, что

$$K(z) = (1 - \alpha_1 z)^{e_1} (1 - \alpha_2 z)^{e_2} \dots (1 - \alpha_s z)^{e_s}. \quad (3.23)$$

Используя последнее равенство, рациональную функцию (3.18), у которой степень числителя меньше степени знаменателя, можно представить в виде суммы частичных дробей

$$F(z) = \frac{C(z)}{K(z)} = \sum_{i=1}^s \sum_{k=1}^{e_i} \frac{\beta_{ik}}{(1 - \alpha_i z)^k}. \quad (3.24)$$

Каждая дробь этой суммы имеет вид $\beta(1 - \alpha z)^{-k}$, поэтому ее можно разложить в степенной ряд следующего вида:

$$\begin{aligned} \beta(1 - \alpha z)^{-k} = & \beta \left[1 + (-k)(-\alpha z) + \dots + \right. \\ & \left. + \frac{(-k) \dots (-k - n + 1)}{n!} (-\alpha z)^n + \dots \right]. \end{aligned}$$

Коэффициент при z^n в этом ряде равен

$$\frac{\beta(n + k - 1) \dots k}{n!} \alpha^n = \beta \binom{n + k - 1}{n} \alpha^n = \beta \binom{n + k - 1}{k - 1} \alpha^n. \quad (3.25)$$

Если заметить, что биномиальный коэффициент

$$\binom{n + k - 1}{k - 1},$$

входящий в последнее равенство, является многочленом степени $k - 1$ по n , то легко проверить, что

$$\left[\sum_{k=1}^{e_i} \beta_{ik} \binom{n + k - 1}{k - 1} \right] \alpha_i^n = p_i(n) \alpha_i^n, \quad (3.26)$$

где $p_i(n)$ — многочлен от n степени самое большее (e_i-1) . Следовательно,

$$F(z) = \sum_{n=0}^{\infty} \left(\sum_{i=1}^s p_i(n) \alpha_i^n \right) z^n \quad (3.27)$$

и

$$f_n = \sum_{i=1}^s p_i(n) \alpha_i^n \quad (3.28)$$

является общим решением линейного рекуррентного уравнения (3.14).

Упражнение 3.3. Описанным общим методом найдите общий член последовательности чисел Фибоначчи.

Решение. Уравнение $f_{n+2} - f_{n+1} - f_n = 0$ имеет характеристический многочлен

$$z^2 - z - 1 = (z - \alpha_1)(z - \alpha_2),$$

где α_1 и α_2 — постоянные, определяемые из (3.12). Так как α_1 и α_2 являются простыми корнями, то $e_1=1$, $e_2=1$ и, следовательно, $p_1(n)$ и $p_2(n)$ — многочлены степени 0 от n , т. е. постоянные. Поэтому $f_n = c_1 \alpha_1^n + c_2 \alpha_2^n$, где c_1, c_2 — неопределенные постоянные. Так как $f_0 = f_1 = 1$, то, подставляя $n=0, 1$, получаем $c_1 + c_2 = 1$, $c_1 \alpha_1 + c_2 \alpha_2 = 1$. Решая эти уравнения, находим

$$c_1 = \frac{\alpha_2 - 1}{\alpha_2 - \alpha_1} = \frac{1}{\sqrt{5}} \alpha_1; \quad c_2 = \frac{\alpha_1 - 1}{\alpha_1 - \alpha_2} = -\frac{1}{\sqrt{5}} \alpha_2.$$

Отсюда следует, что

$$f_n = \frac{1}{\sqrt{5}} (\alpha_1^{n+1} - \alpha_2^{n+1}).$$

Решение этого упражнения показывает, что если все характеристические корни $\alpha_1, \dots, \alpha_r$ являются простыми, то общее решение однородного уравнения имеет вид

$$f_n = c_1 \alpha_1^n + c_2 \alpha_2^n + \dots + c_r \alpha_r^n, \quad (3.29)$$

где c_1, c_2, \dots, c_r — это r неопределенных постоянных. Для определения постоянных используются r начальных условий, а именно значения f_0, f_1, \dots, f_{r-1} .

Если α_i является корнем кратности e_i , то $p_i(n)$ представляет собой многочлен степени (e_i-1) :

$$p_i^{(n)} = c_{i0} + c_{i1} n + \dots + c_{ie_i-1} n^{e_i-1}, \quad (3.30)$$

где $c_{i0}, c_{i1}, \dots, c_{ie_i-1}$ — неопределенных постоянных. Как показывает условие (3.21), всего в данном случае имеется r неопределенных постоянных. Эти постоянные можно определить по граничным условиям. Действительно, по граничным условиям однозначно определяются последовательность чисел f_n и производящая функция. Поэтому однозначно определяются также значения β_{ik} , а следовательно, и многочлены $p_i(n)$. Это означает, что граничные условия однозначно определяют все r неопределенных постоянных.

У п р а ж н е н и е 3.4. Найдите решение уравнения $f_{n+2} - f_{n+1} + f_n = 0$ с граничными условиями $f_0 = 1$ и $f_1 = 1$.

Решение. Данное уравнение имеет характеристический многочлен $z^2 - z + 1 = (z - \alpha_1)(z - \alpha_2)$, где

$$\alpha_1 = \frac{1}{2} + i \frac{\sqrt{3}}{2} = e^{i(\pi/3)}; \quad \alpha_2 = \frac{1}{2} - i \frac{\sqrt{3}}{2} = e^{-i(\pi/3)},$$

i — мнимая единица. Общее решение имеет вид

$$f_n = c_1 \alpha_1^n + c_2 \alpha_2^n = c_1 e^{i(\pi/3)n} + c_2 e^{-i(\pi/3)n}.$$

Подставляя $n=0,1$, получаем $c_1 + c_2 = 1$, $c_1 \alpha_1 + c_2 \alpha_2 = 1$. Отсюда, в свою очередь, имеем

$$c_1 = \frac{1}{2} - i \frac{1}{2\sqrt{3}}; \quad c_2 = \frac{1}{2} + i \frac{1}{2\sqrt{3}}.$$

Следовательно,

$$f_n = \left(\frac{1}{2} - i \frac{1}{2\sqrt{3}} \right) \left(\cos \frac{\pi}{3} n + i \sin \frac{\pi}{3} n \right) + \left(\frac{1}{2} + i \frac{1}{2\sqrt{3}} \right) \times \\ \times \left(\cos \frac{\pi}{3} n - i \sin \frac{\pi}{3} n \right) = \cos \frac{\pi}{3} n + \frac{1}{\sqrt{3}} \sin \frac{\pi}{3} n.$$

У п р а ж н е н и е 3.5. Найдите решение уравнения $f_{n+2} - 4f_{n+1} + 4f_n = 0$ с граничными условиями $f_0 = 1$, $f_1 = 4$.

Решение. Так как характеристический многочлен $z^2 - 4z + 4 = (z - 2)^2$ имеет корень $z=2$ кратности 2, то $f_n = (c_0 + c_1 n) \cdot 2^n$. С помощью граничных условий находим: $c_0 = 1$, $2(c_0 + c_1) = 4$, т. е. $c_1 = 1$. Таким образом, решение рассматриваемого уравнения имеет вид $f_n = (1 + n)2^n$.

3.3. Метод решения неоднородного линейного рекуррентного уравнения

Рассмотрим неоднородное линейное рекуррентное уравнение

$$f_{n+r} + a_1 f_{n+r-1} + \dots + a_r f_n = \Phi(n), \quad n=0, 1, 2, \dots, \quad (3.31)$$

где a_1, \dots, a_r — заданные постоянные, причем

$$a_r \neq 0 \quad (3.32)$$

и $\Phi(n)$ — заданная функция n . Для задания начальных условий фиксируем значения f_0, f_1, \dots, f_{r-1} .

По аналогии с методами решения дифференциальных уравнений вначале пренебрежем начальными условиями и предположим, что одно решение уравнения (3.31) найдено. Назовем это решение частным и обозначим через

$$\{f_n^{(p)}\}. \quad (3.33)$$

Положим $f_n = f_n^{(h)} + f_n^{(p)}$. Тогда

$$(f_{n+r}^{(h)} + a_1 f_{n+r-1}^{(h)} + \dots + a_r f_n^{(h)}) + (f_{n+r}^{(p)} + a_1 f_{n+r-1}^{(p)} + \dots + a_r f_n^{(p)}) = \Phi(n). \quad (3.34)$$

Так как второй член в левой части последнего равенства равен правой части, то

$$f_{n+r}^{(h)} + a_1 f_{n+r-1}^{(h)} + \dots + a_r f_n^{(h)} = 0. \quad (3.35)$$

Это означает, что $\{f_n^{(h)}\}$ является решением однородного линейного рекуррентного уравнения, соответствующего $\varphi(n) = 0$.

Таким образом, если найдено частное решение, то с помощью техники, описанной в предыдущем параграфе, можно найти общее решение однородного рекуррентного уравнения, после чего по начальным условиям можно определить неопределенные постоянные для решения f_n , имеющего вид: (общее решение) + (частное решение). Для некоторых функций $\varphi(n)$ специального вида частное решение можно найти достаточно просто, но в общем случае, как и при решении дифференциальных уравнений, можно лишь рекомендовать воспользоваться одним из известных методов определения частного решения, например методом вариации постоянных [3], [4], [5].

Так, в случае если $\varphi(n) = \beta^n$, где β — это характеристический корень (быть может, равный нулю), частное решение можно найти довольно просто следующим образом. Подставляя

$$f_n = c \beta^n, \quad (3.36)$$

где c — постоянная, в уравнение (3.31), получаем $c(\beta^r + a_1 \beta^{r-1} + \dots + a_r) \beta^n = \beta^n$ $c \beta^n g(\beta) = \beta^n$, т. е. $c g(\beta) = 1$ (здесь $g(\beta)$ — характеристический многочлен). Следовательно,

$$f_n = \frac{1}{g(\beta)} \beta^n \quad (3.37)$$

является частным решением.

У п р а ж н е н и е 3.6. Найдите решение уравнения $f_{n+2} + 2f_{n+1} + f_n = 3 \cdot 2^n$ с граничными условиями $f_0 = 1$, $f_1 = 1$.

Решение. Данное уравнение имеет характеристический многочлен $g(z) = z^2 + 2z + 1 = (z+1)^2$. Если бы правая часть уравнения была равна 2^n , то частным решением было $\frac{1}{g(2)} 2^n = \frac{1}{9} \cdot 2^n$. Отсюда легко видеть, что правой части $3 \cdot 2^n$ соответствует частное решение $3 \cdot \frac{1}{9} \cdot 2^n = \frac{1}{3} \cdot 2^n$. Следовательно, $f_n = (c_0 + c_1 n) (-1)^n + \left(\frac{1}{3}\right) \cdot 2^n$.

По граничным условиям находим:

$$c_0 + 1/3 = 1; \quad -(c_0 + c_1) + (1/3) \cdot 2 = 1; \quad c_0 = 2/3; \quad c_1 = -1; \quad f_n = \left(\frac{2}{3} - n\right) \times (-1)^n + \left(\frac{1}{3}\right) \cdot 2^n.$$

Рассмотрим случай, когда $\varphi(n)$ является многочленом от n степени k и 1 не является характеристическим корнем линейного

рекуррентного уравнения. Пусть $g(z)$ — характеристический многочлен. Тогда

$$g(1) = 1 + a_1 + a_2 + \dots + a_r \neq 0. \quad (3.38)$$

Частное решение следует искать в виде многочлена от n степени k :

$$f_n = \sum_{i=0}^k b_i n^i. \quad (3.39)$$

Подставляя этот многочлен в (3.31), получаем

$$\begin{aligned} \sum_{i=0}^k b_i [(n+r)^i + a_1 (n+r-1)^i + \dots + a_r n^i] = \\ = \sum_{i=0}^k b_i (g(1) n^i + \dots) = \varphi(n). \end{aligned}$$

Так как $g(1) \neq 0$, то, сравнивая коэффициенты при высших степенях в левой и правой частях последнего равенства, можно определить значение b_k и далее последовательно коэффициенты b_{k-1}, \dots, b_1, b_0 .

У п р а ж н е н и е 3.7. Найдите решение уравнения $f_{n+1} + 2f_n = n + 1$ с граничным условием $f_0 = 1$.

Решение. Характеристическим в данном случае является многочлен $g(z) = z + 2$. Так как $g(1) = 3 \neq 0$, то частным решением будет многочлен первой степени. Подставляя

$$f_n = b_0 + b_1 n$$

в уравнение, получаем

$$[b_0 + b_1 (n+1)] + 2[b_0 + b_1 n] = (3b_0 + b_1) + 3b_1 n = 1 + n.$$

Отсюда следует, что $b_1 = 1/3$, $b_0 = 2/9$. Найдя частное решение, положим $f_n = \frac{2}{9} + \frac{1}{3}n + c(-2)^n$.

Из начальных условий находим $2/9 + c = 1$, т. е. $c = 7/9$, $f_n = \frac{2}{9} + \frac{1}{3}n + \frac{7}{9}(-2)^n$.

Если $\varphi(n)$ имеет вид β^n , где β — характеристический корень, или если $\varphi(n)$ — многочлен по n и 1 является характеристическим корнем уравнения, то найти частное решение также просто, как раньше, уже не удастся. Эта задача будет рассмотрена в § 3.5.

3.4. Разностные формулы

Рассмотрим функцию $f(x)$, определенную для всех неотрицательных целых чисел. Обозначая значение этой функции, соответствующее целому числу x , через f_x , будем рассматривать, как и ранее, последовательности чисел f_0, f_1, f_2, \dots .

Выражение

$$\Delta f(x) = f(x+1) - f(x) \quad (3.40)$$

назовем приращением f в точке x . Разности типа (3.40) удобно записывать в виде

$$\Delta f_x = f_{x+1} - f_x. \quad (3.41)$$

Приращения соответствуют дифференциалам в анализе. Для произвольных фиксированных чисел a и b имеют место равенства

$$\Delta [af(x) + bg(x)] = a \Delta f(x) + b \Delta g(x); \quad (3.42)$$

$$\Delta [f(x)g(x)] = f(x+1)\Delta g(x) + g(x)\Delta f(x) = f(x)\Delta g(x) + g(x+1)\Delta f(x). \quad (3.43)$$

Первое из них очевидно, а второе следует из соотношения $f(x+1)g(x+1) - f(x)g(x) = f(x+1)g(x+1) - f(x+1)g(x) + f(x+1)g(x) - f(x)g(x)$. В этом соотношении можно сначала вычесть, а затем прибавить вместо $f(x+1)g(x)$ произведение $f(x)g(x+1)$.

Упражнение 3.8. Если c — постоянная, то $\Delta c = 0$; $\Delta x = 1$.

В анализе наряду с обычными дифференциалами рассматриваются дифференциалы высших порядков. Можно определить наряду с обычными приращениями также приращения высших порядков. Так приращением 2-го порядка функции f в точке x является приращение приращения $\Delta f(x)$:

$$\Delta [\Delta f(x)] = \Delta^2 f(x). \quad (3.44)$$

При этом

$$\Delta^2 f(x) = \Delta [f(x+1) - f(x)] = f(x+2) - 2f(x+1) + f(x). \quad (3.45)$$

Если такую операцию проделать r раз, то получится приращение r -го порядка: $\Delta^r f(x)$. Обобщением равенства (3.45) для приращений высших порядков является следующая важная формула:

$$\Delta^r f_i = \sum_{k=0}^r (-1)^{r-k} \binom{r}{k} f_{i+k}. \quad (3.46)$$

Доказательство. При $r=1$ это равенство совпадает с определением приращения. Для доказательства справедливости формулы, в общем случае, следует воспользоваться методом математической индукции, т. е. доказать ее справедливость для приращений порядка $r+1$, допустив, что она верна для приращений порядка r . Имеем

$$\begin{aligned} \Delta^{r+1} f_i &= \Delta [\Delta^r f_i] = \sum_{k=0}^r (-1)^{r-k} \binom{r}{k} \Delta f_{i+k} = \\ &= \sum_{k=0}^r (-1)^{r-k} \binom{r}{k} (f_{i+k+1} - f_{i+k}) = (-1)^{r+1} f_i + \sum_{k=1}^r (-1)^{r+1-k} \times \\ &\times \left[\binom{r}{k} + \binom{r}{k-1} \right] f_{i+k} + f_{i+r+1} = \sum_{k=0}^{r+1} (-1)^{r+1-k} \binom{r+1}{k} f_{i+k} \end{aligned}$$

(здесь использовалось равенство (2.7) из главы 2).

Упражнение 3.9. Покажите, что приращением порядка k многочлена $a_0 x^k + \dots$ степени k является $a_0 k!$, и, следовательно, приращения порядка $k+1$ и выше все равны 0.

Решение. В данном случае

$$\Delta x^k = (x+1)^k - x^k = kx^{k-1} + \dots$$

Так как при вычислении каждого следующего приращения степень многочлена в правой части равенства понижается на единицу, то приращение k -го порядка многочлена степени k будет постоянным, представляющим собой произведение коэффициента a_0 при наивысшей степени многочлена на $k!$. Отсюда и из упражнения 3.8 следует, что все приращения более высоких порядков будут равны 0.

Уравнения вида (3.31), которые назывались линейными рекуррентными уравнениями с постоянными коэффициентами, часто называют также разностными уравнениями. Действительно, воспользовавшись соотношением (3.46), уравнение (3.31) можно записать в виде

$$\Delta^r f_n + a'_1 \Delta^{r-1} f_n + \dots + a'_r \Delta^0 f_n = \varphi(n). \quad (3.47)$$

Однако, по определению,

$$\Delta^0 f_x = f_x. \quad (3.48)$$

Так как в силу (3.46) $f_{n+r} = \Delta^r f_n +$ (выражение, линейное по f_{n+r-1}, \dots, f_n), то после подстановки этого выражения в (3.31) вместо f_{n+r} будет $\Delta^r f_n$. Далее, если вместо f_{n+r-1} подставить $\Delta^{r-1} f_n +$ (линейное по f_{n+r-2}, \dots, f_n выражение) и так далее, то можно получить уравнение (3.47). Для того чтобы (3.47) записать в виде (3.31), следует воспользоваться равенством (3.46).

Функцию $f(x)$, являющуюся многочленом степени k , можно разложить в ряд Тейлора следующим образом:

$$f(x_0 + x) = \sum_{i=0}^k \frac{f^{(i)}(x_0)}{i!} x^i. \quad (3.49)$$

Найдем разностное уравнение, соответствующее этому разложению. Для этого для всех неотрицательных целых чисел ($j \geq 0$) определим обобщенную степень, положив

$$n^{(j)} = n(n-1)\dots(n-j+1); \quad (3.50)$$

$$n^{(0)} = 1 \quad (3.51)$$

для $j=0$ (в частности, $0^{(0)}=1$). Другие свойства этой функции очевидны:

$$0^{(j)} = 0, \quad j^{(j)} = j! \text{ если } j \geq 1; \quad (3.52)$$

$$n^{(j)} = 0, \text{ если } j > n \geq 0. \quad (3.53)$$

Обобщенная степень обладает также следующим интересным свойством:

$$\Delta n^{(j)} = j n^{(j-1)} \text{ при } j \geq 1. \quad (3.54)$$

Это соотношение в анализе аналогично равенству $(x^j)' = jx^{j-1}$ (роль x^j в данном случае играет $n^{(j)}$).

Доказательство.

$$\Delta n^{(i)} = (n+1)n(n-1)\dots(n-j+2) - n(n-1)\dots(n-j+2)(n-j+1) = \\ = jn(n-1)\dots(n-j+2) = jn^{(i-1)}.$$

Пусть $f(x)$ — многочлен от x степени k . Воспользовавшись разложением Тейлора (3.49), можно представить $f(i+n)$ в виде многочлена от n степени k . Замечая, что мультипликативная функция $n^{(j)}$ является многочленом степени j по n , имеем

$$f(i+n) = \sum_{j=0}^k b_j n^{(j)} = b_0 + b_1 n^{(1)} + \dots + b_k n^{(k)}. \quad (3.55)$$

Вычисляя последовательно приращения, получаем

$$\Delta f(i+n) = b_1 n^{(0)} + 2b_2 n^{(1)} + \dots + kb_k n^{(k-1)};$$

$$\Delta^2 f(i+n) = 2!b_2 n^{(0)} + \dots + k(k-1)b_k n^{(k-2)};$$

.....

$$\Delta^k f(i+n) = k!b_k n^{(0)}.$$

Положим в этих равенствах $n=0$. Тогда

$$\Delta^j f_i = j!b_j, \text{ т. е. } b_j = \frac{\Delta^j f_i}{j!}, \quad j=1, \dots, k. \quad (3.56)$$

Подставляя эти выражения в (3.55), получаем

$$f_{i+n} = \sum_{j=0}^k \frac{\Delta^j f_i}{j!} n^{(j)} = \sum_{j=0}^k \binom{n}{j} \Delta^j f_i. \quad (3.57)$$

Эта важная формула соответствует разложению в ряд Тейлора (3.49).

Упражнение 3.10. Покажите, что если корень $\alpha \neq 0$ многочлена r -й степени $f(z) = z^r + a_1 z^{r-1} + \dots + a_r$ имеет кратность l , то

$$r^{(k)} \alpha^r + (r-1)^{(1k)} a_1 \alpha^{r-1} + \dots + 0^{(k)} a_r \begin{cases} = 0, & k=0, 1, \dots, l-1; \\ \neq 0, & k=l. \end{cases}$$

Решение. Имеем $f(\alpha) = f'(\alpha) = \dots = f^{(l-1)}(\alpha) = 0$; $f^{(l)}(\alpha) \neq 0$.

В общем случае $f^{(k)}(\alpha) = r(r-1)\dots(r-k+1)\alpha^{r-k} + a_1(r-1)\dots$

$\dots(r-k)\alpha^{r-1-k} + \dots + a_k k(k-1)\dots 1 \cdot \alpha^{k-k} = r^{(k)} \alpha^{r-k} + (r-1)^{(k)} \times$

$\times a_1 \alpha^{r-1-k} + \dots + k^{(k)} a_k \alpha^0$.

Однако в силу (3.53) выражения $(k-1)^{(k)}$, ..., $0^{(k)}$ все равны нулю и, следовательно, в данном случае

$$\alpha^k f^{(k)}(\alpha) = r^{(k)} \alpha^r + (r-1)^{(k)} a_1 \alpha^{r-1} + \dots + 1^{(k)} a_{r-1} \alpha + 0^{(k)} a_r.$$

3.5. Нахождение частного решения

В § 3.3 было показано, как можно найти частное решение линейного рекуррентного уравнения с постоянными коэффициентами ($a_r \neq 0$):

$$f_{n+r} + a_1 f_{n+r-1} + \dots + a_r f_n = \Phi(n), \quad n = 0, 1, 2, \dots, \quad (3.58)$$

для некоторых специальных функций $\Phi(n)$. В остальных случаях частное решение можно найти, если воспользоваться общими формулами для приращений, полученными в § 3.4.

Вначале найдем частное решение в случае, когда

$$\Phi(n) = \alpha^n \quad (3.59)$$

и $\alpha (\neq 0)$ является характеристическим корнем кратности l . Если α не является характеристическим корнем, то, как следует из § 3.3, частное решение имеет вид $c\alpha^n$. Если же α — характеристический корень, то это выражение частным решением не является. Подставляя $c\alpha^n$ в уравнение, получаем $c\alpha^n (\alpha^r + a_1 \alpha^{r-1} + \dots + a_r) = \alpha^n$ (здесь мы воспользовались тем, что выражение в скобках в левой части равно нулю). Частное решение в рассматриваемом случае имеет вид

$$f_n = cn^{(l)} \alpha^n, \quad (3.60)$$

где $n^{(l)}$ — мультипликативная функция, введенная в предыдущем параграфе. Подставляя (3.60) в (3.58), находим

$$c\alpha^n [(n+r)^{(l)} \alpha^r + (n+r-1)^{(l)} a_1 \alpha^{r-1} + \dots + n^{(l)} a_r] = \alpha^n. \quad (3.61)$$

Следовательно, если показать, что выражение в скобках в левой части — отличная от нуля константа, то, взяв в качестве c значение, обратное этой константе, можно утверждать, что (3.60) является частным решением.

Пусть $h(n)$ — многочлен степени не выше l . Положим

$$h(n) = (n+r)^{(l)} \alpha^r + (n+r-1)^{(l)} a_1 \alpha^{r-1} + \dots + n^{(l)} a_r. \quad (3.62)$$

Из упражнения 3.10 следует, что

$$\Delta^0 h(0) = h(0) = r^{(l)} \alpha^r + (r-1)^{(l)} a_1 \alpha^{r-1} + \dots + 0^{(l)} a_r \neq 0.$$

Далее, вычисляя последовательно приращения относительно (3.62) и полагая $n=0$, получаем

$$\Delta h(0) = l[r^{(l-1)} \alpha^r + (r-1)^{(l-1)} a_1 \alpha^{r-1} + \dots + 0^{(l-1)} a_r];$$

$$\Delta^2 h(0) = l^{(2)} [r^{(l-2)} \alpha^r + (r-1)^{(l-2)} a_1 \alpha^{r-1} + \dots + 0^{(l-2)} a_r];$$

.....

$$\Delta^{(l)} h(0) = l^{(l)} [r^{(0)} \alpha^r + (r-1)^{(0)} a_1 \alpha^{r-1} + \dots + 0^{(0)} a_r].$$

Все эти выражения, как показывает упражнение 3.10, равны нулю. Следовательно, в силу (3.57),

$$h(n) = \sum_{j=0}^l \binom{n}{j} \Delta^{(j)} h(0) = h(0).$$

Поэтому равенство (3.61) переходит в равенство

$$ch(0) \alpha^n = \alpha^n.$$

Таким образом, если положить $c=1/h(0)$, то (3.60) будет частным решением. При этом c задается формулой

$$c = [r^{(l)} \alpha^r + (r-1)^{(l)} a_1 \alpha^{r-1} + \dots + 0^{(l)} a_r]^{-1}. \quad (3.63)$$

Упражнение 3.11. Найдите частное решение уравнения

$$f_{n+2} + 2f_{n+1} + f_n = (-1)^n.$$

Решение. Так как -1 является характеристическим корнем кратности 2, то искомое частное решение будет иметь вид

$$c n^{(2)} (-1)^n = cn(n-1)(-1)^n.$$

Воспользовавшись формулой (3.63), постоянную c можно определить следующим образом: $c = [2^{(2)}(-1)^2 + 1^{(2)} \cdot 2 \cdot (-1)^1 + 0^{(2)} \cdot 1]^{-1} = [2 \cdot 1 + 0 + 0]^{-1} = 1/2$. Таким образом, искомое частное решение

$$\frac{n(n-1)}{2} (-1)^n.$$

Пусть функция $\varphi(n)$ является многочленом от n степени k и 1 — характеристический корень кратности l . При этом тождество из упражнения 3.10 переходит в следующее ($\alpha=1$):

$$r^{(j)} + (r-1)^{(j)} a_1 + (r-2)^{(j)} a_2 + \dots + 0^{(j)} a_r \begin{cases} = 0, & j = 0, 1, \dots, l-1; \\ \neq 0, & j = l. \end{cases} \quad (3.64)$$

Если при $j \geq 1$ подставить

$$f_n = cn^{(j)} \quad (3.65)$$

в левую часть (3.58), то она будет иметь вид

$$c[(n+r)^{(j)} + (n+r-1)^{(j)} a_1 + \dots + n^{(j)} a_r].$$

Многочлен в скобках $h(n)$ имеет степень самое большее j . Вычисляя последовательно приращения, находим

$$\Delta^0 h(0) = h(0) = r^{(j)} + (r-1)^{(j)} a_1 + \dots + 0^{(j)} a_r;$$

$$\Delta h(0) = j^{(1)} [r^{(j-1)} + (r-1)^{(j-1)} a_1 + \dots + 0^{(j-1)} a_r];$$

.....

$$\Delta^j h(0) = j^{(j)} [r^{(0)} + (r-1)^{(0)} a_1 + \dots + 0^{(0)} a_r].$$

Если $j \geq l$, то из (3.64) имеем

$$\Delta^j h(0) = \Delta^{j-1} h(0) = \dots = \Delta^{j-l+1} h(0) = 0;$$

$$\Delta^{j-l} h(0) \neq 0.$$

Отсюда и из (3.57) следует, что $h(n)$ является многочленом от n степени $(j-l)$. Если $j < l$, то $\Delta^j h(0) = \Delta^{j-1} h(0) = \dots = \Delta^0 h(0) = 0$ и, следовательно, $h(n) \equiv 0$.

Таким образом, мы показали, что если $\varphi(n)$ является многочленом степени k и 1 — характеристический корень кратности l , то частным решением будет многочлен степени $k+l$ следующего вида:

$$f_n = c_0 n^{k+l} + c_1 n^{k+l-1} + \dots + c_k n^l. \quad (3.66)$$

Упражнение 3.12. Найдите решение уравнения $f_{n+1} - f_n = n^2 + 2n + 1$ с граничным условием $f_0 = 0$.

Решение. Так как 1 является характеристическим корнем, ищем частное решение в виде $f_n = c_0 n^3 + c_1 n^2 + c_2 n$. Подставляя это выражение в уравнение, получаем

$$c_0 (n+1)^3 + c_1 (n+1)^2 + c_2 (n+1) - c_0 n^3 - c_1 n^2 - c_2 n = 3 c_0 n^2 + (3 c_0 + 2 c_1) n + (c_0 + c_1 + c_2) = n^2 + 2 n + 1.$$

Приравнявая постоянные коэффициенты, находим, что $c_0 = 1/3$, $c_1 = 1/2$, $c_2 = 1/6$. Складывая частное решение с общим решением однородного уравнения $c(1)^n = c$, получаем

$$f_n = \frac{1}{3} n^3 + \frac{1}{2} n^2 + \frac{1}{6} n + c.$$

Из граничного условия следует, что $c = 0$. Таким образом,

$$f_n = (2 n^3 + 3 n^2 + n)/6 = n(n+1)(2n+1)/6.$$

Это выражение совпадает с правой частью формулы (3.2).

3.6. Приложения к задачам, связанным с периодическими структурам

Упражнение 3.13. Найдём значение f_n следующего определителя порядка n :

$$\begin{vmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & 1 & \dots & 0 \\ 0 & \dots & 1 & 1 & \dots & 0 \\ 0 & \dots & 0 & 1 & 1 & 1 \\ 0 & \dots & 0 & 0 & 1 & 1 \end{vmatrix}.$$

Решение. Если разложить определитель по первому столбцу, то получим

$$f_n = f_{n-1} - \begin{vmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & 1 & \dots & 0 \\ 0 & \dots & 1 & 1 & \dots & 0 \\ 0 & \dots & 0 & 1 & 1 & 1 \\ 0 & \dots & 0 & 0 & 1 & 1 \end{vmatrix}.$$

Вновь разлагая определитель по 1-му столбцу, имеем $f_n = f_{n-1} - f_{n-2}$. Это можно записать в виде рекуррентного уравнения $f_{n+2} - f_{n+1} + f_n = 0$, $n = 0, 1, \dots$ Так как $f_1 = 1$, $f_2 = 0$, то, полагая $f_0 = 1$, приходим к рекуррентному уравнению, выполняющемуся при всех $n = 0, 1, 2, \dots$ Задаваемая им последовательность точно совпадает с последовательностью чисел, найденной в упражнении 3.4.

Следовательно, $f_n = \cos \frac{\pi}{3} n + \frac{1}{\sqrt{3}} \sin \frac{\pi}{3} n$.

Упражнение 3.14. Предположим, что при размещении без повторения $n \geq 2$ чисел $1, 2, \dots, n$ на i -ю позицию должно быть помещено обязательно одно из чисел, входящих в i -й столбец следующей таблицы:

$$\begin{array}{ccccccc} 1 & 2 & \dots & n-3 & n-2 & n-1 & \\ 1 & 2 & 3 & \dots & n-2 & n-1 & n \\ 2 & 3 & 4 & \dots & n-1 & n & \end{array}$$

Найдите общее число $f(n)$ таких размещений.

Решение. Число n должно быть взято либо из n -го, либо из $(n-1)$ -го столбца приведенной в задаче таблицы. Если n взять из n -го столбца, то оставшиеся числа 1, 2, ..., $n-1$ должны выбираться согласно следующей таблице:

1 2 ... $n-3$ $n-2$
 1 2 3 ... $n-2$ $n-1$
 2 3 4 ... $n-1$

Поэтому общее число размещений, при которых число n берется из n -го столбца, равно f_{n-1} .

Далее, если число n берется из $(n-1)$ -го столбца, то n -ю позицию автоматически занимает число $(n-1)$. Следовательно, оставшиеся $n-2$ чисел 1, 2, ..., $n-2$ должны выбираться согласно следующей таблице:

1 2 ... $n-3$
 1 2 3 ... $n-2$
 2 3 4 ...

Это означает, что число размещений рассматриваемого типа равно f_{n-2} .

Таким образом, мы получаем следующее рекуррентное уравнение: $f_{n+2} - f_{n+1} - f_n = 0$, $n=2, 3, \dots$ Легко проверить, что $f_2=2$ и $f_3=3$. Если положить, кроме того, $f_0=f_1=1$, то приведенное рекуррентное уравнение будет выполняться при всех $n=0, 1, 2, \dots$ Следовательно, искомая последовательность чисел является последовательностью чисел Фибоначчи, рассматривавшейся в упражнениях 3.2 и 3.3.

Упражнение 3.15. Найдите напряжение U_N на входе показанной на рис. 3.1 схемы, по которой протекает постоянный ток. Известно, что напряжение на ее самом правом резисторе равно $U_0=1$ В.

Решение. Обозначим напряжения и токи в звеньях рассматриваемой схемы так, как показано на рис. 3.2. При этом $I_{n+1} = \frac{1}{4}(U_{n+2} - U_{n-1})$; $I_n = \frac{1}{4}(U_{n+1} - U_n)$. Подставляя эти выражения в формулу $I_{n+1} = I_n + \frac{1}{2}U_{n+1}$, получаем $\frac{1}{4}U_{n+2} - U_{n+1} + \frac{1}{4}U_n = 0$, т. е. $U_{n+2} - 4U_{n+1} + U_n = 0$, $n=0, 1, 2, \dots$

Характеристический многочлен $z^2 - 4z + 1$ имеет два простых корня $2 \pm \sqrt{3}$, а поэтому общим решением в данном случае будет

$$U_n = c_1(2 + \sqrt{3})^n + c_2(2 - \sqrt{3})^n.$$

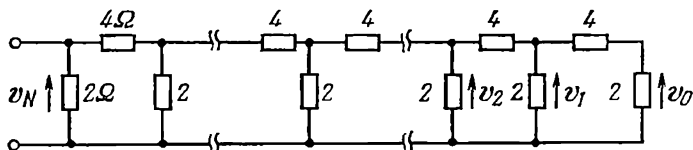


Рис. 3.1. Электрическая схема из упражнения 3.15

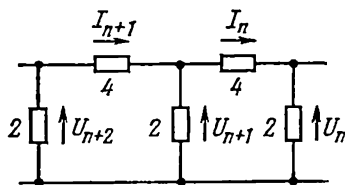


Рис. 3.2. Связь напряжений и токов

Для определения констант c_1 и c_2 необходимо воспользоваться заданным значением $U_0=1$ В, а кроме того, значением U_1 , которое можно легко вычислить. Действительно, как видно из рис. 3.1, через крайний правый резистор протекает ток $U_0/2$, и, следовательно, $U_1=6U_0/2=3U_0=3$ В.

Отсюда, в свою очередь, получаем, что $c_1+c_2=1$, $c_1(2+\sqrt{3})+c_2(2-\sqrt{3})=3$, т. е. $c_1=\frac{1}{2}(1+\frac{1}{\sqrt{3}})$, $c_2=\frac{1}{2}(1-\frac{1}{\sqrt{3}})$.

Таким образом, искомое напряжение на входе схемы

$$U_N = \frac{1}{2} \left(1 + \frac{1}{\sqrt{3}} \right) (2 + \sqrt{3})^N + \frac{1}{2} \left(1 - \frac{1}{\sqrt{3}} \right) (2 - \sqrt{3})^N.$$

Задачи

3.1. Найдите решение уравнения $u_{n+2}-2u_{n+1}-3u_n=1$ с граничными условиями $u_0=u_1=1$.

3.2. Сколько размещений с повторениями четырех букв, a , b , c и d по n позициям содержат букву a четное число раз?

3.3. На плоскости имеется круг и n прямых. Известно, что все прямые пересекаются внутри круга и что никакие три прямые не пересекаются в одной точке. На сколько частей прямые делят внутренность круга?

3.4. Найдите частное решение уравнения $\Delta^r f_n = n^{(k)}$ и общее решение однородного уравнения $\Delta^r f_n = 0$.

3.5. Найдите значение определителя порядка n :

$$\begin{vmatrix} 4 & 2 & 0 & 0 & \dots & 0 \\ 2 & 4 & 2 & 0 & \dots & 0 \\ 0 & 2 & 4 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 2 & 4 & 2 & 0 \\ 0 & \dots & 0 & 2 & 4 & 2 \\ 0 & \dots & 0 & 0 & 2 & 4 \end{vmatrix}$$

Глава 4

Графы

4.1. Понятие графа

Граф можно представлять себе как конфигурацию, подобную тем, что изображены на рис. 4.1,а и, представляющую собой совокупность некоторых линий, соединяющих между собой некоторые точки. Точки называются *вершинами*, а линии — *ребрами*. На рис. 4.1,а имеется ребро, которое начинается и заканчивается в одной и той же вершине; такие ребра называют *петлями*. Однако далее будут рассматриваться графы, не содержащие петель. Некоторые пары различных вершин графа могут быть соединены несколькими ребрами, как это изображено, например, на

рис. 4.1, б; такие графы называют графами с кратными ребрами. Если любые две различные вершины графа соединены не более чем одним ребром, то такой граф часто называют *линейным*. Если это в данной книге не оговорено особо, то рассматриваются произвольные графы с кратными ребрами.

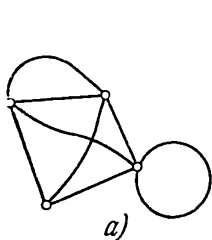


Рис. 4.1. Графы:
а — граф с петлей; б — параллельные ребра

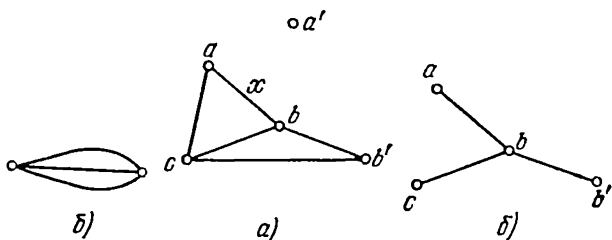


Рис. 4.2. Примеры графов:
а — граф; б — истинный подграф графа

Математически граф G определяется двумя множествами (V , E), где V — конечное множество вершин, а E — некоторое множество неупорядоченных пар вершин (a, b) . Неупорядоченная пара вершин (a, b) называется *ребром*, соединяющим две различные вершины a и b . Так как порядок вершин не играет роли, то наряду с записью (a, b) допустима и запись (b, a) . На рис. 4.2, а ребро x соединяет вершины a и b ; эти вершины называют концевыми точками или концами ребра x . При этом также говорят, что ребро x инцидентно вершинам a и b , или что вершины a и b являются смежными. Совокупность ребер, инцидентных вершине a , называется *звездой* вершины a и обозначается $\delta(a)$. Наиболее важное соотношение теории графов

$$|E| = \frac{1}{2} \sum_{a \in V} \deg(a). \quad (4.1)$$

Оно утверждает: число ребер в графе равно половине суммы степеней его вершин. Это равенство следует из того, что в графе нет петель и каждое ребро учитывается в сумме правой части равенства дважды (так как оно инцидентно ровно двум различным вершинам).

Упражнение 4.1. Покажите, что в любом графе число вершин нечетной степени четно.

Решение. Сумма в правой части равенства (4.1) является четным числом.

Вершина a называется *изолированной вершиной*, если $\delta(a) = \emptyset$, т. е. если в графе нет ребер, которые были бы инцидентны a . Вершина a называется *концевой*, если $|\delta(a)| = 1$, т. е. если имеется ровно одно ребро, которое ей инцидентно. Например, вершина a' на рис. 4.2, а — изолированная, а вершина b' на рис. 4.2, б — концевая.

Граф $G'=(V', E')$ называется подграфом графа $G=(V, E)$, если $V' \subseteq V$ и $E' \subseteq E$. Одним из подграфов графа G является сам граф G . Граф, который не имеет ни одной вершины и ни одного ребра, называется *пустым*. Пустой граф является подграфом произвольного графа G . Подграф графа G , не совпадающий с самим графом G и не являющийся пустым, называется *собственным* подграфом графа G . Граф, изображенный на рис. 4.2,б, является собственным подграфом графа, изображенного на рис. 4.2,а.

Путь длины l называется последовательность

$$a_0, x_1, a_1, x_2, \dots, a_{l-1}, x_l, a_l, \quad (4.2)$$

в которой ребро $x_i (i=1, \dots, l)$ соединяет вершины a_{i-1} и a_i . Если $a_0 \neq a_l$, то путь называется *незамкнутым* (открытым). Если же $a_0 = a_l$, то путь называется *замкнутым* или *циклическим*. Путь (4.2) называется *простым*, если все его ребра x_1, \dots, x_l различны. Открытый путь, в котором все вершины a_0, \dots, a_l различны, называется *цепью*, соединяющей вершины a_0 и a_l , или просто $(a_0 - a_l)$ -*цепью*. Замкнутый простой путь, все вершины a_0, a_1, \dots, a_l которого различны, называется *простым циклом* (или просто *циклом*). На рис. 4.3 приведено несколько примеров введенных понятий.

Если задан путь (4.2), соединяющий вершины a_0 и a_l ($a_0 \neq a_l$), то выбором некоторого подмножества ребер из совокупности x_1, \dots, x_l можно построить цепь. Это можно доказать, например, следующим образом. Среди всех путей, соединяющих вершины a_0 и a_l , множество ребер которых является подмножеством совокупности x_1, \dots, x_l , выберем самый короткий путь (содержащий наименьшее число ребер). Пусть это будет путь

$$a_0 = a'_0, x'_1, a'_1, \dots, a'_{k-1}, x'_k, a'_k = a_l.$$

Покажем, что вершины a'_0, a'_1, \dots, a'_k этого пути различны, т. е. он представляет собой $(a_0 - a_l)$ -цепь. Действительно, если $a'_i = a'_j (i < j)$, то вместо выбранного пути длины k можно взять путь

$$a_0 = a'_0, x'_1, \dots, a'_i, x'_{j+1}, a'_{j+1}, \dots, a'_k = a_l,$$

который также соединяет вершины a_0 и a_l , но имеет длину $k - (j - i) < k$. Это противоречит предположению о минимальности k . Таким образом, мы доказали следующую теорему.

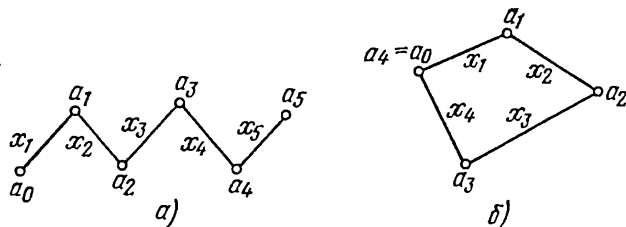


Рис. 4.3. Цепи и циклы:
а — $(a_0 - a_5)$ -цепь; б — цикл

Теорема 4.1. Если существует путь, соединяющий две различные вершины a и b , то из ребер этого пути можно построить $(a-b)$ -цепь.

Следующая теорема является аналогом предыдущей для замкнутых путей.

Теорема 4.2. Если существует замкнутый простой путь, то из его ребер можно составить цикл.

Доказательство. Предположим, что последовательность (4.2) является замкнутым простым путем, т. е. $a_0 = a_l$. Если все вершины a_0, a_1, \dots, a_{l-1} различны, то этот путь будет циклом. В противном случае из совокупности его ребер выберем подмножество, которое образует замкнутый путь минимальной длины, и возьмем его в качестве последовательности (4.2). Можно показать, что все вершины a_0, a_1, \dots, a_{l-1} этого пути различны, а следовательно, он является циклом. Действительно, если какая-либо из вершин входит в совокупность a_0, a_1, \dots, a_{l-1} дважды, то найдется целое число k ($1 < k < l-1$), обладающее следующими свойствами: все вершины a_0, a_1, \dots, a_k различны и a_{k+1} совпадает с одной из вершин a_0, a_1, \dots, a_k . Пусть $a_i = a_{k+1}$ ($0 \leq i \leq k$). Тогда путь $a_i, x_{i+1}, \dots, a_k, x_{k+1}, a_{k+1}$ будет замкнутым простым путем, длина которого $k+1-i \leq l-1$, т. е. меньше длины рассматриваемого пути (4.2), т. е. получено противоречие.

Обе приведенные теоремы кажутся очевидными. Но при рассмотрении несколько более сложных свойств необходимо пользоваться стандартными методами доказательств, принятыми в теории графов, например подобными использованным ранее. Так, если из формулировки теоремы 4.2 исключить требование простоты пути, то теорема будет неверна. Примером, показывающим это, является последовательность a_0, x, a_1, x, a_0 , в которую ребро x входит 2 раза и которая, следовательно, простой не является. Очевидно, что только из ребра x построить цикл невозможно. Если же путь a_0, x_1, a_1, x_2, a_0 является простым, то, очевидно, два его параллельных ребра, x_1 и x_2 ($x_1 \neq x_2$), образуют цикл.

Теорема 4.3. Пусть $G = (V, E)$ — граф, все вершины которого имеют четную степень. Тогда для произвольного его ребра x_1 в графе существует замкнутый простой путь, содержащий это ребро x_1 .

Доказательство. Пусть a_0, a_1 — концевые точки ребра x_1 . Так как степень a_1 — четное число, то найдется другое ребро, отличное от x_1 и инцидентное a_1 . Пусть таким ребром будет x_2 и a_2 — его концевая точка, отличная от a_1 . Путь a_0, x_1, a_1, x_2, a_2 является простым. В общем случае пусть уже построенный простой путь заканчивается в вершине, отличной от a_1 . Существует инцидентное этой вершине ребро, которое ранее не использовалось, и наш путь можно продлить еще на одно ребро. Строя таким образом последовательность $a_0, x_1, a_1, x_2, a_2, x_3, a_3, \dots$, мы получаем последовательность простых путей увеличивающейся длины. Поскольку общее число ребер в графе конечно, то через некоторое время продлить путь уже будет невозможно. Это зна-

чит, что путь заканчивается в вершине a_0 . Это доказывает существование замкнутого простого пути, содержащего ребро x_1 .

Если удлинить простой путь так, как это делалось при доказательстве последней теоремы, строя последовательность $a_0, x_1, a_1, x_2, a_2, \dots$, то некоторые из вершин могут входить в путь 2 и более раз. Предположим, что вершины a_0, a_1, \dots, a_{k-1} все различны, а a_k совпадает с вершиной a_i ($0 \leq i \leq k-1$). Как видно из рис. 4.4, ребра x_{i+1}, \dots, x_k образуют цикл. Обозначим через E' множество ребер, которое получается при удалении из множества ребер E графа $G = (V, E)$ совокупности $\{x_{i+1}, \dots, x_k\}$. Граф $G' = (V, E')$, как и исходный, обладает тем свойством, что все его вершины имеют четную степень. Следовательно, если E' не является пустым множеством, то для каждого отдельного ребра описанный процесс можно повторить. Выполняя эту процедуру достаточное число раз, приходим к следующей теореме.

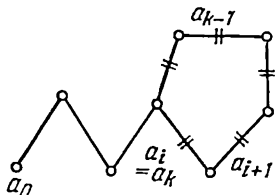


Рис. 4.4. Выделение цикла

Теорема 4.4. Если степени всех вершин графа $G = (V, E)$ четны и множество его ребер E не пусто, то E совпадает с объединением множеств ребер некоторого числа циклов, никакие два из которых не имеют общих ребер.

Эта теорема имеет следствие, усиливающее теорему 4.3.

Следствие. Пусть $G = (V, E)$ — граф, все вершины которого имеют четную степень. Для произвольного ребра x графа G существует цикл в G , содержащий x .

Упражнение 4.2. Покажите, что если в графе существуют две различные цепи, P_1 и P_2 , содержащие одни и те же различные вершины a и b , то из некоторого подмножества множества ребер P_1 и P_2 можно построить цикл.

Решение. Пусть V' — множество вершин, входящих в P_1 и P_2 , E_1 — множество ребер P_1 и E_2 — множество ребер P_2 . Обозначим через E' симметрическую разность E_1 и E_2 , т. е. положим $E' = E_1 \oplus E_2 = (E_1 - E_2) \cup (E_2 - E_1)$. Заметим, что E' не пусто (так как цепи P_1, P_2 различны).

Покажем, что граф (V', E') удовлетворяет условиям теоремы 4.4, т. е. что все его вершины имеют четную степень. Из этого будет следовать, что некоторое подмножество множества E' образует простой цикл.

Пусть x_1 и y_1 — ребра из P_1 и P_2 соответственно, которые инцидентны общей конечной вершине a этих цепей. Возможны два случая: $x_1 \neq y_1$ и $x_1 = y_1$. В первом оба ребра, x_1 и y_1 , принадлежат E' , так что степень a равна 2. Во втором E' не содержит ребер, инцидентных a , так что степень a равна 0. Аналогично можно показать, что другая общая конечная вершина цепей b также имеет четную степень. Рассмотрим далее промежуточные вершины цепей. Если промежуточная вершина c принадлежит только одной из цепей P_1, P_2 , то, очевидно, число ребер в E' , инцидентных c , равно 2. Предположим, что c принадлежит обоим цепям, P_1 и P_2 .

Пусть x, x' — ребра P_1 , инцидентные c , а y, y' — ребра P_2 , инцидентные c . Если эти ребра не совпадают, то они принадлежат E' и, следовательно, число инцидентных c ребер равно 4. Если же некоторые два ребра совпадают, например $x = y$, то E' будут принадлежать только ребра x' и y' и в E' число

инцидентных с ребер будет равно 2. При полном совпадении, т. е. когда $x=y$ и $x'=y'$, E' не принадлежит ни одно ребро и степень s равна 0. Таким образом, степень s всегда является четным числом.

4.2. Связность

Граф $G=(V, E)$ называется *связным*, если между произвольными различными двумя его вершинами существует цепь. Как следует из теоремы 4.1, вместо цепи можно также говорить о пути.

Пусть S — непустое подмножество множества вершин V графа $G=(V, E)$, а $E(S)$ — совокупность ребер графа, обе концевые точки которых принадлежат S . Подграф $(S, E(S))$ графа G называется подграфом, порожденным S .

Предположим, что задан граф $G=(V, E)$. На множестве вершин этого графа определим отношение \equiv , считая, что $a \equiv b$, если $a=b$, или в графе существует цепь, соединяющая a и b . Нетрудно видеть, что \equiv является отношением эквивалентности. Действительно, $a \equiv a$ и отношение \equiv рефлексивно. Если $a \equiv b$, то $b \equiv a$, и отношение \equiv симметрично. Если $a \equiv b$ и $b \equiv c$, то в графе существуют как $(a-b)$ -цепь, так и $(b-c)$ -цепь. Если эти цепи соединить в вершине b , то, очевидно, получится путь, соединяющий вершины a и c . Следовательно, $a \equiv c$, т. е. закон транзитивности здесь также выполняется.

Пусть V_1, \dots, V_k — классы эквивалентности множества вершин V графа $G=(V, E)$, определяемые отношением эквивалентности \equiv (см. § 1.4). Подграфы $(V_1, E(V_1)), \dots, (V_k, E(V_k))$ графа G называются связными компонентами этого графа, а их число k называется *степенью связности* (рис. 4.5). Очевидно, что степень связности связного графа равна 1 и единственная связная компонента совпадает с самим графом. Если $a \in V_i$, $b \in V_j$ и $i \neq j$, то $a \not\equiv b$ и в графе G нет ребер, соединяющих вершины a и b . Таким образом, каждая вершина может принадлежать одной и только одной связной компоненте. Аналогично каждое ребро графа принадлежит ровно одной связной компоненте.

Пусть x — некоторое ребро графа $G=(V, E)$. Графом $G-x$ называется граф $(V, E-\{x\})$. Другими словами, граф $G-x$ получается путем удаления из G ребра x (обе концевые точки x в графе остаются).

Теорема 4.5. Если в связном графе $G=(V, E)$ ребро x принадлежит некоторому циклу, то граф $G-x$ также является связным.

Доказательство. Рассмотрим цикл $a_0, x, a_1, y_2, a_2, \dots, y_l, a_l(a_0=a_l)$, содержащий ребро x . Покажем, что если удалить из графа ребро x , то между любыми двумя вершинами будет существовать некоторый путь.

Рассмотрим произвольный путь, соединяющий две вершины графа G . Очевидно, что если часть пути a_0, x, a_1 заменить на $a_0=a_l, y_l, \dots, a_2, y_2, a_1$, то вновь получится путь, соединяющий эти вершины.

Верна также и обратная теорема.

Теорема 4.6. Пусть $G=(V, E)$ — связный граф, и x — его ребро (рис. 4.6). Если граф $G-x$ является связным, то в G существует цикл, содержащий x .

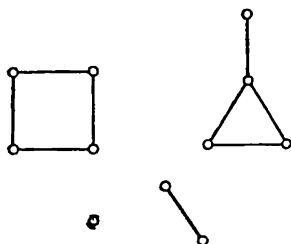


Рис. 4.5. Пример графа со степенью связности 4

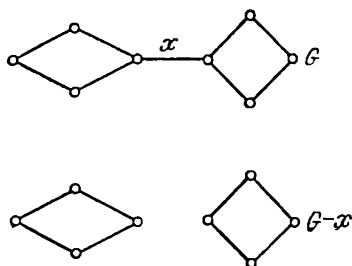


Рис. 4.6. Исключение ребра

Доказательство. Рассмотрим концы a, b ребра x , а также цепь в графе $G-x$: $a=a_0, y_1, a_1, \dots, y_l, a_l=b$, которая соединяет a и b . Построенный по этой цепи путь $a=a_0, y_1, a_1, \dots, y_l, a_l=b, x, a$ в G является циклом.

Ребро x графа $G=(V, E)$ называется *мостом* или *разделяющим ребром*, если связность $G-x$ больше связности G . Если граф G не является связным, то при удалении ребра x видоизменяется лишь связная компонента G , которая содержит x . Отсюда и из теорем 4.5 и 4.6 следует необходимое и достаточное условие того, что некоторое ребро является мостом.

Теорема 4.7. Ребро графа является мостом тогда и только тогда, когда в графе нет циклов, содержащих это ребро.

Упражнение 4.3. Найдите все мосты графа, изображенного на рис. 4.7.

Решение. Мосты рассматриваемого графа — ребра x_4, x_8, x_9 .

Разомкнутой (замкнутой) эйлеровой цепью графа называется простой разомкнутый (замкнутый) путь, включающий все ребра графа.

Теорема 4.8. Граф имеет замкнутую эйлерову цепь тогда и только тогда, когда а) он является связным и б) все его вершины имеют положительные четные степени.

Доказательство. Так как необходимость очевидна, то покажем только достаточность приведенных условий. Согласно теореме 4.3, в рассматриваемом графе $G=(V, E)$ имеется замкнутый простой путь. Рассмотрим замкнутый простой путь

$$a_0, x_1, a_1, \dots, a_l, \dots, x_l, a_l = a_0, \quad (4.3)$$

содержащий наибольшее число ребер. Допустив, что в графе имеются ребра, не принадлежащие этому пути, получим проти-

воречие. Покажем сначала, что в графе нет вершин, не лежащих на этом пути.

Предположим, что некоторая вершина b не лежит на рассматриваемом пути. Так как граф является связным, то существуют цепи между b и каждой из вершин a_0, \dots, a_{l-1} . Рассмотрим наибо-

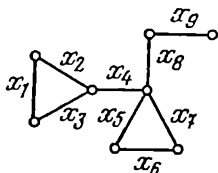


Рис. 4.7. Пример графа

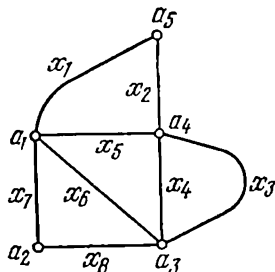


Рис. 4.8. Граф из упражнения 4.4

лее короткую цепь $b = b_0, y_1, b_1, \dots, y_m, b_m = a_i$. Очевидно, что вершины b_0, \dots, b_{m-1} не могут совпадать с вершинами a_0, \dots, a_{l-1} . Следовательно, ребро y_m , соединяющее b_{m-1} и b_m , не входит в путь (4.3). Из множества E ребер графа удалим ребра x_1, \dots, x_l , а полученное в результате множество ребер обозначим через E' . В графе (V, E') все вершины имеют четную степень. Отсюда и из теоремы 4.3 следует, что в этом графе существует простой замкнутый путь, включающий ребро y_m . Пусть это будет путь

a_i, y_m, \dots, a_i . (4.4)

Если в (4.3) вместо a_i подставить (4.4), то, очевидно, получится простой замкнутый путь, длина которого будет больше l . Это противоречит тому, что l — максимальная длина простого замкнутого пути в графе.

Таким образом, мы показали, что простой замкнутый путь максимальной длины (4.3) проходит через все вершины графа. Предположим, что он содержит не все ребра графа. Тогда в графе найдется ребро y , инцидентное одной из вершин a_i и в то же время не входящее в путь (4.3). В этом случае точно так же, как ранее, нетрудно прийти к противоречию.

Следствие. Граф имеет открытую эйлерову цепь тогда и только тогда, когда он а) является связным и б) содержит две вершины нечетной степени.

Доказательство. Так как необходимость очевидна, то покажем только достаточность приведенных условий. Пусть a и b — две вершины графа, имеющие нечетную степень. Соединим вершины a и b еще одним дополнительным ребром. После этого все вершины графа будут иметь четную степень. Так как любые две вершины графа соединены цепью, то вершин степени 0 в графе нет. Следовательно, степень всех вершин является положитель-

ным четным числом. По теореме 4.8 в графе существует замкнутая эйлерова цепь. Если из этой цепи удалить ребро x , то получится открытый простой путь, соединяющий вершины a , b и содержащий все ребра исходного графа.

Наряду с эйлеровой цепью существует также понятие гамильтоновой цепи, которая определяется следующим образом. Разомкнутой (замкнутой) гамильтоновой цепью графа называется замкнутая цепь, которая проходит через все вершины графа.

Для того чтобы на графе, удовлетворяющем условиям теоремы 4.8, найти эйлерову цепь, следует воспользоваться алгоритмом, описанным при доказательстве теорем 4.3 и 4.8.

Упражнение 4.4. Найдите замкнутую эйлерову цепь на графе, изображенном на рис. 4.8.

Решение. Начинать построение можно с любого ребра графа. Начнем с a_1, x_1, a_6 . Получаем: $a_1, x_1, a_6, x_2, a_4, x_3, a_3, x_4, a_5, x_5, a_1, x_6, a_3, x_6, a_2, x_7, a_1$. В данном случае мы сразу получили эйлерову цепь. Если в графе остаются ребра, которые нельзя использовать для продолжения имеющегося пути, то следует начинать строить простой замкнутый путь из уже пройденной вершины и инцидентного ей ребра, если только последнее ранее не использовалось. Продолжая этот процесс, в конце концов мы построим эйлерову цепь, содержащую все ребра графа.

4.3. Деревья

Связный граф, не содержащий простых циклов, называется *деревом*; произвольный граф, не содержащий циклов, называется *лесом*. Очевидно, что связные компоненты леса являются деревьями (рис. 4.9).

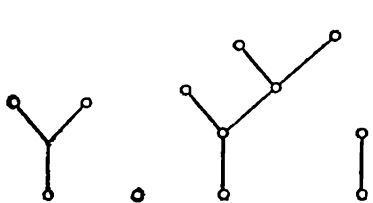
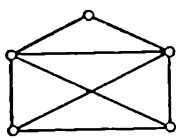
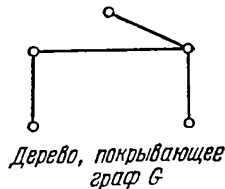


Рис. 4.9. Примеры деревьев



Граф G



Дерево, покрывающее граф G

Рис. 4.10. Граф G и дерево, его покрывающее

Так как дерево, по определению, связно, то любые его две вершины можно соединить цепью. Отсутствие циклов в дереве приводит к тому, что между двумя вершинами дерева имеется только одна цепь (упражнение 4.2). Ясно также, что если между двумя различными вершинами существует ровно одна цепь, то граф является связным и не содержит циклов. Из изложенного следуют теоремы.

Теорема 4.9. Граф является деревом тогда и только тогда, когда между любыми двумя его различными вершинами существует одна и только одна цепь.

Теорема 4.10. Дерево, содержащее две или более вершин, имеет, по крайней мере, две концевые вершины.

Доказательство. Предположим, что граф $G = (V, E)$, имеющий более двух вершин, является деревом и $a_0, x_1, a_1, \dots, x_l, a_l$ представляет собой цепь максимальной длины этого дерева.

Покажем, что конечные вершины этой цепи a_0 и a_l являются концевыми, т. е. что вершинам a_1 и a_l инцидентны соответственно только ребра x_1 и x_l . Доказательство аналогично для обеих вершин, поэтому проведем его только для вершины a_l . Предположим, что существует ребро, отличное от x_l и инцидентное a_l . Пусть это будет ребро x_{l+1} и a_{l+1} — его концевая точка, отличная от a_l . Так как в дереве не может быть циклов, то a_{l+1} не совпадает ни с одной из вершин a_0, a_1, \dots, a_l , т. е. последовательность $a_0, x_1, a_1, \dots, x_l, a_l, x_{l+1}, a_{l+1}$ также является цепью и имеет длину $l+1$. Это противоречит тому, что l — максимальная длина цепи в графе.

Лемма 4.1. После удаления из дерева одной из его концевых вершин вместе с инцидентным ей ребром вновь получится дерево.

Доказательство. Пусть a — концевая вершина дерева, и x — инцидентное ей ребро. Предположим, что ребро x соединяет вершины a и b . Очевидно, граф $(V - \{a\}, E - \{x\})$, который получается удалением из рассматриваемого дерева вершины a и ребра x , не содержит циклов. Также очевидно, что цепь, соединяющая две различные вершины $a', b' \in V - \{a\}$ исходного графа, не содержит ни вершину a , ни ребро x . Следовательно, в исходном графе цепь, соединяющую вершины a' и b' , можно построить только из ребер, принадлежащих множеству $E - \{x\}$. Это означает, что граф $(V - \{a\}, E - \{x\})$ является связным.

Теорема 4.11. Дерево с p вершинами имеет $p-1$ ребер.

Доказательство. При $p=1$ дерево состоит только из одной изолированной вершины и, следовательно, число ребер в нем равно 0. Предположим, что теорема верна при $p=k$. Согласно теореме 4.10, дерево с $k+1$ вершинами имеет концевую вершину. Если исключить из дерева эту концевую вершину вместе с инцидентным ей ребром, то, согласно лемме 4.1, вновь получится дерево, имеющее уже k вершин. По предположению, это дерево имеет $k-1$ ребер. Это означает, что любое дерево с $k+1$ вершинами имеет k ребер. Отсюда и из принципа индукции следует справедливость теоремы.

Пусть $G = (V, E)$ — связный граф. Дерево, являющееся подграфом графа G и содержащее все его вершины, называется деревом, покрывающим граф G (рис. 4.10).

Теорема 4.12. В связном графе всегда существует покрывающее его дерево.

Доказательство. Если в графе нет циклов, то он является деревом и, следовательно, сам себя покрывает. При наличии циклов исключим из графа одно ребро, принадлежащее какому-нибудь циклу. Из теоремы 4.5 следует, что получившийся в результате граф будет связным. Исключая ребра до тех пор, пока

в получающихся графах будут существовать замкнутые циклы, мы приходим к связному подграфу, не содержащему циклов. Этот подграф будет искомым деревом, покрывающим исходный граф.

Следствие 1. Пусть F — подмножество ребер графа $G=(V, E)$ такое, что совокупность ребер ни одного цикла графа не лежит полностью в F . Тогда существует дерево (V, E') , покрывающее G , для которого $F \subset E'$.

Доказательство. Так как в графе нет циклов, состоящих только из ребер множества F , то при доказательстве теоремы 4.12 можно исключать из циклов те ребра, которые не принадлежат множеству F . Это приводит к указанному следствию.

Следствие 2. Связный граф с p вершинами и $(p-1)$ ребрами является деревом.

Доказательство. Если бы такой граф содержал цикл, то при доказательстве теоремы 4.12 после исключения некоторого числа $m(m>0)$ ребер получилось бы дерево, покрывающее заданный граф. Но тогда число вершин было бы равно p , а число ребер $p-1-m < p-1$, что противоречит теореме 4.11.

Следствие 3. Если связный граф имеет p вершин и q ребер; то $q \geq p-1$ или, что то же самое, $q-p+1 \geq 0$.

Доказательство. Так как при доказательстве теоремы 4.12 после исключения некоторого числа $m(m \geq 0)$ ребер получается дерево, то $q-m = p-1$, а следовательно, $q \geq p-1$.

Числа

$$\Phi = p - k, \quad \mu = q - p + k, \quad (4.5)$$

где p — число вершин, q — число ребер, а k — связность графа, называются соответственно *корангом* и *цикломатическим числом*. Пусть $(V_1, E_1), \dots, (V_k, E_k)$ — связные компоненты графа $G=(V, E)$, имеющие соответственно по p_1, \dots, p_k и q_1, \dots, q_k ребер каждая. Коранг и цикломатическое число для каждой связной компоненты определяются равенствами

$$\Phi_i = p_i - 1, \quad \mu_i = q_i - p_i + 1 \quad (i = 1, \dots, k). \quad (4.6)$$

Очевидно,

$$\Phi = \sum_{i=1}^k \Phi_i, \quad \mu = \sum_{i=1}^k \mu_i. \quad (4.7)$$

Из теоремы 4.11, следствий 2 и 3 из теоремы 4.12 и равенств (4.7) следует теорема.

Теорема 4.13. Цикломатическое число графа является целым неотрицательным числом. Граф является деревом тогда и только тогда, когда его цикломатическое число равно 0.

Упражнение 4.5. Покажите, что в графе с цикломатическим числом 1 существует ровно один цикл.

Решение. Если граф является несвязным, то, как следует из равенства (4.7) и теоремы 4.5, одна из его связных компонент имеет цикломатическое число 1, а остальные связные компоненты — число 0. Следовательно, достаточно доказать утверждение упражнения для связного графа $G=(V, E)$ с цикло-

матическим числом 1. Так как граф G не является деревом, то, естественно, в нем существует цикл. Однако пока мы не можем утверждать, что он единственный. Допустим, что существует два различных цикла с множествами ребер C_1 и C_2 соответственно. Предположим, что имеется ребро x такое, что $x \in C_1$, $x \notin C_2$. Если исключить x из G , то получится связный граф с цикломатическим числом 0, все еще содержащий простой цикл C_2 . Получили противоречие.

Рассмотрим связный граф $G=(V, E)$ с p вершинами и q ребрами, а также покрывающее его дерево $T=(V, E')$. Ребра $x \in E'$, принадлежащие дереву, будем называть *ветвями*, а остальные ребра $x \in E-E'$ — *хордами*. Если к дереву T добавить хорду $x \in E-E'$, соединяющую вершины a и b , то x и T образуют граф с циклом, который определяется единственным образом цепью, соединяющей вершины a и b . Этот цикл называется *главным циклом*, определяемым хордой x . Так как имеется $p-1=\mu$ ветвей дерева и $q-p+1=\mu$ хорд, то число главных циклов равно μ (рис. 4.11).

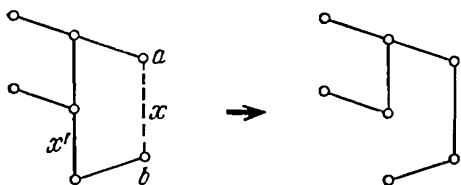


Рис. 4.11. Элементарное преобразование дерева

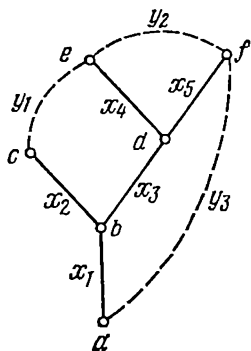


Рис. 4.12. Граф из упражнения 4.6

Упражнение 4.6. Найдите все главные циклы графа, изображенного на рис. 4.11.

Решение:

$y_1: c, y_1, e, x_4, d, x_3, b, x_2, c;$

$y_2: e, y_2, f, x_5, d, x_4, e;$

$y_3: a, y_3, f, x_5, d, x_3, b, x_1, a.$

Обратимся к рис. 4.12 и рассмотрим ветвь x' главного цикла, определяемого хордой x . Если к дереву T добавить хорду x , то единственным простым циклом на полученном графе будет главный цикл, определяемый x . Если из него исключить хорду x' , то простых циклов не будет и образуется другое дерево, покрывающее G . Эта операция называется *элементарным преобразованием дерева*. Суммируя изложенное, получаем следующую теорему.

Теорема 4.14. Предположим, что задан связный граф $G=(V, E)$ и покрывающее его дерево $T=(V, E')$. Пусть $x' \in E'-$

ветвь главного цикла, определяемого хордой $x \in E - E'$ и $E'' = E' \cup \{x\} - \{x'\}$. Тогда граф (V, E'') также является деревом, покрывающим G .

4.4. Коциклы

Рассмотрим подмножество C множества ребер E связного графа $G = (V, E)$ такое, что граф $(V, E - C)$ уже несвязен. Такое минимальное множество называется *коциклом*. Другими словами, C называется коциклом, его граф $(V, E - C)$ несвязен, но для любого собственного подмножества C' множества C граф $(V, E - C')$ связен. Для графа, изображенного на рис. 4.13, множество ребер $\{x_1, x_2, x_3\}$ не является коциклом, а множество $\{x_1, x_2\}$ — коцикл.

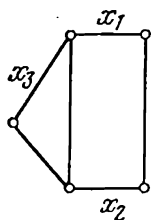


Рис. 4.13. Коцикл $\{x_1, x_2\}$

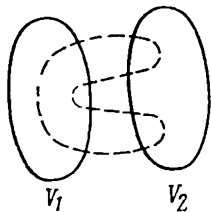


Рис. 4.14. Цикл и коцикл

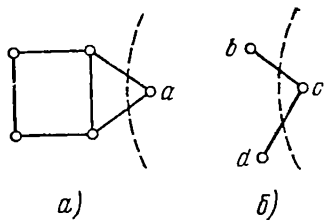


Рис. 4.15. Решение упражнения 4.7

Предположим, что подмножество вершин C связного графа $G = (V, E)$ — коцикл. Граф $(V, E - C)$, по определению, является несвязным. Предположим, что степень связности* этого графа больше 2; обозначим через (V_1, E_1) , (V_2, E_2) , (V_3, E_3) , ... его связные компоненты. Так как исходный граф связен, то найдутся две связные компоненты, (V_i, E_i) , (V_j, E_j) , $i \neq j$, и ребро $x \in C$, которое соединяет одну вершину из V_i и одну вершину из V_j . Для простоты будем считать, что ребро x соединяет одну вершину из V_1 и одну из V_2 . При этом множество $C' = C - \{x\}$ является собственным подмножеством C , но граф $(V, E - C')$ несвязен. В самом деле, вершины из V_3 и вершины из V_1 не связны между собой. Это противоречит минимальности C . Таким образом, мы доказали, что если C — коцикл, то степень связности графа $(V, E - C)$ равна 2.

Пусть C — коцикл связного графа $G = (V, E)$, а (V_1, E_1) и (V_2, E_2) — связные компоненты $(V, E - C)$. В C нет ребер, которые соединяли бы какие-либо две вершины в V_1 или V_2 . Действительно, если $x \in C$ соединяет две вершины в V_1 , то $C' = C - \{x\}$ — собственное подмножество C , а граф $(V, E - C')$ имеет две связные компоненты. Это противоречит минимальности C . Обозначим

* В дальнейшем число связных компонент графа называется степенью связности G . (Прим. ред.)

через $E(V_1, V_2)$ множество ребер из E , каждое из которых соединяет некоторую вершину из V_1 с некоторой вершиной из V_2 . Как следует из изложенного, $C = E(V_1, V_2)$. Таким образом, мы доказали следующую теорему.

Теорема 4.15. Если C — коцикл связного графа $G = (V, E)$, то степень связности $(V, E - C)$ равна 2. Если $(V_1, E_1), (V_2, E_2)$ — связные компоненты графа $(V, E - C)$, то $C = E(V_1, V_2)$.

Следствие. Любой цикл и любой коцикл связного графа имеют четное число общих ребер.

Доказательство. Для произвольного коцикла, так же как и в формулировке теоремы, положим $C = E(V_1, V_2)$. Если цикл проходит только по вершинам V_1 или только по вершинам V_2 , то он не имеет общих ребер с C . Если же цикл проходит как по вершинам V_1 , так и по вершинам V_2 , то он имеет некоторое число общих с C ребер. Однако если цикл выходит из вершины, принадлежащей V_1 , то он должен вернуться вновь в вершину из V_1 , а следовательно, он всегда проходит через четное число ребер из C (рис. 4.14).

Упражнение 4.7. Приведите примеры: а) графа, в котором звезда некоторой вершины была бы коциклом, б) в котором звезда некоторой вершины не является коциклом.

Решение. Звезда $\delta(a)$ графа, изображенного на рис. 4.15, а, является коциклом. Звезда $\delta(c)$ графа, изображенного на рис. 4.15, б, коциклом не является, так как в результате исключения $\delta(c)$ получается граф со степенью связности 3.

Рассмотрим связный граф $G = (V, E)$ и дерево $T = (V, E)$, его покрывающее. Пусть $x \in E'$ — ветвь произвольного дерева. Так как x — мост, то граф $(V, E' - \{x\})$ несвязен. Следовательно, множество $\{x\}$, состоящее только из одного ребра x , является коциклом дерева T . Пусть $(V_1, E'_1), (V_2, E'_2)$ — связные компоненты $(V, E - \{x\})$. Обозначим через E'' множество ребер, которые являются хордами и соединяют вершины из V_1 и V_2 ; $E'' = (E - E') \cap E(V_1, V_2)$. Очевидно, что $\{x\} \cup E''$ — коцикл G . Кроме того, ясно, что E'' можно определить как множество хорд, главные циклы которых включают ветвь x . Этот коцикл $\{x\} \cup E''$ называют главным коциклом, определяемым ветвью x . Так как число ветвей равно φ , то существует φ главных коциклов.

Упражнение 4.8. Найдите все главные коциклы графа, изображенного на рис. 4.12.

Решение:

$x_1 : \{x_1, y_3\}$;

$x_2 : \{x_2, y_1\}$;

$x_3 : \{x_3, y_1, y_3\}$;

$x_4 : \{x_4, y_1, y_2\}$;

$x_5 : \{x_5, y_2, y_3\}$.

В теореме 4.4 было введено понятие объединения совокупностей ребер циклов, никакие два из которых не имеют общих ре-

бер. Непустое объединение множеств ребер некоторого числа циклов, никакие два из которых не имеют общих ребер, называется *границей*. Объединение множеств ребер некоторого числа сечений, никакие два из которых не имеют общих ребер, называется *кограницей*. Далее мы увидим, что понятия границы и кограницы являются двойственными.

Лемма 4.2. Если подмножество $B \subseteq E$ ребер связного графа $G = (V, E)$ не содержит целиком ни одного коцикла, то дерево, покрывающее G , может содержать только ребра из множества $E - B$ (рис. 4.16).

Доказательство. Пусть $T = (V, E')$ — дерево, покрывающее граф G и имеющее наибольшее число ребер из $E - B$. Покажем, что $E' \subseteq E - B$.

Допустим, что существует ребро $x \in E' \cap B$. Рассмотрим главный коцикл, определяемый этим ребром. Так как, по предположению, не существует сечений, которые состояли бы только из ребер множества B , то найдется хорда $x' \in E - B$, которая входит в главный коцикл, определяемый ребром x . Другими словами, главный цикл, определяемый хордой x' , содержит ветвь x . Если с помощью описанного в теореме 4.14 элементарного преобразования дерева поменять местами x и x' , то у нового дерева число ребер из $E - B$ на 1 больше, чем у T . Полученное противоречие показывает, что $E' \subseteq E - B$.

Теорема 4.16. Если непустое подмножество B ребер связного графа $G = (V, E)$ имеет четное число общих ребер с любым циклом, то B является объединением множеств ребер некоторых коциклов, никакие два из которых не имеют общих ребер.

Доказательство. Если B не содержит в качестве подмножества ни одного такого коцикла, то, согласно лемме 4.2, существует дерево $T = (V, E)$, покрывающее G , такое, что $E' \subseteq E - B$, т. е. $E' \cap B = \emptyset$. Так как $B \neq \emptyset$, то найдется ребро $x \in B$. Поскольку ребро $x \notin E'$, оно является хордой. Главный цикл, определяемый хордой x , не имеет других ветвей, общих с B , кроме x , а поэтому x оказывается единственным ребром, общим для этого цикла и B . Однако это противоречит предположению.

Таким образом, существует коцикл $C_1 \subset B$. Если разность $B - C_1$ непуста, то условия теоремы вновь выполняются и, следовательно, существует коцикл $C_2 \subset B - C_1$. Продолжая этот процесс, можно получить совокупность коциклов, никакие два из которых не имеют общих ребер:

$$B = \bigcup_{i=1}^l C_i.$$

Следствие. Звезда $\delta(a)$ произвольной вершины a является кограницей.

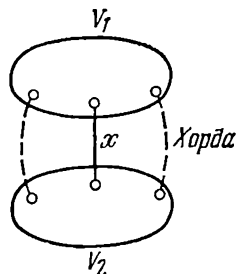


Рис. 4.16. Главный коцикл, определяемый ветвью x

Доказательство. В самом деле, у $\delta(a)$ и произвольного цикла либо нет общих ребер, либо есть ровно два общих ребра.

Теорема 4.17. Если некоторое непустое подмножество ребер $B \subseteq E$ связного графа $G = (V, E)$ имеет четное число общих ребер с произвольным коциклом графа, то B является объединением некоторого числа циклов, никакие два из которых не имеют общих ребер, т. е. является границей.

Доказательство. Рассмотрим произвольную вершину a . Из теоремы 4.16 следует, что если звезда $\delta(a)$ вершины a — непустое множество, то она представляет собой объединение совокупностей ребер коциклов C_1, \dots, C_i , никакие два из которых не имеют общих ребер. Отсюда и из того, что B имеет четное число общих ребер с каждым из коциклов C_1, \dots, C_i , следует, что B имеет четное число общих ребер также и со звездой $\delta(a)$. Это означает, что все вершины графа имеют четную степень. Отсюда и из теоремы 4.4 следует, что B — граница.

Из двух последних теорем и следствия из теоремы 4.15 вытекает следующая теорема.

Теорема 4.18. Граница и кограница имеют четное число общих ребер. Для того чтобы некоторое множество ребер было границей (кограницей), необходимо и достаточно, чтобы оно имело четное число общих ребер с любой кограницей (границей).

Предположим, что множества B_1 и B_2 — границы. Покажем, что $B_1 \oplus B_2$ также является границей. Имеем $B_1 \oplus B_2 = (B_1 - B_2) \cup (B_2 - B_1)$. Обозначим через n_1 , n' и n_2 общее число ребер некоторой кограницы соответственно с $B_1 - B_2$, $B_1 \cap B_2$ и $B_2 - B_1$. Согласно последней теореме, числа $n_1 + n'$ и $n_2 + n'$ четные. Следовательно, четной является также сумма $n_1 + n_2 = (n_1 + n') + (n_2 + n') - 2n'$, что доказывает сделанное выше утверждение. Аналогичный вывод можно сделать и для кограницы.

Следствие. Если B_1, B_2 — границы (кограницы), то симметрическая разность $B_1 \oplus B_2$ также является границей (кограницей).

4.5. Векторные пространства, связанные с графами

Рассмотрим связный граф $G = (V, E)$ с p вершинами и q ребрами. Перенумеруем ребра графа числами $1, \dots, q$ и сопоставим множеству ребер A вектор

$$a = (a_1, a_2, \dots, a_q), \quad (4.8)$$

положив

$$a_i = 1, \text{ если } i \in A, \text{ и } a_i = 0 \text{ в противном случае.} \quad (4.9)$$

Символы 0 и 1 — компоненты векторов в данном разделе — будем трактовать соответственно как единичный элемент 0 по сложению и единичный элемент 1 по умножению поля Z_2 , описанного в гл. 1. Сложение и умножение определяются в этом поле равенствами: $0+0=1$, $0+1=1+0=1$, $1+1=0$, $0 \cdot 0=0 \cdot 1=1 \cdot 0=$

$=0, 1 \cdot 1 = 1$. Рассмотрим вектор \mathbf{a} из q -мерного векторного пространства $V_q(Z_2)$. Пусть

$$\mathbf{b} = (b_1, \dots, b_q) \quad (4.10)$$

— вектор из того же векторного пространства. Определим сложение векторов с помощью соотношения

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, \dots, a_q + b_q). \quad (4.11)$$

Кроме того, определим произведение вектора на элемент c , положив

$$c\mathbf{a} = (ca_1, \dots, ca_q). \quad (4.12)$$

Сложение соответствует симметрической разности множеств. Это следствие того, что в (4.11) равенство $a_i + b_i = 1$ выполняется тогда и только тогда, когда $a_i = 1, b_i = 0$ или $a_i = 0, b_i = 1$. Другими словами, сумме

$$\mathbf{a} + \mathbf{b} \quad (4.13)$$

соответствует множество

$$A \oplus B = (A - B) \cup (B - A). \quad (4.14)$$

Внутреннее произведение векторов \mathbf{a} и \mathbf{b} определяется соотношением

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^q a_i b_i. \quad (4.15)$$

Равенство $(\mathbf{a}, \mathbf{b}) = 0$ означает, что четное число произведений $a_i b_i$ равно 1. Для таких произведений $a_i = 1, b_i = 1$, и, следовательно, соответствующие векторам \mathbf{a}, \mathbf{b} множества ребер A, B имеют четное число общих ребер.

Подмножество \mathcal{A} векторного пространства $V_q(Z_2)$ называется *подпространством* (линейным подпространством), если для произвольных векторов $\mathbf{a}, \mathbf{b} \in \mathcal{A}$ их сумма $\mathbf{a} + \mathbf{b}$ также лежит в \mathcal{A} . Множество \mathcal{L} векторов, соответствующих любой границе, и множество \mathcal{C} векторов, соответствующих когранице, как следует из теоремы 4.18, являются линейными подпространствами. Два векторных пространства называются *ортогональными*, если для произвольных элементов $\mathbf{a} \in \mathcal{A}$ и $\mathbf{b} \in \mathcal{B}$ их внутреннее произведение (\mathbf{a}, \mathbf{b}) равно 0. Множество \mathcal{A}' всех векторов, ортогональных векторам подпространства \mathcal{A} , является подпространством и называется пространством, двойственным к \mathcal{A} . На языке векторных пространств теорему 4.18 можно сформулировать следующим образом.

Теорема 4.19. Пусть \mathcal{L} — подпространство, представляющее собой совокупность векторов, соответствующих некоторой границе. Тогда двойственное подпространство \mathcal{C} состоит из векторов, соответствующих некоторой когранице. Верно также и обратное утверждение.

Векторы $a^{(1)}, \dots, a^{(k)}$ называются линейно независимыми, если равенство

$$c_1 a^{(1)} + \dots + c_k a^{(k)} = 0; \quad c_i \in Z_2 (i = 1, \dots, k) \quad (4.16)$$

имеет место в том и только в том случае, если $c_1 = \dots = c_k = 0$. Если векторы не являются линейно независимыми, то они называются линейно зависимыми. Размерность векторного подпространства называется максимальное число линейно независимых векторов, принадлежащих этому подпространству. Если пространство имеет размерность k , то совокупность любых k его линейно независимых векторов называется *базисом*. Линейной комбинацией называется выражение, подобное левой части равенства (4.16) и представляющее собой сумму некоторого числа векторов, умноженных на скаляры. Произвольный вектор пространства представляется единственным образом в виде линейной комбинации векторов базиса. Единственность представления в виде линейной комбинации означает, что единственным образом определяются значения констант c_1, c_2, \dots линейной комбинации. Если векторные подпространства \mathcal{A}, \mathcal{B} пространства $V_q(Z_2)$ двойственны, то, как хорошо известно, сумма их размерностей равна q^* .

Предположим, что задано дерево $T(V, E')$, покрывающее граф G . Рассмотрим матрицу M , составленную из векторов, соответствующих $\mu = q - p + 1$ главным циклам. Перенумеруем хорды и ветви графа соответственно числами $1, \dots, \mu$ и $\mu + 1, \dots, q$. Если векторное представление главного цикла, определяемого хордой i , взяты в качестве i -й строки матрицы, то получится матрица M размера $\mu \times q$. Левая ее подматрица размера $\mu \times \mu$ имеет элементы 1 на главной диагонали и 0 на остальных позициях. Это является следствием того, что каждый главный цикл содержит одну и только одну хорду. Матрица M называется матрицей главных циклов, она имеет вид

$$M = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} = [I_\mu, F], \quad (4.17), (4.18)$$

где I_μ — единичная матрица размера $\mu \times \mu$; F — матрица размера $\mu \times \varphi$.

Так как μ строк матрицы M линейно независимы, то размерность пространства \mathcal{L} не меньше чем μ . Другими словами, если $\dim \mathcal{L}$ — это размерность \mathcal{L} , то

$$\dim \mathcal{L} \geq \mu. \quad (4.19)$$

Матрица главных коциклов Q определяется следующим обра-

* Это следует из того, что введенное ранее внутреннее произведение невырожденно, т. е. если $(a, b) = 0$ для всех $b \in V_q(Z_2)$, то $a = 0$. (Прим. ред.)

зом. В качестве i -й строки Q берется главный коцикл, соответствующий хорде $\mu + i$. При этом получается матрица размера $\varphi \times q$:

$$Q = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} = [K, I_\varphi], \quad (4.20), (4.21)$$

где I_φ — это единичная матрица размера $\varphi \times \varphi$, элементы главной диагонали которой равны 1, а остальные равны 0; K — матрицы размера $\varphi \times \mu$. Матрица Q имеет такой вид вследствие того, что каждый главный коцикл содержит одну и только одну ветвь. Так как строки Q , очевидно, линейно независимы, то

$$\dim C \geq \varphi. \quad (4.22)$$

Из неравенств (4.19), (4.22) получаем, что

$$\dim \mathcal{L} + \dim C \geq \mu + \varphi = q. \quad (4.23)$$

Ясно, что левая часть последнего равенства не может превосходить q , а следовательно,

$$\dim \mathcal{L} = \mu, \quad \dim C = \varphi. \quad (4.24)$$

Таким образом, доказана следующая теорема.

Теорема 4.20. Размерность \mathcal{L} равна цикломатическому числу μ ; μ векторов-строк матрицы главных циклов M образуют базис \mathcal{L} . Размерность C равна корангу φ ; φ векторов-строк матрицы главных сечений Q являются базисом C .

Вновь воспользовавшись теоремой 4.18, нетрудно убедиться, что

$$MQ^T = 0, \quad (4.25)$$

где символ T обозначает транспонирование, а правая часть представляет собой нулевую матрицу размера $\mu \times \mu$. Это равенство можно также переписать в виде

$$[I_\mu, F] \begin{bmatrix} K^T \\ I_\varphi \end{bmatrix} = K^T + F = 0.$$

Отсюда следует, что

$$K = -F^T = F^T. \quad (4.26)$$

Таким образом, доказана следующая теорема.

Теорема 4.21. Пусть G — связный граф с корангом φ и цикломатическим числом μ , а T — покрывающее его дерево, хордам и ветвям которого сопоставлены соответственно числа $1, \dots, \mu$ и $\mu + 1, \dots, q$. Матрица главных циклов M , образованная векторами, соответствующими μ главным циклам, и матрица главных коциклов Q , образованная векторами, соответствующими φ главным коциклам, имеют следующий вид:

$$M = [I_\mu, F]; \quad Q = [F^T, I_\varphi].$$

Упражнение 4.9. Постройте матрицу главных циклов и матрицу главных коциклов, соответствующие графу и дереву, изображенным на рис. 4.12.

Решение. Если присвоить ребрам номера в соответствии с их положением в последовательности $y_1, y_2, y_3, x_1, x_2, x_3, x_4, x_5$, то, используя решения упражнений 4.6 и 4.8, получим

$$M = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & 1 & \cdot & 1 & \cdot & 1 & \cdot & 1 \end{bmatrix};$$

$$Q = \begin{bmatrix} \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$

(в этих матрицах проставлены только символы 1, а нули заменены точками).

4.6. Двудольные графы

В связном графе $G=(V, E)$ введем понятие расстояния $d(a, b)$ между вершинами a и b . Для этого положим $d(a, a)=0$, а для двух несовпадающих вершин a и b определим $d(a, b)$ как длину самой короткой $(a-b)$ -цепи. Можно проверить, что введенное таким образом расстояние удовлетворяет всем аксиомам расстояния, приведенным в § 2.2. Действительно, очевидно, что а) $d(a, b) \geq 0$, б) равенства $d(a, b)=0$ и $a=b$ эквивалентны. Далее, если существуют $(a-b)$ -цепь длины $d(a, b)$ и $(b-c)$ -цепь длины $d(b, c)$, то между вершинами a и c существует путь длины $d(a, b)+d(b, c)$. Но тогда, как следует из теоремы 4.1, существует также $(a-c)$ -цепь длины не более чем $d(a, b)+d(b, c)$. Следовательно, $d(a, b)+d(b, c) \geq d(a, c)$. В общем случае для вершины a и множеств вершин $V' \subset V$ положим

$$d(a, V') = \min_{b \in V'} d(a, b). \quad (4.27)$$

Кроме того, для множеств вершин V', V'' положим

$$d(V', V'') = \min_{a \in V', b \in V''} d(a, b). \quad (4.28)$$

Упражнение 4.10. Покажите, что если два множества вершин V', V'' связного графа $G=(V, E)$ таковы, что $d(V', V'')=r>0$, то цепь длины r , соединяющая вершину из V' и вершину из V'' , не содержит никаких других вершин из множеств V' и V'' , кроме своих концов.

Решение. Пусть $V' \ni a_0, x_1, a_1, \dots, x_r, a_r \in V''$ — цепь длины r ; $r \geq 2$. Допустим, что $a_i \in V'$ при некотором $0 < i < r$. В этом случае последовательность $V' \ni a_i, x_{i+1}, \dots, x_r, a_r \in V''$ также является цепью, соединяющей вершину a_i из V' и вершину a_r из V'' . Длина этой цепи $r-i < r$. Следовательно, $d(a_i, a_r) < r$, что противоречит равенству $d(V', V'')=r$.

Упражнение 4.11. Покажите, что если в связном графе $G=(V, E)$ цепь $a_0, x_1, a_1, \dots, x_l, a_l$ является кратчайшей (a_0-a_l) -

цепью, то для произвольной ее вершины $a_i (1 \leq i \leq l-1)$ путь $a_i, x_{i+1}, \dots, x_l, a_l$ также является кратчайшей $(a_i - a_l)$ -цепью.

Решение. Предположим, что между вершинами a_i и a_l существует более короткая цепь. Пусть это будет цепь $a_i = b_0, y_1, \dots, y_k, b_k (k < l-i)$. Путь $a_0, x_1, \dots, a_i, y_1, \dots, y_k, a_l$ соединяет вершины a_0, a_l и имеет длину $i+k < i+(l-i) = l$. Но тогда, как следует из теоремы 4.1, существует цепь $(a_0 - a_l)$, длина которой меньше l . Получили противоречие.

Пусть V — множество вершин графа $G = (V, E)$. Предположим, что существует разбиение V на два непересекающихся подмножества V_1 и V_2 , при котором каждая вершина графа принадлежит ровно одному из множеств V_1 или V_2 , ни одно из множеств V_1, V_2 не является пустым и каждое из ребер множества E соединяет некоторую вершину из множества V_1 с некоторой вершиной из V_2 . Тогда граф G называется *двудольным* (рис. 4.17).

Теорема 4.22. Граф, имеющий более одной вершины, является двудольным тогда и только тогда, когда все его простые циклы содержат четное число ребер.

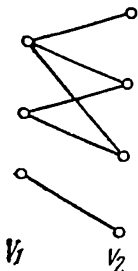


Рис. 4.17. Двудольный граф

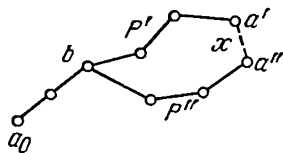


Рис. 4.18. Общая вершина b

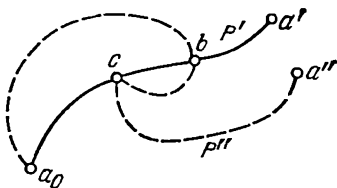


Рис. 4.19. Общая вершина c

Доказательство. Так как необходимость сформулированного условия очевидна, то докажем только его достаточность. Без ограничения общности можно предположить, что граф является связным. Если же граф несвязен, то следует рассмотреть каждую его связную компоненту.

Возьмем произвольную вершину a_0 и определим множества $V_1 = \{a \in V | d(a_0, a) \text{ четно} \}$ и $V_2 = \{a \in V | d(a_0, a) \text{ нечетно} \}$. Так как $a_0 \in V_1$, то V_1 непусто. Поскольку граф G является связным и имеет не менее двух вершин, то непусто также и множество V_2 .

Допустим, что некоторое ребро $x \in E$ соединяет две вершины $a', a'' \in V_1$, и придем к противоречию. Если $a', a'' \in V_2$, доказательство проводится аналогично. Пусть P' и P'' — кратчайшие цепи, соединяющие соответственно a_0 с a' и a_0 с a'' (рис. 4.18). Так как обе цепи имеют четную длину, то рассматриваемое ребро x не принадлежит P', P'' . Предположим, что цепь P' имеет несколько общих вершин с P'' и b — та из них, которая находится ближе всего к a' . Согласно упражнению 4.11, часть $P'[a_0, b]$

цепи P' от вершины a_0 до b является кратчайшей $(a_0—b)$ -цепью, аналогично $P''[a_0, b]$ — кратчайшая $(a_0—b)$ -цепь. Следовательно, длины $P'[a_0, b]$ и $P''[a_0, b]$ равны. Далее заметим, что кроме вершины b цепь $P''[b, a'']$ не имеет других общих вершин с P' . Действительно, допустим, что $P''[b, a'']$ имеет кроме b другую вершину c , также лежащую в P' (рис. 4.19). Тогда, по определению b , вершина c должна принадлежать $P'[a_0, b]$, так что длина цепи $P'[a_0, c]$ меньше, чем длина $P[a_0, b] = P''[a_0, b]$. С другой стороны, длина $P''[a_0, c]$ больше, чем длина $P'[a_0, b]$. Но, согласно упражнению 4.11, обе цепи $P'[a_0, c]$ и $P''[a_0, c]$ являются кратчайшими $(a_0—c)$ -цепями (см. рис. 4.19). Получили противоречие.

Таким образом, путь $P'[b, a']$, x , $P''[b, a'']$ оказывается циклом, а его длина $d(a_0, a') + d(a_0, a'') + 1 - 2d(a_0, b)$ — нечетным числом. Это противоречит предположению.

Раскрашиванием графа называют процесс сопоставления его вершинам цветов, при котором любые две соседние вершины оказываются разноцветными. Минимальное число цветов, необходимых для такой раскраски графа, называют *хроматическим*. Заметим, что граф имеет хроматическое число 2, т. е. может быть раскрашен с помощью всего двух цветов, в том и только в том случае, если он двудольный. Так как деревья и леса не имеют циклов, то они могут быть раскрашены с помощью не более чем двух цветов. Графом, допускающим раскраску одним цветом, может быть лишь совокупность изолированных вершин, которая вообще не имеет ни одного ребра.

Задачи

4.1. Предположим, что только две вершины a и b графа имеют нечетные степени. Покажите, что в графе существует $(a—b)$ -цепь.

4.2. Докажите, что две цепи максимальной длины P_1 и P_2 в связном графе всегда имеют общие вершины.

4.3. Покажите, что сплошные линии на рис. 4.20 являются хордами, а пунктирные линии — ветвями дерева, покрывающего граф.

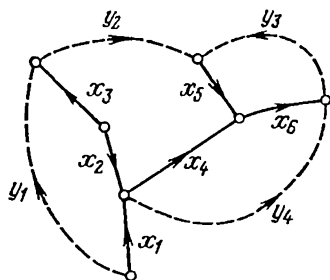


Рис. 4.20. К задаче 4.3

4.4. Предположим, что имеется два дерева, $T=(V, F)$ и $T'=(V, F')$, покрывающих связный граф $G=(V, E)$ с n вершинами, и что число общих ребер у множеств F и F' равно $(n-2)$. Докажите, что в этом случае дерево T' может быть получено из T элементарным преобразованием.

4.5. Вершина a графа $G=(V, E)$ называется разделяющей, если степень связности графа $G'=(V-\{a\}, E-\{\delta(a)\})$, полученного из G исклю-

чением вершины a и инцидентных ей ребер, больше степени связности G . Покажите, что вершина a — разделяющая тогда и только тогда, когда в графе существуют две различные от a вершины b и c такие, что все $(b-c)$ -цепи проходят через a .

Глава 5

Ориентированные графы

5.1. Ориентированные пути

Граф, каждое ребро которого имеет ориентацию, называется *ориентированным графом* и обозначается через $\vec{G}(V, E)$. Ориентация ребра обычно обозначается стрелкой, например, как это показано на рис. 5.1. Ребро \vec{x} называется ориентированным в направлении от a к b или выходящим из a и входящим в b . При этом вершины a и b называются соответственно начальной (началом) и конечной точкой (концом) ребра \vec{x} . Формально ориентированный граф $\vec{G} = (V, E)$ определяется как множество вершин V и множество E упорядоченных пар вершин $[a, b]$. В отличие от неориентированного графа, в котором каждая вершина a имеет одну звезду $\delta(a)$, в ориентированном графе у каждой вершины имеется две звезды: $\delta^+(a)$ — множество (ориентированных) ребер, выходящих из a , и $\delta^-(a)$ — множество ребер, входящих в a . Числа $|\delta^+(a)|$ и $|\delta^-(a)|$ называются соответственно положительной и отрицательной степенями вершины a .

Упражнение 5.1. Покажите справедливость следующих равенств:

$$|\vec{E}| = \sum_{a \in V} |\delta^+(a)| = \sum_{a \in V} |\delta^-(a)|. \quad (5.1)$$

Решение. Приведенные равенства можно получить, если сопоставить каждому ребру его начало или конец.

Понятие (ориентированного) подграфа ориентированного графа вводится точно также как и ранее (разд. 4.1). Далее, если пренебречь ориентацией ребер в ориентированном графе, то получится (неориентированный) граф, который называется (неориентированным) графом, соответствующим исходному ориентированному графу. *Направленным путем* длины l называется последовательность

$$a_0, \vec{x}_1, a_1, \vec{x}_2, \dots, a_{l-1}, \vec{x}_l, a_l, \quad (5.2)$$

где \vec{x}_i — ориентированное ребро, выходящее из вершины a_{i-1} и

входящее в вершину $a_i (i=1, 2, \dots, l)$, т. е. $\vec{x}_i = [a_{i-1}, a_i]$. Незамкнутый путь, замкнутый путь, простой путь определяются так же, как и ранее. Открытый путь, все вершины a_0, a_1, \dots, a_l которого различны, называется ориентированной $(a_0 \rightarrow a_l)$ -цепью. Замкнутый простой путь, все вершины a_0, a_1, \dots, a_l которого различны, называется *ориентированным циклом* (рис. 5.2). В теории гра-

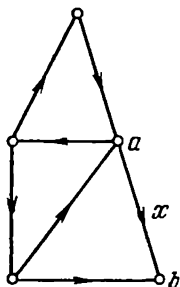


Рис. 5.1. Ориентированный граф

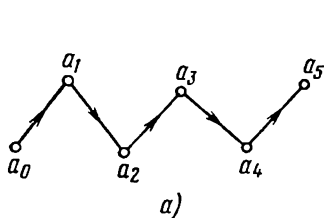


Рис. 5.2. Ориентированная цепь и ориентированный цикл:
а — $(a_0 \rightarrow a_5)$ -цепь; б — цикл

фов иногда используется запись $a \Rightarrow b$ для указания того, что вершины a и b ориентированного графа либо совпадают, либо соединены $(a \rightarrow b)$ -цепью. При этом также говорят о том, что b можно достичь из a .

Теоремам 4.1 и 4.2, сформулированным и доказанным для неориентированных графов, соответствуют две теоремы для ориентированных графов.

Теорема 5.1. Если две различные вершины a и b соединены ориентированным путем, выходящим из a и входящим в b , то из ориентированных ребер этого пути можно построить $(a \rightarrow b)$ -цепь.

Теорема 5.2. Если в графе существует замкнутый ориентированный путь, то из ребер этого пути можно построить ориентированный цикл.

Обе эти теоремы доказываются почти так же, как и ранее. Однако для доказательства теоремы 5.2, в отличие от теоремы 4.2, условие простоты не требуется. Это связано с тем, что в ориентированных графах нет замкнутых путей, подобных приведенному ранее контрпримеру a_0, x, a_1, x, a_0 . В пути $a_0, \vec{x}, a_1, \vec{y}, a_0$ два ребра $\vec{x} = [a_0, a_1]$ и $\vec{y} = [a_1, a_0]$, очевидно, являются различными.

Теореме 4.3 соответствует следующая теорема для ориентированных графов.

Теорема 5.3. Пусть $\vec{G} = (V, \vec{E})$ — ориентированный граф, у которого положительная и отрицательная степени каждой вершины совпадают. Тогда для произвольного ориентированного ребра

\vec{x}_1 в \vec{G} существует замкнутый простой ориентированный путь, содержащий \vec{x}_1 .

Доказательство этой теоремы отличается от доказательства теоремы 4.3 лишь тем, что в данном случае всегда следует двигаться согласно ориентации ребер. Если $\vec{x}_1 = [a_0, a_1]$, то, отправляясь из a_1 и следуя вдоль \vec{x}_1 , мы приходим в a_1 . Так как положительная и отрицательная степени вершины a совпадают, то существует ориентированное ребро $\vec{x}_2 = [a_1, a_2]$, выходящее из a_1 и отличное от \vec{x}_1 . Следуя из a_1 вдоль этого ребра, мы придем в a_2 . Повторяя эту процедуру, можно построить замкнутый простой ориентированный путь, возвращающийся в a_0 и содержащий \vec{x}_1 .

Доказательство следующей теоремы 5.4 можно получить, если все ребра, фигурирующие в доказательстве теоремы 4.4, считать ориентированными.

Теорема 5.4. Предположим, что положительные и отрицательные степени всех вершин ориентированного графа $\vec{G} = (V, \vec{E})$ совпадают. Если множество ориентированных ребер \vec{E} непусто, то оно представляет собой объединение множеств ребер некоторого числа ориентированных циклов, никакие два из которых не имеют общих ребер.

Следствие. Предположим, что положительные и отрицательные степени всех вершин ориентированного графа $\vec{G} = (V, \vec{E})$ совпадают. Тогда для произвольного ребра \vec{x} в \vec{G} существует ориентированный цикл, содержащий \vec{x} .

Рассмотрим ориентированный граф $\vec{G} = (V, \vec{E})$. Если соответствующий ему неориентированный граф $G = (V, E)$ является связным, то ориентированный граф \vec{G} называется слабо связным. Другими словами, слабая связность означает, что, пренебрегая ориентацией ребер, можно из любой вершины графа по ребрам дойти до любой другой вершины.

Разомкнутой (замкнутой) эйлеровой цепью называется разомкнутый (замкнутый) простой ориентированный путь, содержащий все ориентированные ребра графа. Теореме 4.8 соответствует следующая теорема.

Теорема 5.5. Ориентированный граф имеет эйлерову цепь тогда и только тогда, когда: а) он является слабосвязным, б) положительная и отрицательная степени каждой его вершины равны между собой.

Доказательство этой теоремы аналогично доказательству теоремы 4.8, однако в данном случае нужно следить за ориентацией ребер. Следствие из теоремы 4.8 соответствует также следствие из теоремы 5.5.

Следствие. Ориентированный граф имеет разомкнутую эйлерову цепь тогда и только тогда, когда: а) он является слабо связным, б) в нем существуют одна вершина, положительная степень которой на 1 больше отрицательной степени, и одна вершина, отрицательная степень которой на 1 больше положительной, и в) положительная и отрицательные степени любой другой вершины равны между собой.

Доказательство отличается лишь тем, что дополнительно вводятся ориентированное ребро $[b, a]$, где a — вершина, для которой (положительная степень) — (отрицательная степень) = 1, и b — вершина, для которой (отрицательная степень) — (положительная степень) = 1.

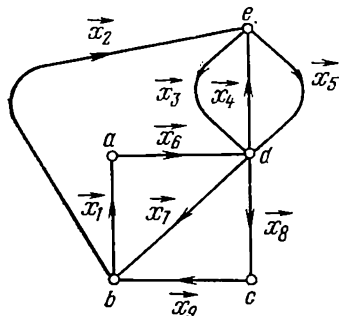


Рис. 5.3. Ориентированный граф из упражнения 5.2

Эйлерова цепь в ориентированном графе, удовлетворяющем условиям теоремы 5.5, может быть построена точно так же, как и при решении упражнения 4.4; при этом на ориентацию ребер не следует обращать внимания.

Упражнение 5.2. Найдите замкнутую эйлерову цепь в ориентированном графе, изображенном на рис. 5.3.

Решение 5.2. Начиная с ребра $x_1 = [b, a]$, получаем $b, \vec{x}_1, a, \vec{x}_6, d, \vec{x}_8, c, \vec{x}_9, b, \vec{x}_2, e, \vec{x}_5, d, \vec{x}_4, e, \vec{x}_3, d, \vec{x}_7, b$.

5.2. Сильная связность

Ориентированный граф $\vec{G} = (V, \vec{E})$ называется сильно связным, если для произвольных двух его вершин, a и b , существуют $(a \rightarrow b)$ - и $(b \rightarrow a)$ -цепи, т. е. $a \Rightarrow b$ и $b \Rightarrow a$.

В множество вершин ориентированного графа можно ввести отношение \equiv , положив $a \equiv b$ для тех пар вершин графа a и b , для которых одновременно выполняются оба соотношения $a \Rightarrow b$ и $b \Rightarrow a$. Отношение \equiv , как легко видеть, является отношением эквивалентности. Действительно, оно рефлексивно, так как $a \Rightarrow a$. Так как из $a \equiv b$ следует, что $b \equiv a$, оно, очевидно, также симметрично. Предположим, что $a \equiv b$ и $b \equiv c$. Так как при этом вершина b достижима из a и c из b , то c достижима из a , т. е. $a \Rightarrow c$. Точно так же из достижимости c из b , а b из a следует достижимость a из c , т. е. $c \Rightarrow a$. Следовательно, $a \equiv c$ и отношение \equiv транзитивно.

Обозначим через S_1, \dots, S_k классы эквивалентности, получающиеся при разбиении множества вершин V сильно связного ориентированного графа $\vec{G} = (V, \vec{E})$ с помощью отношения эквивалентности, введенного ранее. Графы $(S_1, \vec{E}(S_1)), \dots, (S_k, \vec{E}(S_k))$, получающиеся из множеств S_1, \dots, S_k , называются сильно связными.

ми компонентами графа \vec{G} . Множество $\vec{E}(S_i)$ представляет собой множество ориентированных ребер из \vec{E} , начинающихся и заканчивающихся в вершинах множества S_i . Если граф \vec{G} является сильносвязным, то, очевидно, его единственной сильносвязной компонентой будет он сам. Если граф имеет две или более сильносвязные компоненты, то ребра, которые не принадлежат ни одной из сильносвязных компонент, связывают вершины различных компонент. Предположим, что $i \neq j$, $a \in S_i$, $b \in S_j$, $x = [a, b] \in E$, и покажем, что в графе не существует ребер вида $[b', a']$, где $a' \in S_i$ и $b' \in S_j$. Допустим, что такое ребро существует. Тогда, так как S_i и S_j — классы эквивалентности, то $a' \Rightarrow a$, $a \Rightarrow b$, $b \Rightarrow b'$, $b' \Rightarrow a'$ и, следовательно, выполняются оба соотношения $a \Rightarrow b$ и $b \Rightarrow a$, т. е. $a \equiv b$. Это противоречит тому, что S_i и S_j — различные классы эквивалентности.

Упражнение 5.3. Найдите сильносвязные компоненты ориентированного графа, изображенного на рис. 5.4.

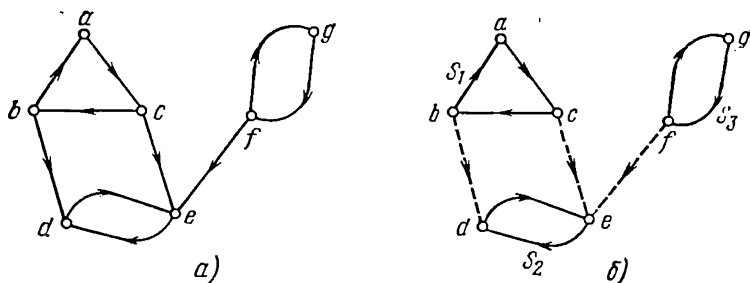


Рис. 5.4. Графы из упражнения 5.4:

a — заданный ориентированный граф; b — связанные компоненты

Решение. Так как $S_1 = \{a, b, c\}$, $S_2 = \{d, e\}$, $S_3 = \{f, g\}$, то сильносвязными компонентами будут три ориентированных подграфа, которые получаются после удаления из рассматриваемого графа ориентированных ребер, показанных на рис. 5.4, b пунктирными линиями.

Рассмотрим сильносвязный граф $\vec{G} = (V, \vec{E})$ с сильносвязными компонентами $(S_1, \vec{E}(S_1))$, ..., $(S_k, \vec{E}(S_k))$. Сопоставим ему ориентированный граф следующим образом. Множествам S_1 , ..., S_k сопоставим соответственно вершины a_1 , ..., a_k . Если существует ребро, выходящее из вершины множества S_i и входящее в вершину из S_j , то сопоставим ему в новом графе ориентированное ребро $[a_i, a_j]$, выходящее из a_i и входящее в a_j . Полученный таким образом граф назовем *сжатием* графа G . Сжатие графа, рассмотренного в упражнении 5.3 (см. рис. 5.4, a), показано на рис. 5.5.

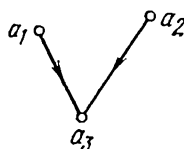


Рис. 5.5. Сжатие графа из упражнения 5.4

Упражнение 5.4. Покажите, что сжатие слабосвязного графа является слабосвязным и не содержит ориентированных циклов.

Решение. Так как слабая связность сжатия почти очевидна, то докажем только отсутствие ориентированных циклов. Пусть $(S_1, \vec{E}(S_1)), \dots, (S_k, \vec{E}(S_k))$ — сильносвязные компоненты заданного слабосвязного ориентированного графа $\vec{G} = (V, \vec{E})$. Допустим, что сжатие этого графа имеет ориентированный цикл. Для простоты обозначений предположим, что таковым является $a_1, [a_1, a_2], a_2, \dots, [a_{l-1}, a_l], a_l = a_1$. Тогда в исходном графе \vec{G} существуют вершины $b_1, b'_1 \in S_1, b_2, b'_2 \in S_2, \dots, b_{l-1}, b'_{l-1} \in S_{l-1}$ такие, что ориентированные ребра $[b'_1, b_2], [b'_2, b_3], \dots, [b'_{l-1}, b_1]$ принадлежат \vec{E} . Так как при этом $b_1 \Rightarrow b_2, b_2 \Rightarrow b'_2, b'_2 \Rightarrow b_3, \dots, b'_{l-1} \Rightarrow b_1, b_1 \Rightarrow b'_1$, то $b_1 \Rightarrow b_2, b_2 \Rightarrow b_1$, т. е. $b_1 \equiv b_2$. С другой стороны, вершины b_1 и b_2 принадлежат различным классам эквивалентности S_1 и S_2 . Получили противоречие.

5.3. Ориентированные деревья

Вершина a ориентированного графа \vec{G} называется его *корнем*, если любая другая его вершина достижима из a . Ориентированный граф $\vec{G} = (V, \vec{E})$ называется *ориентированным деревом*, если он имеет корень и соответствующий ему неориентированный граф G является обычным деревом.

Теорема 5.6. Ориентированное дерево имеет только один корень.

Доказательство. Если предположить, что ориентированное дерево $\vec{G} = (V, \vec{E})$ имеет два различных корня a_1, a_2 , придем к противоречию. Так как существуют $(a_1 \rightarrow a_2)$ - и $(a_2 \rightarrow a_1)$ -цепи, то, пренебрегая ориентацией ребер, можно доказать существование в неориентированном графе $G = (V, E)$ двух различных цепей, соединяющих вершины a и b . Однако это противоречит теореме 4.9.

Лемма 5.1. Если вершина a является корнем ориентированного дерева $\vec{G} = (V, \vec{E})$, то отрицательный порядок вершины a равен 0, а отрицательный порядок любой отличной от a вершины равен 1.

Доказательство. Если отрицательный порядок вершины является положительным числом, то найдется вершина a' такая, что $\vec{x} = [a', a] \in \vec{E}$, т. е. $a' \Rightarrow a$. Так как a — корень, то для произвольной вершины $b \in V$ $a \Rightarrow b$. Следовательно, для произвольной вершины $b \in V$ $a' \Rightarrow b$. Это означает, что корень a' отличен от a , что противоречит теореме 5.6.

Так как для произвольной отличной от корня a вершины b имеется $(a \rightarrow b)$ -цепь, то отрицательный порядок b является положительным числом. Если он равен или больше 2, то найдутся две различные вершины a_1, a_2 , такие, что $[a_1, b] \in \vec{E}, [a, b] \in \vec{E}$. Так как, кроме того, $a \Rightarrow a_1$ и $a \Rightarrow a_2$, то существуют также цепи

\vec{P}_1 и \vec{P}_2 , соединяющие соответственно a с a_1 и a с a_2 (рис. 5.6). Далее, поскольку по условию леммы соответствующий графу \vec{G} неориентированный граф G является деревом, то \vec{P}_1 не проходит через вершину b . По этой же причине через b не проходит и \vec{P}_2 . Следовательно, пути $((a \rightarrow a_1)$ -цепь \vec{P}_1 , $[a_1, b]$, b) и $((a \rightarrow a_2)$ -цепь \vec{P}_2 , $[a_2, b]$, b) являются $(a \rightarrow b)$ -цепями, причем различными. Однако это противоречит тому, что неориентированный граф G является деревом (теорема 4.9).

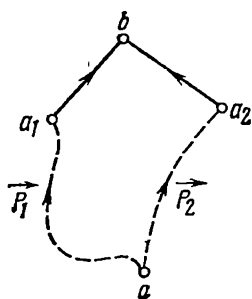


Рис. 5.6

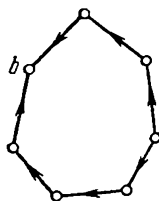


Рис. 5.7

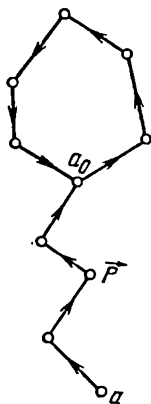


Рис. 5.8.

Рис. 5.6. $(a \rightarrow a_1)$ -цепь \vec{P}_1 и $(a \rightarrow a_1)$ -цепь \vec{P}_2

Рис. 5.7. Существование вершины b с отрицательной степенью 2 и выше

Рис. 5.8. Отрицательная степень a_0 не меньше двух

Теорема 5.7. Сильносвязный ориентированный граф является деревом тогда и только тогда, когда он имеет ровно одну вершину, отрицательная степень которой равна 0, а все остальные вершины имеют отрицательную степень 1.

Доказательство. Так как необходимость условий теоремы уже доказана (теорема 5.6), то докажем только их достаточность. Пусть a — вершина с отрицательной степенью 0 и b — любая другая отличная от нее вершина графа. Так как соответствующий ориентированному графу $\vec{G} = (V, \vec{E})$ неориентированный граф $G = (V, E)$ является связным, то в нем существует следующая $(a \rightarrow b)$ -цепь: $a = a_0, x, a_1, \dots, a_{l-1}, x_0, a_l = b$.

Поскольку отрицательная степень вершины a равна 0, то ребро x_1 в \vec{G} имеет ориентацию $[a_0, a_1]$. При этом никакое другое ребро, кроме x_1 , не может входить в a_1 , а следовательно, x_2 также имеет ориентацию $[a_1, a_2]$. Повторяя эти рассуждения, можно

показать, что $x_i = [a_{i-1}, a_i] \in \vec{E} (i=1, \dots, l)$, т. е. в \vec{G} существует $(a \rightarrow b)$ -цепь. Таким образом, вершина a — корень \vec{G} .

Далее, покажем, что неориентированный граф G является деревом. Если граф G — не дерево, то он имеет цикл $a_0, x_1, a_1, \dots, a_{l-1}, x_l, a_l = a_0$. Так как отрицательная степень каждой вершины не превосходит 1, то этому циклу в \vec{G} должен соответствовать ориентированный цикл. В противном случае, как показано на

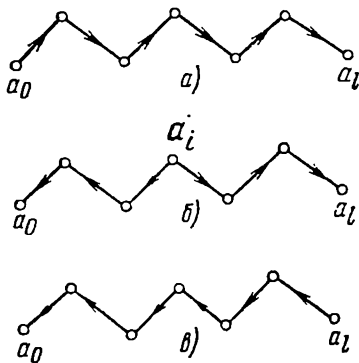


Рис. 5.9. Возможные ориентации

рис. 5.7, нашлась бы вершина с отрицательной степенью 2. Следовательно, каждая вершина этого цикла имеет отрицательную степень 1 и корень a не может принадлежать этому циклу.

Рассмотрим кратчайшую из ориентированных цепей $(a \rightarrow a_0)$, $(a \rightarrow a_1)$, ..., $(a \rightarrow a_l)$. Для простоты предположим, что такой цепью является $(a \rightarrow a_0)$ -цепь \vec{P} . Очевидно, что эта цепь не проходит через вершины a_1, \dots, a_{l-1} . Но тогда, как видно из рис. 5.8, отрицательная степень вершины a не меньше 2. Это противоречит предположениям теоремы.

Упражнение 5.5. Пусть $a_0, x_1, a_1, \dots, a_{l-1}, x_l, a_l$ — цепь в неориентированном графе $G = (V, E)$, соответствующем ориентированному графу $\vec{G} = (V, \vec{E})$. Покажите, что для некоторой вершины a_i этой цепи $a_i \rightarrow a_j$ ($j=0, 1, \dots$).

Решение. Так как отрицательная степень каждой из вершины не превосходит 1, то в ориентированном дереве \vec{G} ребра этой цепи могут иметь только три типа ориентации, показанные на рис. 5.9. Искомые вершинами a_i в этих трех случаях $(a \rightarrow b)$ являются соответственно вершины a_0, a_i и a_l .

Пусть a и b — это соответственно корень и произвольная вершина ориентированного дерева $\vec{G} = (V, \vec{E})$. Пусть S — множество вершин, достижимых из вершины b . Рассмотрим сечение $(S, \vec{E}(S))$. Нетрудно проверить, что этот граф является ориентированным деревом с корнем b . Действительно, b' — произвольная отличная от b вершина из S и $b = b_0, \vec{y}_1, b_1, \dots, \vec{y}_{l-1}, b_l = b'$ — $(b \rightarrow b')$ -цепь в \vec{G} . Так как вершины b_0, b_1, \dots, b_l этой цепи принадлежат S , то все ориентированные ребра $\vec{y}_1, \dots, \vec{y}_l$ принадлежат $\vec{E}(S)$. Следовательно, приведенная $(b \rightarrow b')$ -цепь является также $(b \rightarrow b')$ -цепью в сечении $(S, \vec{E}(S))$. Это означает, что вершина b — корень сечения. Если пренебречь ориентацией ребер, то из

отсутствия циклов в исходном графе следует отсутствие циклов и сечений. Это означает, что рассматриваемое сечение — ориентированное дерево с корнем b . Это дерево называется *ориентированным поддеревом* графа G .

Для ориентированного дерева, изображенного на рис. 5.10, а, на рис. 5.10, б и в приведены поддеревья, корнями которых являются соответственно вершины a , b и i .

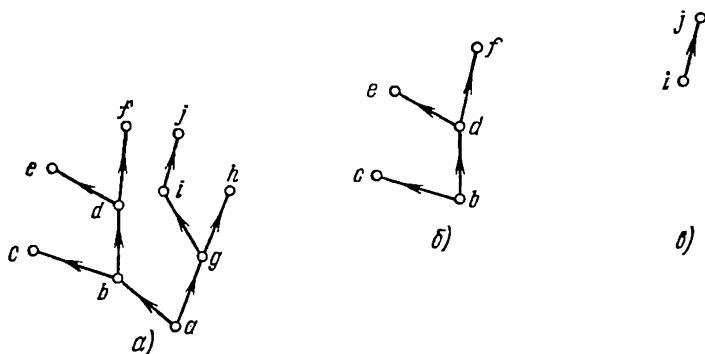


Рис. 5.10. Ориентированные поддеревья:
 a — ориентированное дерево с корнем a ; b — ориентированное
 поддерево с корнем b ; v — ориентированное дерево с корнем i

5.4. Матрицы инцидентности

Предположим, что сильносвязный ориентированный граф $\vec{G} = (V, \vec{E})$ имеет p вершин и q ребер. Пронумеруем вершины и ребра графа целыми числами $1, \dots, p$ и $1, \dots, q$ соответственно. Матрицей инцидентности графа \vec{G} называется матрица A_a размера $p \times q$, элементами которой являются действительные числа:

$$A_a = [\delta_{ij}]_{i=1, \dots, p; j=1, \dots, q}, \quad (5.3), (5.4)$$

где $\delta_{ij} = +1$, если ребро j выходит из вершины i ; $\delta_{ij} = -1$, если ребро j входит в вершину i , и $\delta_{ij} = 0$, если ребро j не инцидентно вершине i .

Каждый столбец этой матрицы имеет один элемент $+1$, один элемент -1 , а все остальные элементы столбца равны 0.

Упражнение 5.6. Постройте матрицу инцидентности ориентированного графа, изображенного на рис. 5.11.

Решение. Искомая матрица имеет следующие ненулевые компоненты:

$$A_a = \begin{bmatrix} \cdot & -1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ -1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & -1 & -1 \\ \cdot & \cdot & \cdot & -1 & -1 & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}. \quad (5.5)$$

Так как каждый столбец матрицы инцидентности содержит один элемент $+1$ и один элемент -1 , то после удаления из матрицы инцидентности одной произвольной строки информация не теряется. Другими словами, по усеченной таким образом матрице исходная матрица всегда может быть восстановлена однозначно. Матрица размера $(p-1) \times q$, которая получается из мат-

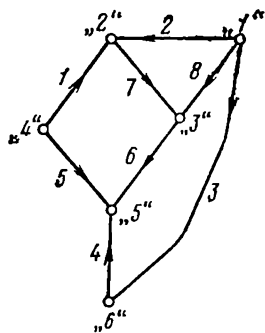


Рис. 5.11. Ориентированный граф из упражнения 5.6

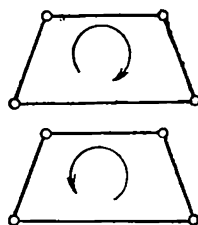


Рис. 5.12. Две ориентации ориентированного цикла

рицы инцидентности A_a удалением одной произвольной строки, называется *приведенной матрицей инцидентности* и обозначается через A . Вершина графа, соответствующая выкинутой строке, называется *базисной точкой*. Если в качестве базисной точки в приведенном примере, взять вершину 6, то после удаления последней строки получится следующая приведенная матрица инцидентности:

$$A = \begin{bmatrix} \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ -1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & -1 & -1 \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & -1 & -1 & \cdot & \cdot & \cdot \end{bmatrix} \quad (5.6)$$

Матрицу инцидентности можно определить и для неориентированных графов. Для этого достаточно символы -1 заменить на 1. Так же определяется и приведенная матрица инцидентности.

Строки матрицы инцидентности и приведенной матрицы инцидентности можно рассматривать как векторы q -мерного векторного пространства $V_q(R)$ над полем действительных чисел R .

Цикл в неориентированном графе можно ориентировать, как это показано на рис. 5.12. Выбрав ориентацию и задав в соответствии с ней ориентацию всех ребер, можно построить ориентированный цикл. Векторным представлением такого цикла называется вектор

$$(e_1, e_2, \dots, e_q), \quad (5.7)$$

где $\varepsilon_i = 1$, если ребро i принадлежит этому ориентированному циклу и ориентация цикла совпадает с заданной ориентацией в исходном ориентированном графе; $\varepsilon_i = -1$, если ребро i принадлежит ориентированному циклу и ориентация цикла не совпадает с заданной ориентацией в исходном ориентированном графе; $\varepsilon_i = 0$, если ребро i не принадлежит ориентированному циклу.

Например, в неориентированном графе, соотнесенном ориентированному графу, изображенному на рис. 5.11, последовательность «4», 1, «2», 7, «3», 6, «5», 5, «4» является циклом. Если ориентировать этот цикл по часовой стрелке, то получится ориентированный цикл, показанный на рис. 5.13. Векторное представление ориентированного цикла (1, ·, ·, ·, -1, 1, 1, ·).

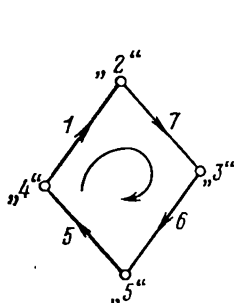


Рис. 5.13. Ориентированный цикл с ориентацией по ребру 1

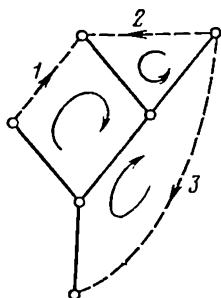


Рис. 5.14. Главный ориентированный цикл

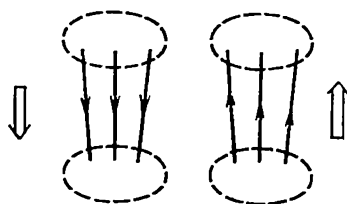


Рис. 5.15

Дерево, покрывающее неориентированный граф, соответствующий ориентированному, изображенному на рис. 5.11, показано на рис. 5.14 сплошной линией (см. рис. 4.12). Ребра являются хордами и показаны пунктирными линиями. Если для каждого главного цикла задать ориентацию в соответствии с ориентацией, определяющей его хорды, то получится *главный ориентированный цикл*. Матрица размера $\mu \times q = 3 \times 8$, строками которой являются три вектора, представляющих эти $\mu = 3$ главных ориентированных цикла, называется *матрицей главных ориентированных циклов* (см. упражнение 4.9):

$$M = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & -1 & 1 & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & 1 & -1 \\ \cdot & \cdot & 1 & \cdot & \cdot & -1 & \cdot & -1 \end{bmatrix}. \quad (5.8)$$

Для коциклов неориентированного графа можно рассмотреть два варианта ориентации, показанные на рис. 5.15, и точно так же, как и для ориентированных циклов, определить их векторное представление. Например, главный коцикл дерева, определяемый ветвью 6 (см. рис. 5.14), состоит из ветви 6 и хорд 1 и 3. Если ориентировать этот коцикл в соответствии с ориентацией ветви

6, то получится *главный ориентированный коцикл*, определяемый ветвью 6 (рис. 5.16). Его векторным представлением является $(-1, \cdot, 1, \cdot, \cdot, 1, \cdot, \cdot)$.

Матрица размера $\varphi \times q = 5 \times 8$, строками которой являются векторные представления этих $\varphi = 5$ главных ориентированных сечений, называется *матрицей главных ориентированных коциклов* (см. упражнение 4.9):

$$Q = \begin{bmatrix} \cdot & \cdot & -1 & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ -1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot \\ -1 & -1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}. \quad (5.9)$$

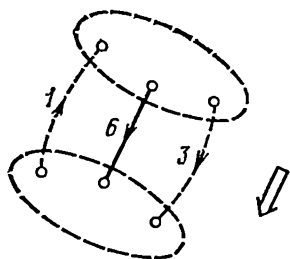


Рис. 5.16. Главный ориентированный коцикл, определяемый ветвью 6

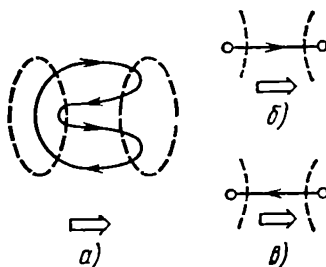


Рис. 5.17. Ориентированный цикл и ориентированный коцикл

Теорема 5.8. Векторные представления произвольного ориентированного цикла и произвольного ориентированного коцикла ортогональны.

Доказательство. Если рассматриваемые цикл и коцикл имеют общие ребра, то, как показано на рис. 5.17,а (см. следствие из теоремы 4.15), существует одинаковое число общих ребер типов б и в, показанных соответственно на рис. 5.17,б и в. Обозначим через $(\epsilon_1, \dots, \epsilon_q)$, $(\delta_1, \dots, \delta_q)$ векторные представления соответственно ориентированного цикла и ориентированного коцикла. Если i — ребро типа б, то $\epsilon_i = \delta_i = 1$ или $\epsilon_i = \delta_i = -1$, т. е. $\epsilon_i \delta_i = 1$. Если же i — ребро типа в, то $\epsilon_i = -\delta_i = 1$ или $\epsilon_i = \delta_i = -1$, т. е. $\epsilon_i \delta_i = -1$.

Следствие. Матрица главных ориентированных циклов M и матрица главных ориентированных коциклов Q таковы, что $MQ^T = 0$.

5.5. Закон Кирхгофа

$$\text{Вектор } v = (v_1, \dots, v_q), \quad (5.10)$$

называется вектором напряжений, если для произвольного ориентированного цикла

$$\varepsilon = (\varepsilon_1, \dots, \varepsilon_q) \quad (5.11)$$

выполняется равенство (закон Кирхгофа для напряжений)

$$(v, \varepsilon) = 0. \quad (5.12)$$

Множество всех векторов напряжений называется пространством напряжений \mathcal{V} . Если обозначить через v_k — потенциал вершины, из которой выходит ориентированное ребро k относительно вершины, в которую это ребро входит, то левая часть равенства (5.12) $\varepsilon_1 v_1 + \varepsilon_2 v_2 + \dots + \varepsilon_q v_q$ будет представлять собой сумму напряжений вдоль рассматриваемого ориентированного цикла. Требование равенства нулю этой суммы является законом Кирхгофа для напряжений.

$$\text{Вектор } i = (i_1, \dots, i_r) \quad (5.13)$$

называется вектором токов, если для произвольной строки

$$\delta = (\delta_1, \dots, \delta_q) \quad (5.14)$$

матрицы смежности A_a выполняется равенство (закон Кирхгофа для токов)

$$(i, \delta) = 0. \quad (5.15)$$

Множество векторов токов называется пространством токов и обозначается через \mathcal{I} . Если ориентацию ребра k принять за положительное направление тока (рис. 5.18), то левая часть равен-

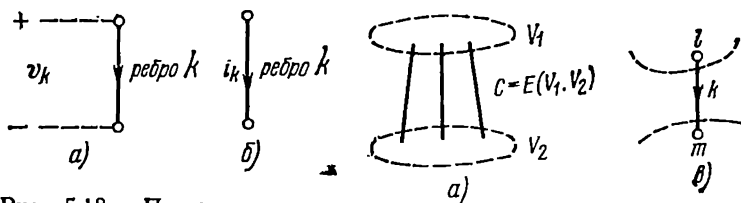


Рис. 5.18. Положительные направления напряжения и тока

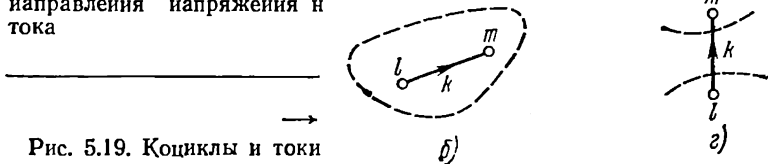


Рис. 5.19. Коциклы и токи

ства (5.15) $\delta_1 i_1 + \delta_2 i_2 + \dots + \delta_q i_q$ будет представлять алгебраическую сумму токов, вытекающих из рассматриваемого узла. Требование равенства нулю этой суммы является законом Кирхгофа для токов.

Возьмем произвольный коцикл C и исключим его из неориентированного графа $G = (V, E)$, соответствующего исходному ориентированному графу $\vec{G} = (V, \vec{E})$. Обозначим через (V, E_1) и (V, E_2) две связанные компоненты получившегося в результате графа. При этом $C = E(V_1, V_2)$. Предположим для простоты, что V_1 со-

стоит из вершин $1, \dots, n$, а V_2 — из $n+1, \dots, p$. Обозначим j -ю строку матрицы инцидентности Λ_a через $\delta^{(j)}$. Если i — вектор токов, то

$$(i, \delta^{(1)}) + \dots + (i, \delta^{(n)}) = 0. \quad (5.16)$$

Если ориентированное ребро k выходит из вершины $l \in V_1$ и входит в вершину $m \in V_1$, как показано на рис. 5.19,б, то в последнем равенстве слагаемыми, в которые входит i_k , являются $(i, \delta^{(l)})$, $(i, \delta^{(m)})$. Первое из них равно $+i_k$, а второе $-i_k$. Следовательно, токи i_k для ребер k , соединяющих вершины V_1 между собой, не входят в сумму в левой части (5.16). Далее, если ориентированное ребро k выходит из вершины $l \in V_1$ и входит в вершину $m \in V_2$, как это показано на рис. 5.19,в, то слагаемым в левой части (5.16), связанным с i_k , является произведение $(i, \delta^{(l)})$, которое представляет собой ток $+i_k$. Наоборот, если ориентированное ребро k выходит из вершины $l \in V_2$ и входит в вершину $m \in V_1$, как это показано на рис. 5.19,г, то в левой части (5.16) ток i_k представлен слагаемым $(i, \delta^{(m)})$, которое представляет собой ток $-i_k$. Таким образом, равенство (5.16) можно записать в виде

$$\sum_{k \in C} \epsilon_k i_k = 0, \quad (5.17)$$

где $\epsilon_k = 1$, если ориентированное ребро k выходит из V_1 и входит в V_2 ; и $\epsilon_k = -1$, если ориентированное ребро k выходит из V_2 и входит в V_1 .

Если коцикл C ориентирован таким образом, что его ребра выходят из V_1 и входят в V_2 и ϵ — векторное представление коцикла, то равенство (5.17) эквивалентно равенству

$$(i, \epsilon) = 0. \quad (5.18)$$

Далее, предположим, что равенство (5.18) справедливо для всех ориентированных коциклов, и покажем, что в этом случае i — вектор токов. Для произвольной вершины j множество инцидентных ей ребер называется звездой вершины j и обозначается через $\delta(j)$. Так как в неориентированном графе, согласно теореме 4.16, звезда $\delta(j)$ является кограницей, то в нем существуют такие коциклы без общих ребер, что объединение множеств их вершин совпадает с $\delta(j)$. Если все коциклы ориентированы таким образом, что их ребра выходят из вершины j , то, очевидно,

$$\delta^{(j)} = \epsilon^{(1)} + \dots + \epsilon^{(l)}, \quad (5.19)$$

где $\delta^{(j)}$ — j -я строка матрицы инцидентности Λ_a , соответствующая вершине j , а $\epsilon^{(1)}, \dots, \epsilon^{(l)}$ — векторные представления ориентированных коциклов, которые получаются при ориентации C_1, \dots, C_l . Следовательно, из равенств $(i, \epsilon^{(1)}) = \dots = (i, \epsilon^{(l)}) = 0$ следует, что $(i, \delta^{(j)}) = 0$.

Из изложенного следует, что вектор i , удовлетворяющий равенству (5.18) для произвольного ориентированного коцикла, является вектором токов, удовлетворяющим закону Кирхгофа для токов. Таким образом, справедлива следующая теорема.

Теорема 5.9. Вектор является вектором токов тогда и только тогда, когда он ортогонален векторным представлениям всех ориентированных коциклов.

Лемма 5.2. Пространство токов \mathcal{I} и пространство напряжений \mathcal{V} являются подпространствами векторного пространства $V_q(R)$.

Доказательство. Пусть $i^{(1)}, i^{(2)} \in \mathcal{I}$, $\alpha \in R$ (α — действительное число). Так как для векторного представления ϵ произвольного ориентированного коцикла $(i^{(1)}, \epsilon) = (i^{(2)}, \epsilon) = 0$, то $(i^{(1)} + i^{(2)}, \epsilon) = 0$, $(\alpha i^{(1)}, \epsilon) = 0$. Отсюда следует, что \mathcal{I} — подпространство. Для доказательства того, что \mathcal{V} является подпространством, следует рассмотреть ориентированные циклы.

Пусть G — неориентированный граф, соответствующий ориентированному графу \vec{G} . Определим дерево, покрывающее G , как это было описано в предыдущем параграфе. Тогда можно определить матрицу главных ориентированных циклов M и матрицу главных ориентированных сечений Q .

Согласно теореме 5.8, произвольная строка M и произвольная строка Q ортогональны. Так как μ строк матрицы M линейно независимы, то подпространство \mathcal{I}' , натянутое на μ строк матрицы M , и подпространство \mathcal{V}' , натянутое на φ линейно независимых строк Q , имеют соответственно размерности μ и φ . Заметив, что $\mu + \varphi = q$, приходим к тому, что подпространства \mathcal{I}' и \mathcal{V}' взаимно ортогональны. Следовательно, имеет место следующая лемма.

Лемма 5.3. Подпространство \mathcal{I}' , натянутое на μ строк матрицы главных ориентированных циклов, и подпространство \mathcal{V}' , натянутое на φ строк матрицы главных ориентированных коциклов, имеют соответственно размерности μ , φ и взаимно ортогональны.

Из теорем 5.8 и 5.9 следует, что каждая строка матрицы M — вектор токов. При этом из теоремы 5.8 также следует, что каждая строка матрицы Q является вектором напряжений. Это означает, что $\mathcal{I}' \subset \mathcal{I}$, $\mathcal{V}' \subset \mathcal{V}$.

Наоборот, из ортогональности подпространств \mathcal{I}' и \mathcal{V}' следует, что \mathcal{I}' представляет собой множество векторов, ортогональных каждой из строк матрицы Q . Так как, согласно теореме 5.9, этим свойством обладает произвольный вектор тока, то $\mathcal{I} \subset \mathcal{I}'$, т. е. $\mathcal{I} = \mathcal{I}'$.

Аналогично можно показать, что $\mathcal{V} \subset \mathcal{V}'$, т. е. $\mathcal{V} = \mathcal{V}'$.

Таким образом, мы доказали следующую теорему.

Теорема 5.10. Пространство токов \mathcal{I} и пространство напряжения \mathcal{V} являются взаимно ортогональными. Базисом \mathcal{I} являются μ вектор-строк матрицы главных ориентированных циклов, а базисом \mathcal{V} — φ вектор-строк матрицы главных ориентированных коциклов.

Следствие (теорема Телегена). Для произвольных вектора тока i и вектора напряжений v справедливо равенство $(i, v) = 0$.

По аналогии с доказательством теоремы 4.21 с помощью следствия из теоремы 5.8 можно получить следующую теорему.

Теорема 5.11. Предположим, что хорды и ветви дерева, покрывающего граф G , соотносены ориентированному сильносвяз-

ному графу \vec{G} с корангом φ и цикломатическим числом μ , перенумерованы целыми числами $1, \dots, \mu$ и $\mu+1, \dots, q$ соответственно. Тогда матрица главных ориентированных циклов M и матрица главных ориентированных коциклов Q имеют вид

$$M = [I_\mu, F]; \quad Q = [-F^T, I_\varphi].$$

Задачи

5.1. Докажите, что ориентированный граф без ориентированных циклов имеет, по крайней мере, одну вершину, положительная степень которой равна 0.

5.2. Пусть для произвольных двух вершин a и b ориентированного графа \vec{G} существует вершина c такая, что $c \Rightarrow a$ и $c \Rightarrow b$. Докажите, что в этом случае \vec{G} имеет корень.

5.3. На ориентированном графе, изображенном на рис. 5.20, сплошными линиями показаны хорды, а пунктирными — ветви. Найдите матрицу главных ориентированных циклов и матрицу главных ориентированных сечений.

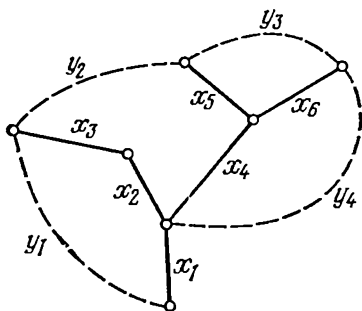


Рис. 5.20. К задаче 5.3

5.4. Пусть A — приведенная матрица инцидентности размера $\varphi \times q$ сильносвязного ориентированного графа. Пусть F — это некоторое подмножество ребер неориентированного графа, соответствующего ориентированному графу с матрицей инцидентности A . Покажите, что если F содержит циклы, то вектор-столбцы матрицы A , соответствующие ребрам этих циклов, являются линейно зависимыми.

5.5. Как и ранее, пусть T — дерево, покрывающее неориентированный граф, соответствующий ориентированному графу. Докажите, что φ вектор-столбцов, соответствующих ребрам дерева T , являются базисом φ -мерного векторного пространства $V_\varphi(R)$ над полем действительных чисел R .

Глава 6

Булевы алгебры

6.1. Определение булевой алгебры [2,10]

Рассмотрим два множества элементов A и B . Множество, состоящее из элементов множества A и элементов множества B , т. е. полученное объединением множеств A и B , называется об-

единением или *суммой* множеств A и B и обозначается через $A \cup B$. Знак \cup называется знаком объединения. Множество, состоящее из элементов, входящих одновременно и в множество A и в множество B , называется *пересечением* или *произведением* множеств A и B и обозначается через $A \cap B$. Знак \cap называется знаком пересечения. Совокупность элементов, входящих в множество A , но не входящих в множество B , называется *разностью* множеств A и B и обозначается через $A - B$. Число элементов конечного множества A обозначается через $|A|$.

Совокупность всех подмножеств некоторого множества U , включая само множество U и пустое множество Φ , обозначается через $p(U)$. Для произвольного множества $A \in p(U)$ можно рассмотреть множество $U - A$, являющееся *дополнением* A (до U); далее дополнение множества A обозначается через A^c . Для произвольных множеств $A, B, C \in p(U)$ выполняются следующие хорошо известные соотношения.

1. Закон тождественности: $A \cup A = A, A \cap A = A$.

2. Закон ассоциативности: $(A \cup B) \cup C = A \cup (B \cup C), (A \cap B) \cap C = A \cap (B \cap C)$.

3. Закон коммутативности: $A \cup B = B \cup A, A \cap B = B \cap A$.

4. Закон поглощения: $(A \cup B) \cap A = A, (A \cap B) \cup A = A$.

5. Закон дистрибутивности: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C), A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

6. Закон дополнения: а) $A \cup U = U, A \cap U = A$; б) $A \cup \emptyset = A, A \cap \emptyset = \emptyset$; в) $A \cup A^c = U, A \cap A^c = \emptyset$.

Задача 6.1. Убедитесь в справедливости приведенных выше соотношений.

Прямым произведением $A \times B$ двух множеств, A и B , (возможно, совпадающих) является совокупность всех упорядоченных пар элементов (a, b) , в которых элемент a берется из множества A , а элемент b берется из B . Аналогично определяется прямое произведение трех и большего числа множеств. Если A — непустое множество, то отображение φ прямого произведения $A \times A$ на A называется *бинарной операцией*. При этом элемент множества A , соответствующий паре $(a_1, a_2) \in A \times A$, обозначают обычно через $\varphi(a_1, a_2)$ или $a_1 \varphi a_2$. Аналогично, отображение множества A на A называют *унарной операцией*.

Пусть M — непустое множество, на котором определены две бинарные операции « \cup » и « \cap » (эти же символы были использованы ранее для обозначения операций над множествами; вообще говоря, здесь \cup и \cap — произвольные операции) такие, что для произвольных элементов $a, b, c \in M$ выполняются следующие соотношения.

1. Закон тождественности: $a \cup a = a, a \cap a = a$.

2. Закон ассоциативности: $(a \cup b) \cup c = a \cup (b \cup c), (a \cap b) \cap c = a \cap (b \cap c)$.

3. Закон коммутативности: $a \cup b = b \cup a, a \cap b = b \cap a$.

4. Закон поглощения: $(a \cup b) \cap a = a, (a \cap b) \cup a = a$.

Тогда говорят, что множество M образует *решетку* относительно операций \cup и \cap (часто M называют просто решеткой).

Задача 6.2. Выведите закон тождественности из законов коммутативности и поглощения.

Решение. Если в равенстве $(a \cap b) \cup a = a$ вместо b подставить $a \cup b$, то получим $a = (a \cap (a \cup b)) \cup a = ((a \cup b) \cap a) \cup a = a \cup a$. Заменяя далее \cup на \cap , имеем $a \cap a = a$.

Законы ассоциативности, коммутативности и поглощения называют *аксиомами решетки*. Если для решетки верно какое-либо общее утверждение, то из него с помощью приведенных аксиом можно получить так называемое двойственное утверждение, поменяв местами в исходном утверждении знаки \cup и \cap . Это свойство решетки называют *законом двойственности*.

Если для произвольных элементов a, b, c решетки выполняется также

5) закон дистрибутивности: $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$, $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$, то M называется *дистрибутивной решеткой*. Если для дистрибутивной решетки M выполняется, кроме того,

6) закон дополнения: в M существуют такие элементы I и O , что: а) $a \cup I = I$, $a \cap I = a$, $a \cup O = a$, $a \cap O = O$; б) для произвольного элемента a из M в M найдется элемент a' такой, что $a \cup a' = I$, $a \cap a' = O$, то M называется *булевой алгеброй*. Элемент a' из условия б) закона дополнения называется *дополнительным* элементом a . Решетки такого типа называют *булевыми алгебрами* в честь Джорджа Буля, который предпринял попытку алгебраизации теории решеток путем введения операций над элементами.

Совокупность всех подмножеств $p(U)$ некоторого множества U образует булеву алгебру относительно операций над множествами \cup и \cap . Эту совокупность $p(U)$ называют *алгеброй множеств*. В качестве элементов I и O в законе дополнения следует взять соответственно множества U и \emptyset ; при этом дополнительным для множества $A \in p(U)$ является множество A^c . В § 7.2 будет показано, что булевой алгеброй является также совокупность всех логических функций от n переменных.

Задача 6.3. Покажите, что элементы I и O закона дополнения определяются однозначно.

Решение. Предположим, что существует другой отличный от I элемент J , удовлетворяющий первому соотношению условия а) закона дополнения. Тогда $I = J \cup I = I \cup J = J$. Единственность O доказывается из четвертого соотношения того же условия а).

Упражнение 6.1. Закон дополнения требует существования дополнительного элемента. Покажите, что в булевой алгебре для каждого элемента существует только один дополнительный элемент. Вообще говоря, для произвольной дистрибутивной решетки можно показать, что из условий $a \cup x = a \cup y$ и $a \cap x = a \cap y$ вытекает, что $x = y$.

Решение:

$$x = (a \cup x) \cap x = (a \cup y) \cap x \quad (\text{коммутативность, закон поглощения, подстановка});$$

$$x = (a \cap x) \cup (y \cap x) \quad (\text{коммутативность, дистрибутивность});$$

$$x = (a \cap y) \cup (x \cap y) \quad (\text{коммутативность, подстановка});$$

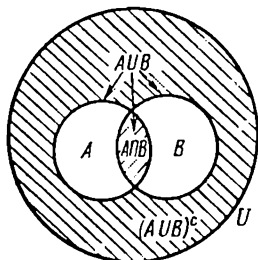
$$x = (a \cup x) \cap y = (a \cup y) \cap y \quad (\text{коммутативность, дистрибутивность, подстановка});$$

$$x = y \quad (\text{коммутативность, закон поглощения}).$$

Замечание. Если выполняется только одно из равенств $a \cup x = a \cup y$ (или $a \cap x = a \cap y$), то, вообще говоря, неверно, что $x = y$. Например, если $a \neq O$, то $a = a \cup a = a \cup O$.

Из существования и единственности дополнительного элемента следует, что операция, сопоставляющая произвольному элементу $a \in M$ дополнительный элемент a' , является унарной операцией.

Упражнение 6.2. В совокупности $p(U)$ всех подмножеств некоторого множества U выполняется хорошо известное правило Де Моргана: для произвольных подмножеств $A, B \in p(U)$



$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$$

(в справедливости этих соотношений можно убедиться, воспользовавшись рис. 6.1).

Рис. 6.1. Правило Де Моргана

Правило Де Моргана для произвольной булевой алгебры M , а именно для соотношения $(a \cup b)' = a' \cap b'$, $(a \cap b)' = a' \cup b'$, (где a, b — произвольные элементы алгебры M) можно доказать следующим образом:

$$\begin{aligned} (a \cup b) \cup (a' \cap b') &= ((a \cup b) \cup a') \cap ((a \cup b) \cup b') \quad (\text{дистрибутивность}); \\ &= ((a \cup a') \cup b) \cap (a \cup (b \cup b')) \quad (\text{коммутативность, ассоциативность}); \\ &= (I \cup b) \cap (a \cup I) \quad (\text{закон дополнения}); \\ &= (b \cup I) \cap I \quad (\text{коммутативность, закон дополнения}); \\ &= I \cap I = I \quad (\text{закон дополнения}); \end{aligned}$$

$$\begin{aligned} (a \cup b) \cap (a' \cap b') &= (a' \cap b') \cap (a \cup b) \quad (\text{коммутативность}); \\ &= ((a' \cap b') \cap a) \cup ((a' \cap b') \cap b) \quad (\text{дистрибутивность}); \\ &= (b' \cap (a \cap a')) \cup (a' \cap (b \cap b')) \quad (\text{коммутативность, ассоциативность}); \\ &= (b' \cap O) \cup (a' \cap O) \quad (\text{закон дополнения}); \\ &= O \cup O = O \quad (\text{закон дополнения}). \end{aligned}$$

Таким образом, $a' \cap b'$ является дополнительным к элементу $a \cup b$. Из единственности дополнительного элемента получаем первое доказываемое соотношение. Второе соотношение можно получить, если в первом соотношении поменять местами знаки \cup и \cap (проверьте это).

Задача 6.4. Покажите, что для произвольного элемента a булевой алгебры M справедливо равенство $(a')' = a$.

Решение. Из закона дополнения (соотношение б) и коммутативности следует, что a — дополнительный элемент элемента a' . Отсюда и из единственности дополнительного элемента получаем, что $a = (a')'$.

Задача 6.5. Обратными к элементам I и O являются соответственно O и I .

Решение. Справедливость этого утверждения следует непосредственно из закона дополнения.

Задача 6.6. Пусть a, b — два произвольных элемента решетки. Покажите, что $a \cup b = b \Leftrightarrow a \cap b = a^*$.

Решение. Пусть $a \cup b = b$. Тогда $a = (a \cup b) \cap a = b \cap a = a \cap b$. Обратное утверждение доказывается аналогично.

Упражнение 6.3. Если в решетке справедливо соотношение а) $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$, то справедливо также и соотношение б) $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$. Можно, наоборот, из соотношения б) вывести соотношение а).

Решение.

$$\begin{aligned} (a \cup b) \cap (a \cup c) &= ((a \cup b) \cap a) \cup ((a \cup b) \cap c) \text{ (соотношение а);} \\ &= a \cup (c \cap (a \cup b)) \text{ (поглощение, коммутативность);} \\ &= a \cup ((a \cap c) \cup (b \cap c)) \text{ (соотношение а, коммутативность);} \\ &= (a \cup (a \cap c)) \cup (b \cap c) \text{ (ассоциативность);} \\ &= a \cup (b \cap c) \text{ (закон поглощения).} \end{aligned}$$

Справедливость обратного утверждения следует из закона двойственности.

Замечания к законам дистрибутивности и дополнения. Как видно из приведенных примеров для дистрибутивности решетки, в действительности, достаточно выполнение хотя бы одного соотношения закона дистрибутивности. Далее, из задачи 6.6 следует, что в условии а) закона дополнения достаточно, например, требовать лишь выполнения равенств $a \cup I = I$ и $a \cap O = O$.

Так как операции \cup и \cap удовлетворяют закону ассоциативности, то в соотношениях вида $(a \cup b) \cup c = a \cup (b \cup c)$, $(a \cap b) \cap c = a \cap (b \cap c)$ скобки можно опускать и записывать их в виде $a \cup b \cup c$, $a \cap b \cap c$.

6.2. Отношение порядка в булевой алгебре [2,10—12]

Подмножество R прямого произведения $A \times B$ множеств A, B (возможно $A=B$) называется *отношением* на паре множеств A, B (на множестве A , если $A=B$). Соотношение $(a, b) \in R$ часто записывают в виде aRb и говорят, что элемент a связан с b отношением R . Если отношение R на множестве A таково, что для произвольных элементов a, b и c из A выполняются:

- 0.1) закон рефлексивности: aRa ;
- 0.2) закон антисимметричности: если aRb и bRa , то $a=b$;
- 0.3) закон транзитивности: если aRb и bRc , то aRc ,

* Если справедливо утверждение A и из него следует утверждение B , то мы записываем это так: $A \Rightarrow B$. Если $A \Rightarrow B$ и $B \Rightarrow A$, то пишем $A \Leftrightarrow B$.

то, как уже указывалось § 1.1, отношение R называется *отношением порядка**. Пусть R — отношение порядка на множестве A . Если отношение R' определено таким образом, что для произвольных элементов $a, b \in A$, $aR'b \Leftrightarrow bRa$, то для R' будут выполняться законы рефлексивности, антисимметричности, транзитивности и, следовательно, оно также будет отношением порядка. Это отношение R' называют *противоположным* отношению R .

Задача 6.7. Покажите, что R' является отношением порядка.

Предположим, что на множестве M задано отношение порядка R (далее вместо R будем пользоваться знаком \leq). Если $a \leq b$, то говорят, что a меньше или равно b или что b больше или равно a . Если $a \leq b$ и $a \neq b$, то будем писать $a < b$ и говорить, что a (строго) меньше или что b (строго) больше a . Если $a < b$ и в множестве M не существует ни одного элемента x такого, что $a < x < b$, то говорят, что b непосредственно следует за a , и записывают это так: $a \prec b$.

Элементы a и b называются *несравнимыми*, если не выполняется ни одно из соотношений $a \leq b$, $b \leq a$. Если же хотя бы одно из этих соотношений выполняется, то элементы называются *сравнимыми*. Пусть A — подмножество множества M . Элемент $a \in A$ такой, что $x \leq a$ для всех $x \in A$, называется *максимальным элементом* A и обозначается через $\max A$. Аналогично определяется *минимальный элемент* множества A , который обозначается через $\min A$. Вообще говоря, в множестве A элементы $\max A$ и $\min A$ могут не существовать. Если $a \in A$ и в A не существует ни одного элемента, больше a , то a называют *предельным максимальным элементом* множества A . Аналогично определяется *предельный минимальный элемент*. Максимальный (минимальный) элемент множества является предельным максимальным (минимальным) элементом, но обратное утверждение, вообще говоря, не верно. Для элементов a, b ($a \leq b$) определим $[a, b] = \{x | a \leq x \leq b, x \in M\}$.

Задача 6.8. Если в множестве A существует максимальный элемент, то он единственен.

Указание. Следует предположить существование двух максимальных элементов и далее воспользоваться законом антисимметричности отношения порядка.

Упражнение 6.4. В непустом конечном упорядоченном множестве M существуют предельные максимальный и минимальный элементы. Кроме того, для элементов a, b из M таких, что $a < b$, в M существуют элементы a_1, a_2, \dots, a_{m-1} такие, что $a \prec a_1 \prec \dots \prec a_m = b$.

Решение. Доказательство первого утверждения. Обозначим через n число элементов в множестве M и воспользуемся индукцией по n . При $n=1$ утверждение тривиально. Пусть $n \geq 2$. Предположим, что утверждение верно для упорядоченного множества, состоящего из $n-1$ элементов. Возьмем некоторый эле-

* Заметим, что приведенное определение является, строго говоря, определением частичной упорядоченности и отличается от определения полной упорядоченности.

мент a в M и рассмотрим множество $M - \{a\}$. По предположению индукции в этом множестве существует предельный максимальный элемент b . Если $b < a$, то a — предельный максимальный элемент M . Если $a < b$ или a и b несравнимы, то предельным максимальным элементом в M является b . Существование предельного минимального элемента доказывается аналогично.

Доказательство второго утверждения. Так как множество $[a, b] - \{a\}$ непусто, то в нем существует предельный минимальный элемент a_1 . По определению, $a | a_1 \leq b$. Если $a_1 < b$, то множество $[a_1, b] - \{a_1\}$ непусто и в нем существует предельный минимальный элемент a_2 , так что $a | a_1 | a_2 \leq b$. Если $a_2 < b$, то описанную операцию нужно повторить. Так как множество M конечно, то число таких операций также будет конечным. Второе утверждение доказано.

Упорядоченное множество M , имеющее конечное число элементов, удобно представлять с помощью направленного графа $H(M)$, который называют *диаграммой Хассе* и строят следующим образом. Вершины $H(M)$ взаимно однозначно соответствуют элементам M . При этом вершину, соответствующую элементу a , называют просто вершиной a . Если элемент b непосредственно следует за элементом a (и только в этом случае), вершину b соединяют ребром с вершиной a ; при этом ребро ориентируется от b к a . На рис. 6.2 приведено несколько примеров таких диаграмм (в некоторых случаях большие элементы помещают над меньшими, а ребра не ориентируют). Рисунок 6.2,а соответствует вполне упорядоченному множеству; остальные рисунки соответствуют частично упорядоченным множествам.

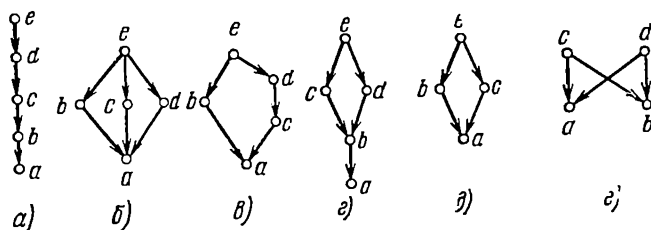


Рис. 6.2. Примеры диаграмм Хассе

Задача 6.9. Соотношение $a < b$ эквивалентно существованию направленного пути из вершины b в вершину a на диаграмме Хассе.

Указание. Существование пути можно вывести из определения диаграммы Хассе и второго утверждения из упражнения 6.4. Обратное утверждение получается из закона транзитивности.

Задача 6.10. Диаграмма Хассе обладает следующими свойствами. 1) Если на диаграмме существует ребро e , идущее из вершины b в вершину a , то никакого другого пути из b в a быть не может. 2) Диаграмма Хассе не имеет замкнутых направленных циклов.

Указание. 1) Предположим, что на диаграмме существует путь, отличный от e . Если длина этого пути равна 1, то приходим к противоречию с определением диаграммы Хассе. Если же длина пути больше или равна 2, то при-

ходим к противоречию с отношением $a|<b$. 2) Если допустить существование на диаграмме замкнутого цикла, то из задачи 6.9 будет следовать, что для вершины a замкнутого цикла выполняется отношение $a<a$, чего быть не может.

Задача 6.11. Для любого конечного направленного графа G , обладающего свойствами 1) и 2) из задачи 6.10, существует упорядоченное множество M , для которого G является диаграммой Хассе.

Указание. Пусть a и b — элементы множества вершин V графа G . Если $a=b$ или на графе существует направленный путь из b в a , то будем считать, что $a\leq b$. Воспользовавшись определением и свойством 2), можно показать, что введенное отношение \leq удовлетворяет законам рефлексивности, антисимметричности и транзитивности. Далее, воспользовавшись свойством 1), можно показать, что если на графе имеется направленное ребро от b к a , то $a|<b$.

Рассмотрим упорядоченное множество M и некоторое его подмножество A . Если в M существует элемент b , обладающий указанными далее свойствами 1) и 2), то b называется *верхней гранью* A и обозначается через $\sup A$:

1) $a\leq b$ для всех элементов $a\in A$;

для каждого элемента $a\in A$ и произвольного элемента $c\in M$ такого, что $a<c$, выполняется соотношение $b\leq c$.

Если в множестве M существует элемент b , обладающий указанными далее свойствами 1') и 2'), то b называется *нижней гранью* A и обозначается через $\inf A$:

1) $b\leq a$ для всех элементов $a\in A$.

2) для произвольного элемента $c\in M$ такого, что $c\leq a$ для всех $a\in A$ выполняется соотношение $c\leq b$.

Заметим, что элементы $\sup A$ и $\inf A$, вообще говоря, могут не принадлежать множеству A . Если в множестве существует элемент $\max A$ ($\min A$), то $\sup A = \max A$ ($\inf A = \min A$).

Пример 6.1. На рис. 6.2,б $e = \sup\{b, c\} = \sup\{b, c, d\} = \sup\{a, b, c, d\}$. На рис. 6.2,в $a = \inf\{b, c\} = \inf\{b, d\} = \inf\{b, c, d\} = \inf\{b, c, d, e\}$. На рис. 6.2,е элементы $\sup\{a, b\}$, $\sup\{c, d\}$, $\inf\{a, b\}$, $\inf\{c, d\}$ не существуют.

Пример 6.2. Отношение включения \subset , определенное между подмножествами некоторого множества U , является отношением порядка на множестве $p(U)$ всех подмножеств U . Воспользовавшись определениями, легко проверить, что для произвольных множеств $A, B \in p(U)$ выполняются следующие соотношения (проверьте).

(M1) $A \subset A \cup B$, $B \subset A \cup B$ и для произвольного множества $C \in p(U)$ такого, что $A \subset C$, $B \subset C$, имеет место включение $A \cup B \subset C$, т. е. $A \cup B = \sup\{A, B\}$.

(M2) $A \cap B \subset A$, $A \cap B \subset B$ и для произвольного множества $C \in p(U)$ такого, что $C \subset A$, $C \subset B$, имеет место включение $C \subset A \cap B$, т. е. $A \cap B = \inf\{A, B\}$.

(M3) $A \subset B \Leftrightarrow A \cup B = B \Leftrightarrow A \cap B = A$.

Как мы увидим далее, эти свойства присущи произвольной решетке.

У п р а ж н е н и е 6.5. Предположим, что в непустом множестве M определено некоторое отношение порядка \leq и для произвольных элементов $a, b \in M$ существуют элементы $\sup\{a, b\}$ и $\inf\{a, b\}$. Если определить операции \cup и \cap , положив $a \cup b = \sup\{a, b\}$, $a \cap b = \inf\{a, b\}$, то для операций \cup и \cap будут выполняться законы ассоциативности, коммутативности и поглощения, т. е. мы получим решетку.

Решение. Если вместо отношения \leq рассмотреть двойственное к нему отношение порядка, то определенные операции \cup и \cap можно поменять местами. Следовательно, в аксиомах структуры достаточно рассмотреть лишь соотношения, соответствующие одной из этих операций. Кроме того, по определению \sup и \inf имеем:

$$а) \ a \leq a \cup b, \ b \leq a \cup b, \ a \cap b \leq a, \ a \cap b \leq b;$$

$$б) \ \text{если } a \leq c, \ b \leq c, \ \text{то } a \cup b \leq c;$$

$$в) \ \text{если } c \leq a, \ c \leq b, \ \text{то } c \leq a \cap b.$$

Ассоциативность. Доказательство равенства $(a \cup b) \cup c = a \cup (b \cup c)$. Из а) следует, что

$$a \leq a \cup b \leq (a \cup b) \cup c. \quad (6.1)$$

Аналогично

$$b \leq (a \cup b) \cup c, \ c \leq (a \cup b) \cup c. \quad (6.2), (6.3)$$

Из (6.2), (6.3) и б) получаем $b \cup c \leq (a \cup b) \cup c$. Отсюда из (6.1) и вновь б) имеем $a \cup (b \cup c) \leq (a \cup b) \cup c$. Совершенно аналогично выводится соотношение $(a \cup b) \cup c \leq a \cup (b \cup c)$. Из закона антисимметричности следует справедливость закона ассоциативности.

Закон коммутативности следует из определения операций \cup и \cap .

Закон поглощения. Доказательство равенства $(a \cup b) \cap a = a$. С одной стороны, из а) следует, что

$$(a \cup b) \cap a \leq a. \quad (6.4)$$

С другой стороны, так как $a \leq a \cup b$ и $a \leq a$, то, в силу в), $a \leq (a \cup b) \cap a$. Из (6.4) и закона антисимметричности следует справедливость закона поглощения.

Наоборот, утверждение (М3) из примера 6.2 можно обобщить на случай произвольных решеток.

У п р а ж н е н и е 6.6. Пусть M — произвольная решетка. Определим отношение \leq на M следующим образом. Для произвольных элементов a и b из M положим

$$a \leq b \Leftrightarrow a \cup b = b. \quad (6.5)$$

Отношение \leq является отношением порядка и обладает следующими свойствами.

1. **Закон двойственности.** Если для решетки верно какое-либо утверждение, то для нее верно также и двойственное утверждение, которое получается, если поменять местами операции \cup и \cap и заменить отношение \leq на противоположное \geq .

2. Для произвольных элементов $a, b, c \in M$

$$a \leq a \cup b; \ a \cap b \leq a. \quad (6.6)$$

3. Если $a \leq c, b \leq c$, то

$$a \cup b \leq c. \quad (6.7)$$

3'. Если $c \leq a$, $c \leq b$, то

$$c \leq a \cap b. \quad (6.8)$$

4. Существуют элементы $\sup\{a, b\}$, $\inf\{a, b\}$, причем

$$\sup\{a, b\} = a \cup b; \quad \inf\{a, b\} = a \cap b. \quad (6.9)$$

Решение. Согласно задаче 6.6, $a \cup b = b \Leftrightarrow a \cap b = a$. Отсюда и (6.5) следует, что

$$b \leq a \Leftrightarrow a \cap b = b. \quad (6.10)$$

(1) Из закона двойственности в решетке и справедливости двойственного к (6.5) соотношения (6.10) следует справедливость закона двойственности, сформулированного в пункте (1). Следовательно, нам достаточно доказать левое соотношение в (6.6), соотношение (6.7) и левое соотношение в (6.9).

(2) Из закона тождественности и соотношения (6.5) следует, что $a \leq a$.

(3) Если $a \leq b$ и $b \leq a$, то из (6.5) следует, что $a \cup b = b$ и $b \cup a = a$. Отсюда и из закона коммутативности $a \cup b = b \cup a$ получаем, что $a = b$.

(4) Если $a \leq b$ и $b \leq c$, то из (6.5) следует, что $a \cup b = b$ и $b \cup c = c$. Отсюда и из закона коммутативности получаем, что $a \cup c = a \cup (b \cup c) = (a \cup b) \cup c = b \cup c = c$. Таким образом, отношение \leq удовлетворяет законам рефлексивности, антисимметричности и транзитивности, а следовательно, является отношением порядка.

(5) В силу законов поглощения и коммутативности

$$a = (a \cup b) \cap a = a \cap (a \cup b). \quad (6.11)$$

При этом из (6.10) следует справедливость левого соотношения из пункта (2)

$$a \leq a \cup b. \quad (6.12)$$

Следовательно, имеет место также соотношение

$$b \leq a \cup b. \quad (6.13)$$

С другой стороны, если c — элемент M такой, что $a \leq c$ и $b \leq c$, то согласно (6.5) $a \cup c = c$, $b \cup c = c$. В силу ассоциативности $(a \cup b) \cup c = a \cup (b \cup c) = a \cup c = c$. Из (6.5) следует справедливость свойства (3), а именно соотношения

$$a \cup b \leq c. \quad (6.14)$$

Из (6.12)—(6.14) и определения верхней грани $\sup\{a, b\}$ получаем левое соотношение из (6.9): $\sup\{a, b\} = a \cup b$.

Еще одно определение решетки. Как видно из упражнений 6.5 и 6.6, понятие решетки можно также определить следующим образом:

«Если для любых двух элементов a и b упорядоченного множества M существуют элементы $\sup\{a, b\}$ и $\inf\{a, b\}$, то M называется решеткой». Это определение полностью эквивалентно предыдущему.

Задача 6.12. Проверьте, что упорядоченные множества с диаграммами Хассе, приведенными на рис. 6.2, а—д, являются решетками, а множество с диаграммой Хассе, приведенной на рис. 6.2, е, решеткой не является.

В дальнейшем при рассмотрении решеток и булевых алгебр мы будем пользоваться только отношением порядка \leq , которое определяется соотношением 6.5 (далее оно называется отношением больше/меньше). Утверждение а) закона дополнения можно переписать в следующем виде: для всех элементов a из M : $0 \leq$

$\leq a \leq I$. Другими словами, первое утверждение закона дополнения гарантирует существование в M максимального элемента I и минимального элемента O . Если решетка удовлетворяет закону дополнения, то этому закону будет удовлетворять и двойственное упорядоченное множество, которое получается при замене отношения больше/меньше \leq противоположным отношением \geq . Кроме того, так как из соотношений двойственности следует также закон дистрибутивности, то для булевой алгебры из справедливости утверждения p следует справедливость *двойственного утверждения*, которое получается при замене местами отношений \leq и \geq , \cup и \cap , а также максимального и минимального элементов. Это свойство булевой алгебры называют *законом двойственности*.

Задача 6.13. Покажите, что если диаграмма Хассе решетки M содержит в качестве своего подграфа граф, изображенный на рис. 6.2,б или в, то для соответствующей решетки не выполняется закон дистрибутивности.

Указание. Для графа, изображенного на рис. 6.2,б, с одной стороны, имеем $d \cap (b \cap c) = d \cap e = d$, а с другой — $d \cap d = a$, $d \cap c = a$, $(d \cap b) \cup (d \cap c) = a$. Для графа, изображенного на рис. 6.2,в, $d \cap (b \cup c) = d \cap e = d$, $d \cap b = a$, $d \cap c = c$, $(d \cap b) \cup (d \cap c) = c$.

Задача 6.14. Решетки с диаграммами Хассе, изображенными на рис. 6.2,а и г, не удовлетворяют требованию существования дополнительного элемента из пункта б) закона дополнения.

Указание. Проверьте, что в данном случае решетка имеет максимальный элемент e и минимальный элемент a , но не имеет обратного элемента для элемента c .

Задача 6.15. Решетки с диаграммами Хассе, изображенными на рис. 6.2,б и в, удовлетворяют закону дополнения, но дополнительный элемент может быть неединственным.

Указание. Элементы c и d оба являются дополнительными элементами для b . Заметим, что при доказательстве единственности дополнительного элемента в упражнении 6.1, по существу, использовался лишь закон дистрибутивности.

Таким образом, задачи 6.12 — 6.15 показывают, что из диаграмм Хассе, приведенных на рис. 6.2, только диаграмма на рис. 6.2,д является представлением булевой алгебры. Она является диаграммой Хассе множества $p(U)$ всех подмножеств множества U , состоящего из двух элементов.

Задача 6.16. Пусть a, b, c, d — элементы булевой алгебры. Покажите, что

- 1) $a \leq b \Leftrightarrow b' \leq a'$;
- 2) $a \leq b \Leftrightarrow a' \cup b = I \Leftrightarrow a \cap b' = O$;
- 3) если $a \leq b$ и $c \leq d$, то $a \cup c \leq b \cup d$, $a \cap c \leq b \cap d$.

Указание. (1) Соотношение 1) следует из (6.5), (6.10) и правила Де Моргана.

(2) $a \leq b \Leftrightarrow a \cup b = b \Rightarrow a' \cup b = a' \cup (a \cup b) = I \cup b = I$, $a' \cup b = I \Rightarrow a = a \cap I = a \cap (a' \cup b) = (a \cap a') \cup (a \cap b) = a \cap b \Rightarrow a \leq b$.

Аналогично получается, что $a \cap b' = 0$.

(3) Из (6.6) и (6.7) имеем $a \leq b \leq b \cup d$, $c \leq d \leq b \cup d \Rightarrow a \cup c \leq b \cup d$.

Доказательство обратных утверждений проводится аналогично.

Упражнение 6.7. Пусть M — булева алгебра и $a, b \in M$, $a \leq b$. Для произвольного элемента $c \in [a, b]$ существует только один элемент $d \in [a, b]$ такой, что $c \cup d = b$, $c \cap d = a$.

Элемент d называют *дополнительным элементом* для c в $[a, b]$. При этом $[a, b]$ — булева алгебра.

Решение. (1) Для произвольных элементов $c, d \in [a, b]$, как следует из (6.6) — (6.8), элементы $c \cup d$ и $c \cap d$ принадлежат $[a, b]$.

(2) Пусть c — заданный элемент из $[a, b]$ и c' — его дополнительный элемент в M . Определим элемент d следующим образом:

$$d = a \cup (c' \cap b) = (a \cup c') \cap (a \cup b) = (a \cup c') \cap b$$

(здесь мы воспользовались тем, что $a \leq b$). Согласно (6.6), $d \in [a, b]$. Далее,

$$c \cup d = c \cup a \cup (c' \cap b) = c \cup (c' \cap b) \quad (\text{поскольку } a \leq b);$$

$$= (c \cup c') \cap (c \cup b) = I \cap (c \cup b) = c \cup b \quad (\text{поскольку } a \leq c);$$

$$= b \quad (\text{поскольку } c \leq b).$$

В силу закона двойственности справедливо также равенство $c \cap d = a$.

(3) Из (1) и (2) следует, что $[a, b]$ — булева алгебра. Отсюда вытекает единственность дополнительного элемента в $[a, b]$.

6.3. Булевы кольца

В множестве $p(U)$ всех подмножеств непустого множества U можно определить кольцевую сумму (симметричную разность) $A \oplus B$ подмножеств A и B множества U , положив

$$A \oplus B = (A - B) \cup (B - A) = (A \cap B^c) \cup (B \cap A^c).$$

Точнее, кольцевая сумма представляет собой множество всех элементов, входящих либо в множество A , либо в множество B , но не в оба множества одновременно.

Задача 6.17. Проверьте, что 1) $(A \oplus B) \oplus C = A \oplus (B \oplus C)$; 2) $A \oplus B = B \oplus A$; 3) $A \oplus \Phi = A$; 4) $A \oplus A = \Phi$; 5) $A \cap (B \oplus C) = (A \cap B) \oplus (A \cap C)$.

Указание. Все приведенные соотношения выводятся непосредственно из определений. Рассмотрим, например, первое из них. Для того чтобы элемент e принадлежал $(A \oplus B) \oplus C$, необходимо, чтобы выполнялись следующие условия:

а) если e входит либо только в A , либо только в B , то он не входит в C ;

б) если e входит в C , то он либо не входит ни в A , ни в B , либо принадлежит пересечению множеств A, B, C .

Другими словами, $e \in (A \oplus B) \oplus C$, если e принадлежит либо только одному из трех множеств A, B, C , либо всем трем одновременно. Аналогично можно проверить, что это же условие необходимо и достаточно для того, чтобы элемент e принадлежал $A \oplus (B \oplus C)$. Отсюда следует утверждение 1).

Условием 5) принадлежности к $A \cap (B \oplus C)$ является принадлежность к A , а кроме того, к одному и только одному из множеств B или C . Таково же условие принадлежности к $(A \cap B) \oplus (A \cap C)$.

Определение булева кольца. Непустое множество M , в котором определены две бинарные операции, «+» (сложение) и «·» (умножение), называется *кольцом* относительно этих операций

сложения и умножения, если выполняются следующие три условия:

(R1) M образует аддитивную группу по сложению. А именно удовлетворяет законам ассоциативности A.1 (см. § 1.5), коммутативности A.2, требованиям существования единичного (нулевого) элемента A.3 и обратного элемента A.4;

(R2) удовлетворяет закону ассоциативности по умножению A.5 (§ 1.5);

(R3) выполняется закон дистрибутивности, а именно для произвольных элементов

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c); \quad (b + c) \cdot a = (b \cdot a) + (c \cdot a). \quad (6.15)$$

Единичный (нулевой) элемент по сложению будем обозначать через 0.

Кольцо M называется *булевым кольцом* (относительно сложения и умножения), если оно, кроме того, удовлетворяет следующим двум условиям:

(BR1) в M существует единичный элемент, обозначаемый далее через 1, такой, что для всех

$$a \cdot 1 = 1 \cdot a = a; \quad (6.16)$$

(BR2) выполняется закон тождественности по умножению, а именно для любого элемента

$$a \cdot a = a. \quad (6.17)$$

Задача 6.18. Множество $p(U)$ образует булево кольцо, если в качестве сложения взять кольцевое сложение \oplus , а в качестве умножения — операцию пересечения множеств \cap .

Указание. Максимальный элемент U является единичным элементом по умножению. Отсюда, из закона тождественности $A \cap A = A$ и результатов задачи 6.17 следует справедливость доказываемого утверждения.

Упражнение 6.8. В булевом кольце M выполняется закон коммутативности по умножению, а кроме того, для всех a :

$$a + a = 0. \quad (6.18)$$

Решение. Так как для произвольных элементов $a, b \in M$ выполняются законы тождественности и дистрибутивности по умножению, а кроме того, коммутативности по сложению, то

$$a + b = (a + b) \cdot (a + b) = ((a + b) \cdot a) + ((a + b) \cdot b) = (a \cdot a) + (b \cdot a) + (a \cdot b) + (b \cdot b) = a + b + (b \cdot a) + (a \cdot b).$$

Пользуясь существованием обратных элементов по сложению, можно сократить элементы a, b в левой и правой частях последнего равенства. При этом оно перейдет в следующее:

$$b \cdot a + (a \cdot b) = 0. \quad (6.19)$$

Полагая в нем $b = 1$ (1 — единичный элемент по умножению), получаем

$$a + a = 0.$$

Это означает, что каждый элемент совпадает со своим обратным по сложению. Из (6.19) и единственности обратного элемента (§ 1.5) следует, что

$$a \cdot b = b \cdot a. \quad (6.20)$$

У п р а ж н е н и е 6.9. Предположим, что множество M является булевой алгеброй относительно операций \cup и \cap . Определим в M сложение и умножение следующим образом. Для произвольных элементов $a, b \in M$ положим:

$$a + b = (a \cap b') \cup (a' \cap b); \quad (6.21)$$

$$a \cdot b = a \cap b. \quad (6.22)$$

Тогда M является булевым кольцом относительно сложения и умножения.

Решение. Пусть a, b, c — произвольные элементы из M . С помощью правила Де Моргана, аксиом булева кольца и соотношения $(a')' = a$ (задача 6.4), получаем

$$\begin{aligned} (a + b)' &= (a \cap b')' \cap (a' \cap b)' = (a' \cup b) \cap (a \cup b'); \\ &= (a' \cap (a \cup b')' \cap (a \cup b'))'; \\ &= (a' \cap a) \cup (a' \cap b') \cup (b \cap a) \cup (b \cap b'); \\ &= (a' \cap b') \cup (a \cap b), \end{aligned} \quad (6.23)$$

(1) Из определения (6.21) и коммутативности операций \cup и \cap следует коммутативность сложения.

(2) С одной стороны, в силу (6.21), (6.23) и аксиом кольца

$$\begin{aligned} (a + b) + c &= ((a + b) \cap c') \cup ((a + b)' \cap c); \\ &= (((a \cap b') \cup (a' \cap b)) \cap c') \cup (((a' \cap b') \cup (a \cap b)) \cap c); \\ &= (a \cap b' \cap c') \cup (a' \cap b \cap c') \cup (a' \cap b' \cap c) \cup (a \cap b \cap c). \end{aligned}$$

С другой стороны, так как $a + (b + c)$ совпадает с $(b + c) + a$, то в последнем выражении элементы a, b, c можно заменить соответственно на b, c, a и новое выражение будет равно старому. Отсюда следует ассоциативность сложения.

(3) Минимальный элемент O является единичным (нулевым) элементом по сложению. В самом деле,

$$a + O = (a \cap O') \cup (a' \cap O) = (a \cap I) \cup O = a.$$

(4) Существование обратного элемента по сложению

$$a + a = (a + a') \cup (a' + a) = O \cup O = O.$$

(5) Законы ассоциативности и тождественности для умножения совпадают с законами ассоциативности и тождественности для операции \cap .

(6) Максимальный элемент I является единичным элементом по умножению. А именно

$$a \cdot I = a \cap I = a = I \cap a = I \cdot a.$$

(7) Закон дистрибутивности. Из (6.22) и коммутативности операции \cap достаточно проверить справедливость лишь одного из соотношений этого закона:

$$\begin{aligned} a \cdot (b + c) &= a \cap ((a \cap b) \cap (a \cap c)') \cup ((a \cap b)' \cap (a \cap c)); \\ (a \cdot b) + (a \cdot c) &= ((a \cap b) \cap (a \cap c')) \cup ((a \cap b)' \cap (a \cap c)); \\ &= (a \cap b) \cap (a' \cup c') \cup ((a' \cup b') \cap (a \cap c)); \\ &= (a \cap b \cap a') \cup (a \cap b \cap c') \cup (a' \cap a \cap c) \cup (b' \cap a \cap c). \end{aligned}$$

Преобразуя эти соотношения с помощью законов дополнения и коммутативности, получим $(a \cap b \cap c') \cup (a \cap b' \cap c)$. Это завершает доказательство закона дистрибутивности.

Так как в булевой алгебре сложение и умножение ассоциативны, то скобки обычно опускаются.

Задача 6.19. Пусть $A, B \in p(U)$. Покажите, что соотношения 1) $A \cup B = A \oplus B \oplus (A \cap B)$ и 2) $A^c = U \oplus A$ следуют непосредственно из определения кольцевой суммы множеств.

В некотором смысле следующее утверждение обратное по отношению к упражнению 6.9.

Упражнение 6.10. В множестве M , которое является булевым кольцом относительно сложения «+» и умножения «·», определим бинарные операции \cup , \cap следующим образом. Для произвольных элементов $a, b \in M$ положим

$$a \cup b = a + b + (a \cdot b); \quad (6.24)$$

$$a \cap b = a \cdot b. \quad (6.25)$$

Тогда M является булевой алгеброй относительно операций \cup и \cap .

Доказательство. (1) Закон ассоциативности. В силу законов дистрибутивности и коммутативности булева кольца имеем

$$a \cup b \cup c = (a + b + (a \cdot b)) + c + (a + b + (a \cdot b)) \cdot c;$$

$$(a \cup b) \cup c = a + b + c + (a \cdot b) + (b \cdot c) + (c \cdot a) + (a \cdot b \cdot c).$$

Аналогично можно показать, что выражение $a \cup (b \cup c)$ также равно правой части последнего равенства. Закон ассоциативности относительно операции \cap является прямым следствием закона ассоциативности по умножению.

(2) Закон коммутативности следует непосредственно из коммутативности сложения и упражнения 6.8.

(3) Закон поглощения. Из законов дистрибутивности, тождественности по умножению булева кольца и упражнения 6.8 следует, что

$$(a \cup b) \cap a = ((a + b) + (a \cdot b)) \cdot a = (a \cdot a) + (b \cdot a) + (a \cdot b \cdot a);$$

$$(a \cup b) \cap a = a + (a \cdot b) + (a \cdot b) = a.$$

Аналогичные операции можно проделать с выражением $(a \cap b) \cup a$. (Проверьте это.)

(4) Закон дистрибутивности. В силу упражнения 6.3 достаточно проверить выполнение лишь одного из соотношений этого закона. Имеем

$$a \cap (b \cup c) = a \cdot (b + c + (b \cdot c)) = (a \cdot b) + (a \cdot c) + (a \cdot b \cdot c) = (a \cdot b) + (a \cdot c) + (a \cdot b \cdot a \cdot c) = (a \cap b) \cup (a \cap c).$$

(5) Закон дополнения. Для единичных элементов 1, 0 соответственно по умножению и сложению выполняются следующие соотношения:

$$(a) \ a \cap 1 = a \cdot 1 = a, \ a \cap 0 = a \cdot 0 = 0 \quad (\text{см. формулу (1.30)});$$

$$(б) \ a \cap (1 + a) = a \cdot (1 + a) = a + (a \cdot a) = a + a = 0;$$

$$a \cup (1 + a) = a + 1 + a + (a \cdot (1 + a)) = 1 + a + a + 0 = 1.$$

Таким образом, считая 1 максимальным элементом, а 0 — минимальным, мы приходим к закону дополнения (см. замечания к закону дополнения в § 6.1). При этом элемент a' , дополнительный к a , определяется равенством

$$a' = 1 + a. \quad (6.26)$$

6.4. Представления булевых алгебр [10, 13]

В § 6.1 — 6.3 было показано, что если совокупность всех подмножеств $p(U)$ непустого множества U обладает некоторыми свойствами, то теми же свойствами обладает и произвольная

булева алгебра. Далее покажем, что в некотором смысле верно и обратное утверждение.

Предположим, что в множестве M определены бинарные $\alpha_1, \alpha_2, \dots, \alpha_m$ и унарные $\beta_1, \beta_2, \dots, \beta_l$ операции, а в множестве M' определены бинарные $\alpha'_1, \alpha'_2, \dots, \alpha'_m$ и унарные операции $\beta'_1, \beta'_2, \dots, \beta'_l$. Если существует взаимно однозначное отображение ρ множества M на M' такое, что для произвольных элементов $a, b \in M$ выполняются соотношения

$$\rho(a \alpha_i b) = \rho(a) \alpha'_i \rho(b), \quad 1 \leq i \leq m;$$

$$\rho(\beta_i a) = \beta'_i \rho(a), \quad 1 \leq i \leq l,$$

то алгебраические системы $A = (M, \alpha_1, \alpha_2, \dots, \alpha_m, \beta_1, \beta_2, \dots, \beta_l)$ и $A' = (M', \alpha'_1, \alpha'_2, \dots, \alpha'_m, \beta'_1, \beta'_2, \dots, \beta'_l)$ называются *изоморфными*, это обозначается $A \approx A'$. При этом ρ называется *изоморфизмом*. Предположим, что имеются две изоморфные алгебраические системы и, например, что бинарная операция α_i удовлетворяет закону коммутативности. Пусть ρ' — обратное к ρ отображение и a', b' — произвольные элементы из M' . Положим $a = \rho'(a')$, $b = \rho'(b')$. Тогда $a' = \rho(a)$, $b' = \rho(b)$, $a' \alpha'_i b' = \rho(a) \alpha'_i \rho(b) = \rho(a \alpha_i b) = \rho(b \alpha_i a) = \rho(b) \alpha'_i \rho(a) = b' \alpha'_i a'$, т. е. для α'_i также выполняется закон коммутативности.

Задача 6.20. Проверьте выполнение соответствующих утверждений для законов ассоциативности, поглощения и дистрибутивности.

Задача 6.21. Пусть $(M_1, \cup, \cap) \approx (M_2, \cup, \cap)$ и ρ_1 — изоморфное отображение M_1 на M_2 . Покажите, что если M_1 удовлетворяет закону дополнения, то при отображении ρ_1 максимальный (минимальный) элемент переходит соответственно в максимальный (минимальный) элемент и для произвольного $a_1 \in M_1$ элемент $\rho_1(a'_1)$ является дополнительным к элементу $\rho_1(a_1)$ (где a'_1 является дополнительным элементом к a_1).

Указание. Все эти утверждения можно получить непосредственно из определений.

Как следует из сказанного, соотношение $(M_1, \cup, \cap) \approx (M_2, \cup, \cap)$ означает, что при изоморфном отображении решетка (булева алгебра) переходит в решетку (булеву алгебру). Таким образом, при наличии изоморфизма элементы некоторой алгебраической системы всегда можно заменить их образами при изоморфном отображении (последние можно рассматривать как представления элементов исходной алгебраической системы). С точностью до обозначения операций и элементов изоморфные алгебраические системы полностью идентичны.

Задача 6.22. Покажите, что отношение изоморфизма « \approx » удовлетворяет законам рефлексивности, симметричности и транзитивности (см. § 1.4), а следовательно, является отношением эквивалентности.

Задача 6.23. Если два конечных множества U и U' состоят из одного и того же числа элементов, то любое взаимно однознач-

ное отображение между ними является изоморфизмом относительно операций объединения, пересечения и дополнения множеств. Следовательно, булевы алгебры $p(U)$ и $p(U')$ изоморфны.

Справедливость всех утверждений непосредственно следует из определений.

Если существует взаимно однозначное отображение ρ упорядоченного множества M на упорядоченное множество M' (отношения порядка в них обозначаются через \leq и \leq' соответственно) и для произвольных элементов $a, b \in M$ выполняется соотношение $a \leq b \Leftrightarrow \rho(a) \leq' \rho(b)$, то упорядоченные множества M и M' называются *изоморфными*.

Задача 6.24. Предположим, что в множествах M, M' определены соответственно отношения R, R' и существует взаимно однозначное отображение ρ множества M на M' такое, что для произвольных элементов $a, b \in M$ имеет место соотношение $aRb \Leftrightarrow \rho(a)R'\rho(b)$. Если R является отношением порядка, то:

1) R' также является отношением порядка (отношения R, R' обозначаются далее знаком \leq);

2) если для элементов a, b из M существует $\sup\{a, b\}$ ($\inf\{a, b\}$), то $\rho(\sup\{a, b\}) = \sup\{\rho(a), \rho(b)\}$, $(\rho(\inf\{a, b\}) = \inf\{\rho(a), \rho(b)\})$;

3) если множество M является решеткой (булевой алгеброй) относительно M' , то это же верно и для множества M' .

Решение следует непосредственно из определений.

Задача 6.25. (Обобщение правила Де Моргана). Пусть $(M, \cup, \cap, ')$ — булева алгебра. Для произвольного элемента $a \in M$ положим $\rho(a) = a'$. При этом ρ является изоморфизмом $(M, \cup, \cap, ')$ на $(M, \cap, \cup, ')$. Следовательно, если в булевой алгебре верно некоторое утверждение, то в ней верно также d — двойственное утверждение*, которое получается при замене каждого элемента дополнительным, отношения \leq противоположным отношением \geq , и \cup на \cap .

Решение. Из правила Де Моргана и равенства $\rho(a') = a = (\rho(a))'$ (задача 6.4) следует справедливость первой половины доказываемого утверждения. Вторая половина утверждения следует из первой и соотношения $a \leq b \Leftrightarrow \rho(b) \leq \rho(a)$ (задача 6.16).

В этом разделе для простоты ограничимся изучением булевых алгебр, состоящих из конечного числа элементов, т. е. *конечных булевых алгебр*. Пусть M — конечная булева алгебра с отношением порядка \leq , задаваемым соотношением (6.5). Предельный минимальный элемент множества M — $\{O\}$, которое получается из M удалением минимального элемента O , называется *предельным минимальным элементом* конечной алгебры M . Если $M \neq \{O\}$, то, согласно упражнению 6.4, существует, по крайней мере, один предельный минимальный элемент. Обозначим через U_m

* Например, d -двойственным утверждением для $a \cap b \leq a$ будет $a' \leq a' \cup b'$. Двойственным утверждением в данном случае является $a \leq a \cup b$.

совокупность всех предельных минимальных элементов булевой алгебры M . Непосредственно из определения следует, что необходимым и достаточным условием выполнения включения $a \in U_m$ является следующее:

$$\text{если } b < a, \text{ то } b = 0, \text{ т. е. } 0 | < a. \quad (6.27)$$

Упражнение 6.11. Если a_1 и a_2 — различные предельные минимальные элементы, то $a_1 \cap a_2 = 0$.

Решение. В силу (6.6) $a_1 \cap a_2 < a_1$. Если $a_1 \cap a_2 = a_1$, то из (6.10) имеем $a_1 < a_2$. Поскольку $a_1 \neq a_2$, то $a_1 < a_2$, но это противоречит тому, что a_1, a_2 — предельные минимальные элементы. Следовательно, $a_1 \cap a_2 < a_1$. Поскольку a_1 является предельным минимальным элементом, то $a_1 \cap a_2 = 0$.

Для $A = (a_1, a_2, \dots, a_i) \in U_m$ определим $u(A)$ следующим образом: $u(A) = a_1 \cup a_2 \cup \dots \cup a_i$; условимся, что $u(\emptyset) = 0$. Из определения следует, что

$$u(A) = 0 \Leftrightarrow A = \emptyset. \quad (6.28)$$

Для произвольных множеств $A, B \in p(U_m)$ имеем:

$$\begin{aligned} u(A) \cup u(B) &= u(A - (A \cap B)) \cup u(A \cap B) \cup u(B - (A \cap B)) \cup u(A \cap B) = \\ &= u(A - (A \cap B)) \cup u(A \cap B) \cup u(B - (A \cap B)) = \\ &= u(A \cup B). \end{aligned} \quad (6.29)$$

Если $A \subset B$, то из равенства $u(A) \cup u(B) = u(B)$ и из (6.5) получаем

$$u(A) \leq u(B). \quad (6.30)$$

Если $A = \{a_1, a_2, \dots, a_i\}$, $B = \{b_1, b_2, \dots, b_j\}$ и $A \cap B = \emptyset$, то, применяя последовательно дистрибутивный закон к полученному соотношению, а кроме того, пользуясь упражнением 6.11, легко проверить, что

$$\begin{aligned} u(A) \cap u(B) &= (a_1 \cup a_2 \cup \dots \cup a_i) \cap (b_1 \cup b_2 \cup \dots \cup b_j) = \\ &= (a_1 \cap b_1) \cup (a_1 \cap b_2) \cup \dots \cup (a_i \cap b_j) = 0. \end{aligned} \quad (6.31)$$

Из (6.31) следует, что для произвольных множеств $A, B \in p(U_m)$

$$\begin{aligned} u(A) \cap u(B) &= u(A) \cap (u(B - (A \cap B)) \cup u(A \cap B)) = \\ &= (u(A) \cap u(B - (A \cap B))) \cup (u(A) \cap u(A \cap B)) = u(A) \cap u(A \cap B). \end{aligned} \quad (6.32)$$

Из (6.30) и (6.10)

$$u(A) \cap u(B) = u(A \cap B). \quad (6.33)$$

Предположим, что $A \neq B$. Без ограничения общности можно считать, что $a \in A - B$, $a \in U_m$. В силу (6.33)

$$u(\{a\}) \cap u(A) = u(\{a\}) = a \neq 0;$$

$$u(\{a\}) \cap u(B) = u(\emptyset) = 0.$$

Отсюда следует, что

$$A \neq B \Leftrightarrow u(A) \neq u(B), \quad (6.34)$$

т. е. u является вложением $p(U_m)$ в M^* . Покажем, что u является взаимно однозначным отображением.

Для произвольного элемента $a \in M$ обозначим через $\varphi(a)$ совокупность всех элементов из U_m , которые меньше a . А именно

$$\varphi(a) = [0, a] \cap U_m. \quad (6.35)$$

Если $a \neq 0$, то $[0, a] - \{0\}$ непусто, и, согласно упражнению 6.4, в множестве $[0, a] - \{0\}$ существует предельный минимальный элемент a_1 . Поскольку $0 \prec a_1 \leq a$, то из (6.27) и (6.35) имеем $a_1 \in \varphi(a)$. По определению, $\varphi(0) = \emptyset$, а следовательно,

$$\varphi(a) = \emptyset \Leftrightarrow a = 0. \quad (6.36)$$

Если, кроме того, $a \leq b$, то

$$\varphi(a) \subset \varphi(b). \quad (6.37)$$

Положим $\varphi(a) = \{a_1, a_2, \dots, a_n\}$. Поскольку $a_i \leq a$ ($1 \leq i \leq n$), то из (6.7) следует, что $a_1 \cup a_2 \leq a$ при $n \geq 2$ и $(a_1 \cup a_2) \cup a_3 \leq a$ при $n \geq 3$. Аналогично можно показать, что

$$u(\varphi(a)) = a_1 \cup a_2 \cup \dots \cup a_n \leq a. \quad (6.38)$$

Упражнение 6.12. Для произвольного элемента $a \in M$

$$u(\varphi(a)) = a. \quad (6.39)$$

Решение. Имеем $u(\varphi(0)) = u(\emptyset) = 0$. Пусть $a \neq 0$ и $c = u(\varphi(a))$. Согласно (6.38), $c \in [0, a]$. Пусть d — дополнительный элемент в $[0, a]$ к элементу c . Тогда $c \cup d = a$; $c \cap d = 0$. Из (6.37) и (6.30) получаем, что $u(\varphi(d)) \leq u(\varphi(a)) = c$. В силу (6.38) $u(\varphi(d)) \leq d$. Из (6.8) вытекает, что $u(\varphi(d)) \leq c \cap d = 0$. Следовательно, $u(\varphi(d)) = 0$. Из (6.28) и (6.36) имеем $d = 0$, т. е. $c = a$.

Согласно (6.39), отображение u является сюръекцией, а следовательно, u задает взаимно однозначное соответствие между $p(U_m)$ и M (u и φ — взаимно обратные отображения). Кроме того, по определению,

$$\varphi(I) = U_m, \quad (6.40)$$

так что

$$= u(U_m). \quad (6.41)$$

Для произвольного $A \in p(U_m)$ из (6.29), (6.31) и (6.41) следует, что

$$u(A) \cup u(A^c) = u(U_m) = I;$$

$$u(A) \cap u(A^c) = 0,$$

а поэтому

$$u(A)' = u(A^c). \quad (6.42)$$

* Отображение p множества A в B называется *вложением*, если $p(a_1) \neq p(a_2)$ для произвольных элементов $a_1, a_2 \in A$, $a_1 \neq a_2$. Отображение p называется *сюръекцией*, если для произвольного $b \in B$ найдется элемент $a \in A$ такой, что $p(a) = b$.

Из (6.29), (6.33) и (6.42) следует, что u является изоморфизмом*. Суммируя изложенное, получаем следующую теорему (вместо обозначений U_m и u далее в формулировке используются соответственно U и ρ).

Теорема 6.1. Пусть M — произвольная конечная булева алгебра. Алгебра M изоморфна множеству $\rho(U)$ всех подмножеств некоторого множества U . А именно существует взаимно однозначное отображение ρ множества $\rho(U)$. На M такое, что для произвольных элементов $A, B \in \rho(U)$

$$\rho(A \cup B) = \rho(A) \cup \rho(B);$$

$$\rho(A \cap B) = \rho(A) \cap \rho(B);$$

$$\rho(A^c) = \rho(A)';$$

$$A \subset B \Leftrightarrow \rho(A) \leq \rho(B).$$

Из этой теоремы и задачи 6.23 получаем следующее утверждение.

Следствие из теоремы 6.1. Две конечные булевы алгебры, имеющие одинаковое число элементов, изоморфны. Число элементов булевой алгебры является некоторой степенью числа 2.

Доказательство. Вторая половина утверждения является следствием равенства $|\rho(U)| = 2^{|U|}$.

Таким образом, при изучении конечных булевых алгебр вместо вывода тех или иных суждений на основании системы аксиом можно пользоваться множеством $\rho(U)$, допускающим более конкретную интерпретацию, и рассматривать утверждения относительно операций « \cup », « \cap » и « c », взяв в качестве последних операций объединения, пересечения и дополнения множеств. Как показывает пример вывода правила де Моргана, такой подход часто оказывается более удобным.

В заключение заметим, что теорема 6.1 может быть обобщена также на бесконечные булевы алгебры [10].

Задачи [2,10]

6.1. Для элементов a, b, c решетки M справедливы следующие соотношения: $a \cup (b \cap c) \leq (a \cup b) \cap (a \cup c)$; $a \cap (b \cup c) \geq (a \cap b) \cup (a \cap c)$.

6.2. Для дистрибутивных решеток выполняется следующий закон модулярности: $a \leq c \Rightarrow a \cup (b \cap c) = (a \cup b) \cap c$.

6.3. Для дистрибутивных решеток выполняется следующее соотношение: $(a \cup b) \cap (b \cup c) \cap (c \cup a) = (a \cap b) \cup (b \cap c) \cup (c \cap a)$.

6.4. Найдите диаграмму Хассе булевой алгебры, число элементов которой равно 2^m .

* Справедливость соотношения (6.41) и (6.42) следует также из задачи (6.21).

Глава 7

Булевы функции

7.1. Определение [12,13]

Пусть V_1 — множество, состоящее из двух элементов, 0 и 1; $V_1 = \{0, 1\}$. Здесь символы 0 и 1 обозначают два абстрактных элемента (величины, объекта). В зависимости от конкретной ситуации это могут быть, например, целые числа 0 и 1, обозначения ложности или истинности высказывания, положительное и отрицательное напряжение на выходе канала связи, замкнутые и разомкнутые контакты выключателя и т. д. Часто оказывается удобным считать V_1 либо булевой алгеброй, состоящей из двух элементов (§ 6.1), либо конечным полем Z_2 (§ 1.5). Пусть n — целое положительное число и V_n — n -кратное прямое произведение множества V_1 на себя. Другими словами, V_n — это совокупность всех последовательностей из n компонент, каждая из которых может быть либо символом 0, либо символом 1:

$$V_n = \{(a_1, a_2, \dots, a_n) | a_i \in V_1, 1 \leq i \leq n\}.$$

Число элементов $|V_n|$ множества V_n равно 2^n . В тех случаях, когда это не приводит к путанице, вместо записи (1, 1, 0, 0) будем пользоваться более простой записью (1100). С помощью элементов V_n можно обозначать, например, совокупность физических состояний, в которых могут находиться n входов (выходов) цифрового канала связи, содержимое n -разрядного двоичного регистра и т. д.

Остановимся на методах конкретного представления элементов множества V_n .

(1) Представление с помощью целых чисел. Если символы 0 и 1 компонент последовательности $A = (a_1, a_2, \dots, a_n)$ из V_n рассматривать как целые числа 0 и 1, то последовательности A можно сопоставить целое число

$$J(A) = \sum_{j=1}^n a_j 2^{n-j}.$$

Это соответствие между V_n и множеством I_n , состоящим из целых чисел от 0 до $2^n - 1$ включительно, является взаимно однозначным. В самом деле, целое число $i \in I_n$ можно представить в двоичном виде с помощью n разрядов. Если символы 0 и 1 такого двоичного представления записывать в виде последовательности (a_1, a_2, \dots, a_n) , начиная со старших разрядов, то эту последовательность можно рассматривать как элемент множества V_n . Для упрощения обозначений вместо последовательности (например, 1, 1, 0, 1) часто пишут соответствующее ей целое число («13»).

(2) Геометрическое описание. Рассмотрим n -мерный единичный куб. Последовательность (a_1, a_2, \dots, a_n) можно рассматривать как представление вершины этого куба, i -я координата которой равна a_i (символы 0 и 1 при этом следует рассматривать как действительные числа). На рис. 7.1 приведено представление множества V_3 в виде вершин единичного куба. В этом смысле последовательности (a_1, a_2, \dots, a_n) называются *двоичными векторами*. Однако они не являются векторами в полном обычном смысле слова, поскольку над ними нельзя выполнять стандартные операции.

Расстоянием Хэмминга между двумя элементами

$A = (a_1, a_2, \dots, a_n)$ и $B = (b_1, b_2, \dots, b_n)$

из V_n называется число пар (a_i, b_i) таких, что $a_i \neq b_i$ (см. § 2.2). Это расстояние обозначается через $d(A, B)$. В частности, если $d(A, B) = 1$, то элементы A и B называются *соседними*. При геометри-

ческом описании вершины, соответствующие соседним элементам, соединяются ребром. Расстояние Хэмминга $d(A, B)$ обладает теми же свойствами, что и евклидово расстояние, и если «идти» из одной вершины в другую вдоль ребер, то оно будет равно минимальному числу ребер в таком пути. Так, в кубе, показанном на рис. 7.1, имеется два кратчайших пути из вершины $(1, 0, 0)$ в вершину $(1, 1, 1)$: либо через вершину $(1, 0, 1)$, либо через $(1, 1, 0)$. Однако в обоих случаях число пройденных ребер, т. е. расстояние Хэмминга, будет равно 2. Число равных 1 компонент вектора $A = (a_1, a_2, \dots, a_n) \in V_n$ называется его *весом* и обозначается через $w(A)$. Например, вес вектора $(1, 0, 1, 1)$ равен 3. Если 0 (соответственно 1) — это элемент (вектор) из V_n , все компоненты которого равны 0 (соответственно 1), то вес элемента A из V_n будет равен $d(A, 0) = n - d(A, 1)$.

Определения высказывания и булевой функции. Отображение q непустого множества M в V_1 называется *высказыванием* (на M). Если q тождественно равно 1 на M , то q называется *истинным высказыванием*; если q тождественно равно 0 на M , то q — *ложное высказывание*.

Отображение q множества V_n в V_1 называется *булевой функцией* от n переменных (или функцией алгебры логики). Множество всех булевых функций от n переменных обозначается через \mathcal{B}_n ; множество всех булевых функций обозначим просто через \mathcal{B} . Переменные, принимающие значения в V_1 , будем называть булевыми (логическими или двоичными) переменными и обозначать: через $x, x_1, x_2, \dots; y, y_1, y_2, \dots; z, z_1, z_2, \dots$. Группы переменных, такие, как, например, $X = (x_1, x_2, \dots, x_n)$ будем обозначать через $X, X_1, X_2, \dots; Y, Y_1, Y_2, \dots; Z, Z_1, Z_2, \dots$.

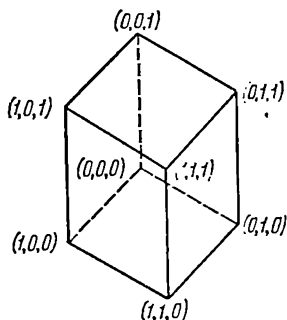


Рис. 7.1. Представление V_3 с помощью единичного куба

Пусть $f \in \mathcal{B}_n$; положим

$$D_0(f) = \{A \mid f(A) = 0, A \in V_n\}; \quad (7.1)$$

$$D_1(f) = \{A \mid f(A) = 1, A \in V_n\}. \quad (7.2)$$

Так как $D_0(f) = V_n - D_1(f)$, то по каждому из этих множеств второе определяется автоматически. Сопоставляя функции f множество $D_1(f)$, получаем отображение \mathcal{B}_n в множество $\rho(V_n)$ всех подмножеств V_n .

У п р а ж н е н и е 7.1. Отображение D_1 является взаимно однозначным, и $|\mathcal{B}_n| = 2^{2^n}$.

Решение. Если $f_1 \neq f_2$, то $D_1(f_1) \neq D_1(f_2)$. Для произвольного подмножества W множества V_n существует функция $f \in \mathcal{B}_n$ такая, что $D_1(f) = W$. Поэтому D_1 задает взаимно однозначное соответствие между \mathcal{B}_n и $\rho(V_n)$, так что $|\mathcal{B}_n| = |\rho(V_n)| = 2^{|V_n|} = 2^{2^n}$.

Обозначения постоянных функций (констант). Функции из \mathcal{B}_n , принимающие одно и то же значение (0 или 1) во всей области определения V_n , можно однозначно идентифицировать с помощью элементов 0 или 1 из V_1 . Поэтому далее эти функции обозначаются просто символами 0 и 1. При этом $D_1(0) = \Phi$, $D_1(1) = V_n$.

Не всюду определенные булевы функции. Отображение непустого подмножества W множества V_n (W может совпадать с V_n) в V_1 называется *частичной булевой функцией от n переменных*; множества W и $V_n - W$ называются соответственно *областью определения* и *областью неопределенности* этой функции. Множество всех частичных булевых функций будем обозначать через \mathcal{B}'_n . Ясно, что $\mathcal{B}_n \subset \mathcal{B}'_n$, $\mathcal{B} \subset \mathcal{B}'$.

Для функции $f \in \mathcal{B}'_n$ также можно определить множества $D_0(f)$ и $D_1(f)$ с помощью соотношений (7.1) и (7.2). Область неопределенности функции f обозначается далее также через $D_{ud}(f)$. Для перечисления элементов множеств $D_0(f)$, $D_1(f)$, $D_{ud}(f)$ часто вместо записи последовательностей из V_n используются их представления с помощью целых чисел.

З а д а ч а 7.1. $|\mathcal{B}'_n| = 3^{2^n} - 1$.

Решение. Имеются три возможности задать частичную булеву функцию на элементе A из V_n : выбрать ее значение равным 0, равным 1 или не определять. Различным таким выборам соответствуют различные функции и наоборот. Следует, однако, исключить случай, когда функция не определена на никаком элементе. Таким образом, общее число возможных способов задания частичной логической функции от n переменных равно $3^{2^n} - 1$.

Булевы функции (в том числе частичные) являются математическими объектами, которые описывают преобразования входных сигналов в выходные, реализуемые схемами из функциональных элементов, описанными в § 8.1.

Отношение порядка. Введем между элементами 0 и 1 из V_1 отношение порядка, полагая

$$0 < 1 \quad (7.3)$$

(т. е. рассмотрим булеву алгебру, состоящую из двух элементов множества V_1 (см. задачу 7.7)). С помощью этого отношения можно ввести отношение порядка между элементами $A=(a_1, a_2, \dots, a_n)$ и $B=(b_1, b_2, \dots, b_n)$ из V_n , полагая $A \leq B$, если $a_i \leq b_i$, $1 \leq i \leq n$. Если при этом $A \neq B$, то будем использовать запись $A < B$. Например, $(0, 0, 1) < (0, 1, 1) < (1, 1, 1)$. Если $A < B$ и соотношение $A < X < B$ не выполняется ни для одного элемента $X \in V_n$, то будем писать $A \lceil B$. Введенное отношение является рефлексивным, антисимметричным и транзитивным (§ 1.1 и 6.2) и, следовательно, является отношением порядка.

Задача 7.2 $(a_1, a_2, \dots, a_n) \lceil (b_1, b_2, \dots, b_n) \Leftrightarrow$ (существует целое число i ($1 \leq i \leq n$) такое, что $a_i = 0$, $b_i = 1$, $a_j = b_j$ ($j \neq i$)).

Решение. Утверждение непосредственно следует из определений.

Задачи и упражнения, связанные с материалом гл. 6, помечены здесь знаком '.

Задача 7.3'. Множество V_n является упорядоченным с отношением порядка « \leq », и как упорядоченное множество изоморфно множеству $p(U)$ всех подмножеств множества U из n элементов. Следовательно, V_n является булевой алгеброй, изоморфной $p(U)$.

Решение. Положим $U = \{1, 2, \dots, n\}$. Определим следующим образом отображение ρ множества V_n на $p(U)$. Для произвольного элемента $A = (a_1, a_2, \dots, a_n)$ положим

$$\rho(A) = \{i \mid a_i = 1, 1 \leq i \leq n\}.$$

Отображение ρ является взаимно однозначным. Далее, для произвольных $A \leq B$, $\rho(A) \subset \rho(B)$. Следовательно, отношение \leq также является отношением порядка, и $V_n \approx p(U)$. Согласно решению задачи 6.24, V_n является булевой алгеброй.

Задача 7.4. Представим V_n в виде n -мерного единичного куба и будем рассматривать вершины и ребра куба как вершины и ребра графа. Тогда этот граф будет представлять собой диаграмму Хассе упорядоченного множества (в которой мы пренебрегаем ориентацией ребер).

Отношение порядка между функциями. Пусть $f, g \in \mathcal{B}_n$. Если для всех элементов $(a_1, a_2, \dots, a_n) \in V_n$, таких, что $f(a_1, a_2, \dots, a_n) = 1$, выполняется также равенство $g(a_1, a_2, \dots, a_n) = 1$, то будем считать, что $f \leq g$. Если, кроме того, $f \neq g$, то используется также запись $f < g$. Запись $f \not\leq g$ означает, что соотношение $f \leq g$ не выполняется. Непосредственно из определений следует, что

$$f \leq g \Leftrightarrow f(A) \leq g(A) \text{ для всех } A; \quad (7.4)$$

$$f \leq g \Leftrightarrow D_1(f) \subseteq D_1(g). \quad (7.5)$$

Упражнение 7.2. Отношение \leq в множестве \mathcal{B}_n является отношением порядка; \mathcal{B}_n и $p(V_n)$ как упорядоченные множества изоморфны. Таким образом, \mathcal{B}_n с введенным отношением порядка образует булеву алгебру, изоморфную $p(V_n)$. Константы 0 и

1 являются соответственно минимальным и максимальным элементами \mathcal{B}_n .

Решение. Из упражнения 7.1 следует, что D_1 — взаимно однозначное отображение \mathcal{B}_n на $p(V_n)$ и отношение включения \subset в $p(U)$ — отношение порядка. Отсюда, из соотношения (7.5) и задачи 6.24 следует, что \mathcal{B}_n образуют булеву алгебру.

Задача 7.5. Если $f \leq g$ или $g \leq f$, то f и g называются *сравнимыми*. Пусть $f \in \mathcal{B}_n$ и $|D_1(f)| = l$. Тогда общее число функций в \mathcal{B}_n , сравнимых с f , равно $2^l + 2^{2^n - l - 1}$.

Решение. Общее число функций в \mathcal{B}_n таких, что $f \leq g$, согласно упражнению 7.2, равно числу подмножеств множества V_n , содержащих $D_1(f)$. Это число, в свою очередь, равно числу подмножеств в $V_n - D_1(f)$, т. е. $2^{2^n - l}$. Аналогично число функций в \mathcal{B}_n таких, что $g \leq f$, равно $2^{|D_1(f)|} = 2^l$. Наконец, заметим, что если $g < f < g$, то в силу закона антисимметричности $g = f$.

Определение монотонной функции. Пусть $f(x_1, x_2, \dots, x_n) \in \mathcal{B}_n$, i — целое число ($1 \leq i \leq n$) и a — элемент из V_1 . Функцию $f(x_1, x_2, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$ от $n-1$ переменных будем обозначать через $f_{x_i=a}$. Если $1 \leq i_1 < i_2 < \dots < i_r \leq n$, $a_j \in V_1$ ($1 \leq j \leq r$) и в функцию $f(x_1, x_2, \dots, x_n) \in \mathcal{B}_n$ подставить $x_{i_1} = a_1, x_{i_2} = a_2, \dots, x_{i_r} = a_r$, то получится функция от $n-r$ переменных, которую мы будем обозначать через $f_{x_{i_1}=a_1, x_{i_2}=a_2, \dots, x_{i_r}=a_r}$.

Если

$$f_{x_i=0} \leq f_{x_i=1}, \quad (7.6)$$

то переменную x_i с номером i функции f называем *положительной*; если же $f_{x_i=1} \leq f_{x_i=0}$, то эту переменную называют *отрицательной*. Когда переменная одновременно является и положительной и отрицательной, то $f_{x_i=0} = f_{x_i=1}$ и значение функции f не зависит от x_i . Функция, все переменные которой являются положительными, называется *положительной* или *монотонной функцией*. Если все переменные функции отрицательные, функция называется *отрицательной*. Если каждая переменная некоторой функции является либо положительной, либо отрицательной, то такая функция называется *смешанной монотонной функцией*.

Задача 7.6. Функция $f \in \mathcal{B}_n$ является монотонной тогда и только тогда, когда $f(A) \leq f(B)$ для произвольных элементов A, B из V_n таких, что $A \leq B$.

Решение. Необходимость. Существуют такие элементы A_1, A_2, \dots, A_m , что $A_1 \leq A_2 \leq \dots \leq A_m \leq B$ (упражнение 6.4)*. Из задачи 7.2 и определения 7.6 получаем

$$f(A) \leq f(A_1) \leq \dots \leq f(A_m) \leq f(B).$$

Достаточность очевидна из определений.

* Это можно доказать непосредственно с помощью определений, не пользуясь упражнением 6.4.

7.2. Операции [12,13]

Введем в V_1 операции, полагая для произвольных элементов a, b из V_1 :

$$1) a \vee b = \max \{a, b\}; \quad (7.7)$$

$$2) a \cdot b = \min \{a, b\}; \quad (7.8)$$

$$3) 0 \oplus 0 = 0, \quad 1 \oplus 0 = 0 \oplus 1 = 1, \quad 1 \oplus 1 = 0; \quad (7.9)$$

$$4) \bar{0} = 1, \quad \bar{1} = 0, \quad (7.10)$$

где $\max \{a, b\}$ и $\min \{a, b\}$ — соответственно максимальный и минимальный в смысле отношения (7.3) элементы в паре a, b .

Элемент $a \vee b$ называется *логической* или *булевой суммой* элементов a и b , элемент $a \cdot b$ — *логическим* или *булевым произведением* (или просто произведением)*, $a \oplus b$ — *исключающим* или *или суммой по модулю 2*, \bar{a} — отрицанием a .

Задача 7.7. Четверка $(V_1, \vee, \cdot, ')$ изоморфна булевой алгебре $(\{0, 1\}, \cup, \cap, ')$, состоящей из двух элементов. Для произвольных элементов s, t, u, v из V_1 справедливы две группы соотношений, приведенные соответственно в табл. 7.1 и 7.2, а также закон двойственности.

Указание. Первая часть задачи решается с помощью определений (7.7), (7.8) и (7.10). Вторая часть следует из аксиомы и свойств булевой алгебры, а также определения (7.3).

Задача 7.8. Система (V_1, \oplus, \cdot) , где \oplus — сложение, а « \cdot » — умножение, изоморфна полю Z_2 (см. § 1.5). Для произвольных элементов s, t, u из V_1 справедлива группа соотношений III (табл. 7.3).

Таблица 7.1

Формула	Соотношение I
(1.1) Закон тождественности	$s \vee s = s, \quad s \cdot s = s$
(1.2) Закон ассоциативности	$(s \vee t) \vee u = s \vee (t \vee u),$ $(s \cdot t) \cdot u = s \cdot (t \cdot u)$
(1.3) Закон коммутативности	$s \vee t = t \vee s, \quad s \cdot t = t \cdot s$
(1.4) Закон поглощения	$(s \vee t) \cdot s = s, \quad (s \cdot t) \vee s = s$
(1.5) Закон дистрибутивности	$s \cdot (t \vee u) = (s \cdot t) \vee (s \cdot u),$ $s \vee (t \cdot u) = (s \vee t) \cdot (s \vee u)$
(1.6)	$s \vee 1 = 1, \quad s \cdot 1 = s$
(1.7) Закон дополнения	$s \vee 0 = s, \quad s \cdot 0 = 0$
(1.8)	$s \vee \bar{s} = 1, \quad s \cdot \bar{s} = 0$
(2.1) Правило де Моргана	$s \vee \bar{t} = \bar{s} \cdot \bar{t}, \quad s \cdot \bar{t} = \bar{s} \vee \bar{t}$
(2.2)	$(s) = \bar{s}$
(2.3)	Если $s \vee t = s \vee u$ и $s \cdot t = s \cdot u$, то $t = u$

* Часто $a \vee b$ называют дизъюнкцией, а $a \cdot b$ — конъюнкцией элементов a и b . (Прим. ред.)

Таблица 7.2

Формула	Соотношение II
(3.1)	$0 \leq s \leq 1$
(3.2)	$s \leq s \vee t, s \cdot t \leq s$
(3.3)	$s \leq t \Leftrightarrow \bar{s} \geq \bar{t}$
(3.4)	$s \leq t \Leftrightarrow s \vee t = t \Leftrightarrow s \cdot t = s$
(3.5)	$s \leq t \Leftrightarrow \bar{s} \vee t = 1 \Leftrightarrow s \cdot \bar{t} = 0$
(3.6)	Если $s \leq t$ и $u \leq v$, то $s \vee u \leq t \vee v$ и $s \cdot u \leq t \cdot v$

Указание. Согласно определениям (7.8) и (7.9), операции \oplus и \cdot являются сложением и умножением по модулю 2. Отсюда следует первое утверждение задачи. Так как в \mathbb{Z}_2 выполняется закон тождественности, то рассматриваемая тройка является булевой алгеброй. Из задачи 7.7 и § 6.3 следует группа соотношений III.

Задача 7.9. Пользуясь непосредственно определениями (7.3), (7.7)—(7.10), покажите, что для произвольных элементов s, t, u, v из V_1 выполняются все три группы соотношений, I, II и III.

Указание. Ассоциативность, коммутативность, закон поглощения и дистрибутивность вытекают соответственно из следующих равенств:

$$\begin{aligned} \max \{ \max \{ s, t \}, u \} &= \max \{ s, t, u \} = \max \{ s, \max \{ t, u \} \}; \quad \max \{ s, t \} = \\ &= \max \{ t, s \}; \quad \min \{ \max \{ s, t \}, s \} = s, \quad \min \{ s, \max \{ t, u \} \} = \\ &= \max \{ \min \{ s, t \}, \min \{ s, u \} \}. \end{aligned}$$

Аналогично можно показать справедливость закона дополнения. При доказательстве первого соотношения правила Де Моргана без ограничения общности можно предположить, что $s \leq t$. При этом из (7.3) и (7.10) будет следовать, что $\bar{t} \leq \bar{s}$, т. е. обе части доказываемого соотношения равны \bar{t} . Для доказательства утверждения (2.3) следует отдельно рассмотреть случаи $s=0$ и 1. Столь же легко из определений вытекает справедливость соотношений группы II. Соотношения (6.1)—(6.5) вытекают из того, что операции \oplus и \cdot являются соответственно сложением и умножением по модулю 2. Для доказательства (5.1), (5.2) и (4) следует отдельно рассмотреть случаи $s=0$ и 1 и воспользоваться уже доказанными ранее соотношениями. Если получена левая формула, то правая формула может быть получена, если поменять местами \max и \min , 0 и 1.

Различные авторы используют различные обозначения: так вместо \vee используются $+$, \cup ; вместо « \cdot » — \wedge , \cap ; вместо $\bar{a} - a'$, $\neg a$, $\sim a$.

Таблица 7.3

Формула	Соотношение III
(4)	$s \oplus t = (s \cdot t) \vee (s \cdot \bar{t})$
(5.1)	$s = 1 \oplus s$
(5.2)	$s \vee t = s \oplus t \oplus (s \cdot t)$
(6.1) Ассоциативность	$(s \oplus t) \oplus u = s \oplus (t \oplus u)$
(6.2) Коммутативность	$s \oplus t = t \oplus s$
(6.3) Дистрибутивность	$s \cdot (t \oplus u) = (s \cdot t) \oplus (s \cdot u)$
(6.4) Существование обратного по сложению \oplus элемента	$s \oplus s = 0$
(6.5) Существование нулевого элемента	$s \oplus 0 = s$

Связь с исчислением высказываний. Предложения, для которых имеет смысл говорить об их истинности или ложности, называются *высказываниями*. Мы будем игнорировать существо высказывания и интересоваться лишь его истинностью или ложностью. При этом истинное значение будем обозначать символом 1, а ложное — символом 0. Если P и Q — высказывания, то из них можно построить следующие сложные высказывания.

1) «Не P » (истинно, когда P ложно, и ложно, когда P истинно; обозначается через \bar{P}).

2) « P или Q » (истинно, когда истинно хотя бы одно из высказываний P или Q , ложно в остальных случаях; обозначается через $P \vee Q$).

3) « P и Q » (истинно, только когда истинны оба высказывания P и Q ; обозначается через $P \cdot Q$).

4) «Несовпадение P и Q » (истинно, когда одно из высказываний истинно, а другое ложно, ложно в остальных случаях).

5) «Если P , то Q » (при истинном P ложно, если ложно Q , в остальных случаях истинно; обозначается через $P \rightarrow Q$ *).

6) « P и Q равны» (истинно тогда и только тогда, когда P и Q оба либо истинны, либо ложны; обозначается через $P = Q$).

7) « P и Q одновременно неверны» (ложно, когда оба высказывания истинны, истинно в остальных случаях; обозначается через $P \downarrow Q$ и называется функцией или штрихом Шеффера). Если x и y рассматривать как логические переменные, показывающие истинность и ложность высказываний P , Q , то истинность и ложность введенных ранее сложных высказываний будут описываться соответственно следующими соотношениями: 1) \bar{x} ; 2) $x \vee y$; 3) $x \cdot y$; 4) $x \oplus y$; 5) $x \cdot y$; 6) $x \oplus y$; 7) $x \cdot y$.

Задача 7.10. Проверьте справедливость последнего утверждения.

Операции «—», « \vee », « \cdot », « \oplus » называются соответственно отрицанием, логической суммой, логическим произведением, исключающим ИЛИ, поскольку они связаны с описанными операциями в V_1 . Иногда операции «—», « \vee », « \cdot » удобно интерпретировать соответственно как НЕ, ИЛИ и И (см. § 8.3).

Введение операций над функциями из \mathcal{B}_n . Как следует из упражнения 7.1, D_1 задает взаимно однозначное отображение \mathcal{B}_n в множество всех подмножеств V_n . Для произвольных функций $f, g \in \mathcal{B}_n$ определим операции « \vee », « \cdot », «—», « \oplus », положив:

$$1) h = f \vee g, \quad (7.11)$$

где $h \in \mathcal{B}_n$ — функция, для которой $D_1(h) = D_1(f) \cup D_1(g)$;

$$2) h = f \cdot g, \quad (7.12)$$

где $h \in \mathcal{B}_n$ — функция, для которой $D_1(h) = D_1(f) \cap D_1(g)$;

$$3) h = \bar{f}, \quad (7.13)$$

* $P \rightarrow Q$ называется импликацией, см. также примечание на стр. 117. (Прим. ред.)

где $h \in \mathcal{B}_n$ — функция, для которой $D_1(h) = D_1(f)^c$;

$$4) \quad h = f \oplus g, \quad (7.14)$$

где $h \in \mathcal{B}_n$ — функция, для которой $D_1(h) = D_1(f) \oplus D_1(g)$.

Вообще говоря, следовало бы использовать различные символы для обозначения операций в V_1 и \mathcal{B}_n , но поскольку смысл обозначений обычно понятен из контекста, то далее для обозначения введенных операций используются те же символы, что и раньше. Получающиеся в результате применения этих операций функции называются соответственно логической суммой, логическим произведением, отрицанием, исключающим ИЛИ и т. д. Значения функций $f \vee g$, $f \cdot g$, \bar{f} , $f \oplus g$ в точке $A \in V_n$ будем обозначать через $f \vee g(A)$, $f \cdot g(A)$, $\bar{f}(A)$, $f \oplus g(A)$.

У п р а ж н е н и е 7.3. Для произвольных элементов $A \in V_n$, $f, g \in \mathcal{B}_n$ справедливы соотношения:

$$1) \quad f \vee g(A) = f(A) \vee g(A);$$

$$2) \quad f \cdot g(A) = f(A) \cdot g(A);$$

$$3) \quad \bar{f}(A) = \overline{f(A)};$$

$$4) \quad f \oplus g(A) = f(A) \oplus g(A)$$

(символ \vee и другие в левой части обозначают операции в \mathcal{B}_n ; те же символы в правой части обозначают операции в V_1).

Решение. Пользуясь определением (7.11), получаем:

$$\vee g(A) = \iff A \in D_1(f \vee g) = D_1(f) \cup D_1(g);$$

$$f \vee g(A) = 1 \iff A \in D_1(f) \text{ или } A \in D_1(g);$$

$$f \vee g(A) = 1 \iff f(A) = 1 \text{ или } g(A) = 1;$$

$$f \vee g(A) = 1 \iff f(A) \vee g(A) = 1 \quad (\text{воспользовались также (7.7)}).$$

Остальные соотношения доказываются аналогично.

Задача 7.11'. Система $(\mathcal{B}_n, \vee, \cdot, ')$ является булевой алгеброй, изоморфной $(p(V_n), \cup, \cap, ^c)$. Следовательно, для произвольных элементов $s, t, u, v \in \mathcal{B}_n$ справедливы группы соотношений I и II закона двойственности.

Указания. D_1 — взаимно однозначное отображение \mathcal{B}_n на $p(V_n)$. Отсюда, из соотношения (7.5) и определений (7.11)—(7.13) следует справедливость первой части утверждения задачи. Из первой части и результатов предыдущей главы следует вторая часть утверждения.

Задача 7.12. Система $(\mathcal{B}_n, \oplus, \cdot)$ является булевым кольцом, изоморфным $(p(V_n), \oplus, \cap)$, и для произвольных элементов $s, t, u \in \mathcal{B}_n$ справедливы соотношения группы III.

Указание. D_1 — взаимно однозначное отображение \mathcal{B}_n на $p(V_n)$. Отсюда, из определений (7.12), (7.14) и задачи (6.18) следует справедливость первого утверждения задачи. Вторая часть утверждения является следствием задачи 7.11 и результатов § 6.3.

Задача 7.13. Пользуясь непосредственно результатами упражнения 7.3 и задачи 7.9, проверьте, что для произвольных элементов $s, t, u, v \in \mathcal{B}_n$ выполняются соотношения групп I, II и III.

Указание. Поскольку значение функции в каждой точке $A \in V_n$ удовлетворяет соотношениям групп I, II и III (задача 7.9), то для решения задачи можно воспользоваться соотношением (7.4) и результатами упражнения 7.3.

Соглашения об обозначениях (порядок выполнения операций и исключение скобок). 1) Так как операции \vee, \cdot, \oplus как в V_1 , так и \mathcal{B}_n ассоциативны, то вместо $(s \vee t) \vee u$ можно пользоваться записью $s \vee t \vee u$, т. е. не писать скобки.

2) Если при выполнении операций выполнять операцию «—» раньше \vee, \cdot, \oplus , а операцию « \cdot » раньше \vee и \oplus , то исключение скобок, как правило, не будет приводить к путанице.

3) Обычно мы будем опускать знак операции \cdot и вместо $s \cdot t$ писать просто st .

Замечание. В тех случаях, когда применяется закон двойственности, всегда должны учитываться соглашения (2) и (3) об исключении скобок.

Задача 7.14. Если $a_i \in V_1 (1 \leq i \leq n)$, то

$$\max \{a_1, a_2, \dots, a_n\} = a_1 \vee a_2 \vee \dots \vee a_n;$$

$$\min \{a_1, a_2, \dots, a_n\} = a_1 \cdot a_2 \cdot \dots \cdot a_n.$$

Указание. Для решения этой задачи следует воспользоваться определениями (7.7), (7.8) и индукцией по n :

$$\max \{a_1, a_2, \dots, a_n\} = \max \{\max \{a_1, a_2, \dots, a_{n-1}\}, a_n\}; \quad \min \{a_1, a_2, \dots, a_n\} = \min \{\min \{a_1, a_2, \dots, a_{n-1}\}, a_n\}.$$

Задача 7.15. Сумма $a_1 \oplus a_2 \oplus \dots \oplus a_n$, где $a_i \in V_1 (1 \leq i \leq n)$, равна 0, если среди символов a_1, a_2, \dots, a_n имеется четное число единиц, и равна 1, если число единиц нечетно.

Указание. Утверждение доказывается индукцией по n .

Задача 7.16. Пусть $s_i \in V_1$ (или \mathcal{B}_n). Докажите следующие соотношения, являющиеся обобщением правила Де Моргана:

$$\overline{(s_1 \vee s_2 \vee \dots \vee s_m)} = \bar{s}_1 \bar{s}_2 \dots \bar{s}_m;$$

$$\overline{s_1 s_2 \dots s_m} = \bar{s}_1 \vee \bar{s}_2 \vee \dots \vee \bar{s}_m;$$

Указание. Следует воспользоваться правилом Де Моргана и индукцией по n .

Задача 7.17. Пусть $f, g, h \in \mathcal{B}_n$. Докажите следующие формулы:

$$1) \quad g \vee g\bar{f} = g \vee f, \quad g(\bar{g} \vee f) = gf;$$

$$2) \quad (f \vee g)(\bar{f} \vee h)(g \vee h) = (f \vee g)(\bar{f} \vee h);$$

$$\bar{f}g \vee \bar{f}h \vee gh = \bar{f}g \vee \bar{f}h.$$

Решение. Формулы (1) следуют из закона дистрибутивности (1.5) и закона дополнения (1.6)—(1.8) группы соотношений I. Для доказательства формул (2) их следует вначале преобразовать с помощью соотношений (1.1)—(1.8). Для произвольного A такого, что $f(A) = 1$, значение левой части первой из этих формул, согласно упражнению 7.3, равно $(1 \vee g(A))(0 \vee h(A))$, $(g(A) \vee h(A)) = h(A)(g(A) \vee h(A)) = h(A)$ (здесь были использованы законы дополнения, коммутативности и поглощения группы соотношений II для элементов V_1). Значение функции, стоящей в правой части, также равно $h(A)$. Аналогично одинаковые значения $g(A)$ принимают левая и правая части этой формулы в.

точках A таких, что $f(A)=0$. Так как значения левой и правой частей равны при всех A , то функции, стоящие в левой и правой частях рассматриваемой формулы, совпадают. Вторая из формул (2) может быть получена либо аналогично, либо из закона двойственности.

Задача 7.18. Пусть $f, g_1, g_2, h \in \mathcal{B}_n$. Тогда:

1) если $f = g \oplus h$, то $g = f \oplus h$, т. е. можно переносить слагаемые из одной части равенства в другую;

2) если $f \oplus g_1 = f \oplus g_2$, то $g_1 = g_2$, т. е. можно производить сокращение;

3) если $f \neq 0$ и $f \vee g_1 = f \vee g_2$, то равенство $g_1 = g_2$, вообще говоря, не выполняется.

Указание. Из соотношения (6.4) группы формул III следует равенство (1): $f \oplus g = g \oplus h \oplus h = g$. Аналогично получается равенство (2). Для доказательства утверждения (3) следует положить $g_1 = 0$ и $g_2 = f$.

7.3. Методы задания булевых функций. Базисные функции [12,13]

Для задания (частичной) булевой функции обычно используются два метода:

1) тем или иным способом задаются два из трех множеств $D_0(f)$, $D_1(f)$, $D_{ud}(f)$ (если $f \in \mathcal{B}$, то определяется либо $D_0(f)$, либо $D_1(f)$);

2) выбирается несколько булевых функций, образующих соответствующий базис, и f определяется как их комбинация.

В этом параграфе рассматривается, главным образом, первый метод и вводятся некоторые базисные функции. Второй метод будет рассмотрен в следующем параграфе.

Пример 7.1. Булевы функции одной переменной. В данном случае $|\mathcal{B}_1| = 2^2 = 4$ и функцию f можно определить одним из следующих четырех способов: 1) $f(0) = f(1) = 0$; 2) $f(0) = f(1) = 1$; 3) $f(0) = 0$, $f(1) = 1$; 4) $f(0) = 1$, $f(1) = 0$. В первом и втором случаях получаются функции-константы, которые (согласно принятым в предыдущем параграфе соглашениям) обозначаются через 0 и 1. Третья функция тождественно равна переменной; если переменной является x , то эта функция обозначается либо через $I(x)$, либо просто через x . Четвертая функция называется функцией отрицания (или инверсии) и обозначается через $N(x)$. По определению, $N(x)$ совпадает с отрицанием \bar{x} тождественной функции x . Вместо $N(x)$ далее будет использоваться главным образом обозначение \bar{x} .

Если функция $f(x_1, x_2, \dots, x_n)$ от n переменных x_1, x_2, \dots, x_n зависит только от группы переменных $x_{i_1}, x_{i_2}, \dots, x_{i_m}$ ($1 \leq i_1 < i_2 < \dots < i_m \leq n$), она нередко записывается в виде $g(x_{i_1}, x_{i_2}, \dots, x_{i_m})$. Однако если переменными являются x_1, x_2, \dots, x_n , функция $g(x_{i_1}, x_{i_2}, \dots, x_{i_m})$ считается функцией всех этих n переменных. Если $f(x_1, x_2, \dots, x_n)$ зависит только от x_i и при $x_i = a \in V_1$ принимает значение a независимо от значений других переменных,

то вместо $f(x_1, x_2, \dots, x_n)$ пишут x_i . При этом x_i обозначает функцию, а не логическую переменную.

Пример 7.2. Функции ИЛИ и И. Если $\max_n(x_1, x_2, \dots, x_n)$ понимать в смысле (7.3), то получается функция от n переменных x_1, x_2, \dots, x_n , которая принимает значение, равное максимальному из значений всех переменных x_1, \dots, x_n . Аналогично $\min_n(x_1, x_2, \dots, x_n)$ является функцией, указывающей минимальное значение входящих в нее переменных. Другими словами, первая функция принимает значение 1, если хотя бы одно из значений переменных равно 1, и 0, если все переменные принимают значение 0. Вторая функция, наоборот, принимает значение 1, если равны 1 значения всех переменных, и 0 в противном случае. Нижний индекс n в обозначениях рассмотренных функций может отсутствовать.

Задача 7.19. Докажите следующие формулы:

$$\max_n(x_1, x_2, \dots, x_n) = x_1 \vee x_2 \vee \dots \vee x_n; \quad (7.15)$$

$$\min_n(x_1, x_2, \dots, x_n) = x_1 \cdot x_2 \cdot \dots \cdot x_n = x_1 x_2 \dots x_n \quad (7.16)$$

(в правой части формул x_i представляют собой функцию из \mathcal{B}_n).

Решение. Из упражнения 7.3 и определений \max_n и x_i для $(a_1, a_2, \dots, a_n) \in V_n$ получаем $\max_n(a_1, a_2, \dots, a_n) = \max_n\{a_1, a_2, \dots, a_n\}$. Последнее выражение, согласно соотношениям из задачи 7.14, равно значению $a_1 \vee a_2 \vee \dots \vee a_n$ функции, стоящей в правой части первой формулы. Аналогично доказывается вторая формула.

Принимая во внимание равенства (7.14) и (7.15), функции \max_n и \min_n можно называть соответственно функциями ИЛИ и И от n переменных. Отрицания \max_n и \min_n называют соответственно функциями ИЛИ—НЕ и И—НЕ от n переменных.

Задача 7.20. Функции \max_n , \min_n являются монотонными.

Указание. Для каждого $1 \leq i \leq n$ имеем $\max_{n, x_i=0} \leq \max_{n, x_i=1}$. Аналогично доказывается монотонность \min_n .

Задача 7.21. $0 \leq \min_n \leq \max_n \leq 1$.

Решение. Очевидно, ввиду равенств $D_1(\min_n) = \{(1, 1, \dots, 1)\}$, $D_1(\max_n) = V_n - \{(0, 0, \dots, 0)\}$.

Пример 7.3. Функция четности. Линейные функции. Булева функция p_n от n переменных x_1, x_2, \dots, x_n называется *функцией четности* от n переменных, если она принимает значения 0, когда четное число переменных принимает значение 1, и 1 в противном случае. Если элементы 0, 1 из V_1 рассматривать как целые числа, а x_1, x_2, \dots, x_n — как целочисленные переменные, то значение функции $p_n(x_1, x_2, \dots, x_n)$ будет равно сумме $x_1 + x_2 + \dots + x_n \pmod{2}$. Отсюда, из упражнения 7.3 и задачи 7.15 получаем:

$$p_n(x_1, x_2, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n. \quad (7.17)$$

Функции p_n , ($n \geq 1$), отрицания p_n , т. е. $\bar{p}_n = 1 \oplus p_n$, и константы называются линейными функциями.

Задача 7.22. При $n \geq 2$ линейные функции не являются ни отрицательными, ни положительными ни по одной из переменных.

Указание. Достаточно показать, что p_n не является ни положительной, ни отрицательной по переменной x_1 . Согласно (7.17), $p_n(x_1, 0, 0, \dots, 0) = x_1$, $p_n(x_1, 1, 0, \dots, 0) = \bar{x}_1$, т. е. x_1 не является ни отрицательной, ни положительной переменной для p_n .

Весом функции $f \in \mathcal{B}_n$ называют число $|D_1(f)|$.

У п р а ж н е н и е 7.4. Выпишите все логические функции от двух переменных, используя функции x_1, x_2 , постоянные 0, 1 и операции $\vee, \cdot, \oplus, -$.

Решение. Воспользуемся классификацией функций f по их весу w , $0 \leq w \leq 4$:

1) если $w=0, 4$, то искомыми функциями могут быть только константы 0, 1;

2) если $w=1$, то множеством $D_1(f)$ может быть одна из следующих пар: $\{(0, 0)\}, \{(0, 1)\}, \{(1, 0)\}, \{(1, 1)\}$. Этим парам соответствуют функции $\bar{x}_1\bar{x}_2, \bar{x}_1x_2, x_1\bar{x}_2, x_1x_2$;

3) если $w=3$, то вес \bar{f} равен 1 и по правилу Моргана искомыми функциями будут $x_1\vee x_2, x_1\vee \bar{x}_2, \bar{x}_1\vee x_2, \bar{x}_1\vee \bar{x}_2$;

4) вес функций $x_1, \bar{x}_1, x_2, \bar{x}_2$ равен 2. Так как $|\mathcal{B}_2|=16$, то осталось всего две функции: $x_1\oplus x_2, 1\oplus x_1\oplus x_2$, имеющие вес 2. Так как все шесть рассмотренных функций веса 2 различны, то мы получили все булевы функции от двух переменных.

З а д а ч а 7.23. В коридоре одна лампочка и два переключателя: один у входа, другой у выхода. Обозначим переключатели через S_1 и S_2 . Два состояния лампочки: горит и не горит — обозначим соответственно через 1 и 0. Положение каждого переключателя также будем описывать символами 0, 1. Если лампочка горит, то переводом любого переключателя в противоположное состояние лампочку можно выключить, и, наоборот, если лампочка не горит, то с помощью любого переключателя лампочку можно включить. Пусть x_1, x_2 и y — переменные, описывающие соответственно состояния S_1, S_2 и лампочки. Найдите функцию f переменных x_1 и x_2 , которая дает значения y , предполагая, что $y=0$ при $x_1=x_2=0$.

Указание. Можно построить непосредственно таблицу значений искомой функции. Однако поскольку $f(0, x_2)=x_2, f(1, x_2)=\bar{f}(0, x_2)=\bar{x}_2$, то, как нетрудно видеть, $f(x_1, x_2)=x_1\oplus x_2$.

Пример 7.4. Мажоритарные функции. Пусть l — неотрицательное целое число и $n=2l+1$. Логическая функция от n переменных x_1, x_2, \dots, x_n , принимающая значения 1, если $l+1$ или более (т. е. более половины) переменных принимают значение 1, и 0 в противном случае, называется *мажоритарной функцией* от n переменных и обозначается через Maj_n . Ясно, что

$$D_1(\text{Maj}_n) = \{A | w(A) \geq n/2, A \in V_n\}. \quad (7.18)$$

З а д а ч а 7.24. Покажите, что

$$\text{Maj}_3(x_1, x_2, x_3) = x_1x_2 \vee x_2x_3 \vee x_3x_1.$$

Указание. Покажите, что (правая часть) равна 1 тогда и только тогда, когда $w(A) > 1, A \in V_3$.

У п р а ж н е н и е 7.5. Сложение двоичных чисел. Рассмотрим процедуру нахождения двоичного числа $s_n s_{n-1} \dots s_0$, являющегося

суммой двух n -разрядных двоичных чисел: $a_{n-1}a_{n-2} \dots a_0$ и $b_{n-1}b_{n-2} \dots b_0$. Разряды a_0, b_0, s_0 будем считать младшими, а $a_{n-1}, b_{n-1}, s_{n-1}$ — старшими.

При обычном вычислении суммы сложение выполняется последовательно, начиная с младших разрядов. При вычислении i -го разряда находятся:

$$1) s_i \equiv a_i \oplus b_i \oplus c_i \pmod{2}, \text{ т.е.}$$

$$s_i = a_i \oplus b_i \oplus c_i, 0 \leq i < n, \quad (7.19)$$

где c_i — символ переноса из предыдущего разряда:

2) символ переноса в следующий разряд c_{i+1} , равный 1, если $a_i + b_i + c_i \geq 2$, и 0, если $a_i + b_i + c_i < 2$, т. е.

$$c_{i+1} = \text{Maj}_3(a_i, b_i, c_i), 0 \leq i < n. \quad (7.20)$$

При этом

$$3) c_0 = 0; \quad 4) s_n = c_n. \quad (7.21), (7.22)$$

Таблица значений функции, таблица истинности и геометрическое представление. Таблица, сопоставляющая каждому из 2^n наборов значений переменных соответствующее ему значение функции, например так, как это сделано в табл. 7.4, называется *таблицей значений* или *таблицей истинности функции*. Если функция не определена на некотором наборе, то вместо значения функции в соответствующей клетке таблицы обычно ставится один из знаков «—», «х», «н» или эта клетка оставляется пустой. Таблица 7.4. является таблицей значений мажоритарной функции Maj_3 .

Таблица 7.4. Таблица значений мажоритарной функции трех переменных

$J(x_1, x_2, x_3)$	$x_1 x_2 x_3$	Maj_3
0	0 0 0	0
1	0 0 1	0
2	0 1 0	0
3	0 1 1	1
4	1 0 0	0
5	1 0 1	1
6	1 1 0	1
7	1 1 1	1

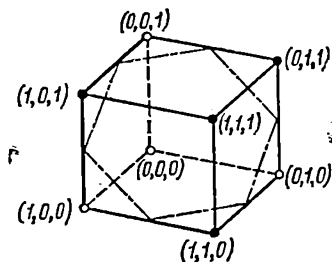


Рис. 7.2. Представление с помощью единичного куба мажоритарной функции от трех переменных

Задача 7.25. Постройте таблицы значений функции \max_2, \min_2, p_2 .

Для наглядности можно воспользоваться представлением с помощью единичного n -мерного куба и пометить черными (белыми) кружочками те из его вершин (a_1, a_2, \dots, a_n) , в которых значения рассматриваемой функции $f(a_1, a_2, \dots, a_n)$ равны 1 (соответственно 0). На рис. 7.2 показано такое представление для функции Maj_3 . Пунктирной линией на этом рисунке показано сечение куба плоскостью. Эта плоскость делит пространство на две области, одна

из которых содержит все черные, а другая — все белые кружочки. Функции, обладающие таким геометрическим свойством, называются *линейно разделимыми* (см. упражнение 7.5). На рис. 7.3 приведено геометрическое представление функции $x_1 \oplus x_2$. Заметим, что в данном случае невозможно провести прямую, которая разделила бы черные и белые кружочки. Этим же свойством обладает и функция $1 \oplus x_1 \oplus x_2$. В § 9.2.2 будут рассмотрены таблицы определения функции (диаграммы Карно), использующие преимущества геометрического представления.

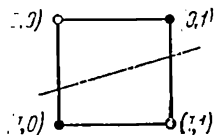


Рис. 7.3. Геометрическое представление $x_1 \oplus x_2$

Пример 7.5. Пороговые функции. Пороговая функция является обобщением мажоритарной функции. Функция $f \in \mathcal{B}_n$ называется пороговой функцией n переменных, если существуют $n+1$ целых чисел c_0, c_1, \dots, c_n таких, что

$$D_1(f) = \{(a_1, a_2, \dots, a_n) \mid \sum_{i=1}^n c_i a_i > c_0, (a_1, a_2, \dots, a_n) \in V_n\}. \quad (7.23)$$

(Здесь в сумме $\sum c_i a_i$ элементы 0, 1 из V_1 рассматриваются, как целые числа 0, 1, а произведения $c_i a_i$ складываются, как целые числа.)

Если воспользоваться представлением V_n с помощью единичного куба, то приведенное определение означает, что существует гиперплоскость $L: \sum_{i=1}^n c_i x_i = c_0$, такая, что $f=1$ для всех звеньев V_n , лежащих по одну сторону от этой плоскости, и $f=0$ элементов V_n , лежащих по другую сторону гиперплоскости. Поэтому пороговые функции также линейно разделимы.

Задача 7.26. Пороговая функция является смешанной монотонной функцией.

Указание. Если $c_i \geq 0$, то переменная x_i является положительной; если же $c_i \leq 0$, то переменная x_i является отрицательной. Отсюда, из (7.6) и (7.23) следует утверждение задачи.

Задача 7.27. Если $f(x_1, x_2, \dots, x_n)$ — пороговая функция, то функции \bar{f} и $f(x_1, \dots, x_{i-1}, \bar{x}_i, x_{i+1}, \dots, x_n)$ также будут пороговыми.

Решение. Пусть c_0, c_1, \dots, c_n — коэффициенты соотношения (7.23) для функции f . Тогда в качестве соответствующих коэффициентов для функции \bar{f} можно взять $-c_0+1, -c_1, -c_2, \dots, c_n$, а для второй из рассматриваемых функций можно взять $c_0-c_i, c_1, \dots, c_{i-1}, -c_i, c_{i+1}, \dots, c_n$.

Задача 7.28. Функции $\max_n, \min_n, \text{Maj}_{2l+1}$ являются пороговыми.

Указание. Для каждой из этих функций положим $c_i=1, 1 \leq i \leq n$. В качестве c_0 для \max_n, \min_n и Maj_n следует взять соответственно 1, n и $l+1$.

Задача 7.29. Какие из 16 функций \mathcal{B}_2 не являются пороговыми?

Решение. Из упражнения 7.4 и задач 7.22, 7.26—7.28 следует, что среди рассматриваемых функций пороговыми не являются $x_1 \oplus x_2$ и $1 \oplus x_1 \oplus x_2$.

Пороговые функции были введены впервые для моделирования функций нервных клеток, но в дальнейшем изучались в теории схем из функциональных элементов, в теории распознавания образов и других областях. Известно [16], что с ростом n отношение числа пороговых функций в \mathcal{B}_n к мощности $|\mathcal{B}_n|$ очень быстро стремится к нулю. Понятие пороговой функции можно обобщить. Функция $f \in \mathcal{B}_n$ называется k -пороговой функцией n переменных, если она удовлетворяет следующему условию: существуют целые числа c_1, c_2, \dots, c_n ; $c^{(1)}, c^{(2)}, \dots, c^{(k)}$ такие, что при четном k

$$D_1(f) = \bigcup_{j=1}^{k/2} \{(a_1, a_2, \dots, a_n) | c^{(2j-1)} \leq \sum_{i=1}^n c_i a_i \leq c^{(2j)} - 1\};$$

при нечетном k

$$D_1(f) = \bigcup_{j=1}^{(k-1)/2} \{(a_1, a_2, \dots, a_n) | c^{(2j-1)} \leq \sum_{i=1}^n c_i a_i \leq c^{(2j)} - 1\} \cup \\ \cup \{(a_1, a_2, \dots, a_n) | c^{(k)} \leq \sum_{i=1}^n c_i a_i\}.$$

При $k \geq 2$ такие функции называются *многопороговыми* и исследовались ранее с точки зрения реализации их на больших интегральных схемах [14].

Задача 7.30. Произвольная функция $f \in \mathcal{B}_n$ является k -пороговой функцией (при некотором $k \leq 2^n$).

Решение. Если положить $c_i = 2^{n-i}$, то $\sum_{i=1}^n c_i a_i = J(a_1, a_2, \dots, a_n)$. Упорядочим элементы $D_1(f)$ по величине представляемых ими целых чисел. Минимальное из этих целых чисел возьмем в качестве $c^{(1)}$. Если среди оставшихся чисел, больших $c^{(1)}$, имеются числа, для которых разность между ними и всеми меньшими числами из $D_1(f)$ не меньше 2, то минимальное из них следует взять в качестве $c^{(2)}$ и положить $c^{(2)} = c^{(3)} - 1$. Продолжая этот процесс, мы построим $c^{(1)}, \dots, c^{(k)}$, $k \leq 2^n$.

Сложность описания логических функций. Как нетрудно подсчитать, $|\mathcal{B}_2| = 16$, $|\mathcal{B}_3| = 256$, $|\mathcal{B}_4| = 65536$. На практике нередко используются булевы функции более чем от десяти переменных, и $|\mathcal{B}_{10}| = 2^{1024} \sim 10^{309}$, хотя реально используемые функции составляют лишь малую часть множества всех функций. Однако описать конструктивно класс используемых булевых функций практически невозможно. Так, схема параллельного сложения 16-разрядных двоичных чисел имеет 32 входа, и если попытаться использовать таблицу истинности, то она должна будет иметь около $2^{32} \approx 10^9$ строк, что, конечно, практически неприемлемо. Удобные методы описания булевых функций, которые могут быть использованы на практике, описаны далее в § 7.4, 8.2 и 8.3.

Задача 7.31. Пусть A — конечное множество, содержащее 2 или более элементов. Пусть L — это некоторое множество последовательностей конечной длины, компонентами которых являются элементы из A (см. § 10.1). Множество L называется описанием булевых функций, если существует отображение φ множества L в множество \mathcal{B} всех булевых функций такое, что для произвольной функции $f \in \mathcal{B}$ найдется, по крайней мере, одна последовательность $\alpha \in L$, для которой $\varphi(\alpha) = f$. Говорят, что последовательность $\alpha \in L$ указывает функцию $\varphi(\alpha)$. Минимальное значение длины последовательности из L , которая указывает функцию $f \in \mathcal{B}$ (вообще говоря, таких последовательностей может быть несколько), обозначим через $l(f)$. Покажите, что:

$$1) \max_{f \in \mathcal{B}_n} l(f) > 2^n / \log_2 |A| - 1;$$

2) число булевых функций, которые могут быть указаны с помощью последовательностей из L длины $< 2^n(1-\varepsilon)/\log_2 |A|$, где $0 < \varepsilon < 1$, не превосходит $(2^{2^n(1-\varepsilon)} - 1) / (|A| - 1)$.

Указание. Левую часть неравенства (1) следует положить равной l . Далее, пользуясь тем, что число последовательностей длины $\leq l$ с компонентами из A не меньше, чем $|A|^l$, неравенство следует переписать в удобной форме и прологарифмировать. Утверждение (2) также доказывается легко.

Как следует из утверждения (2) последней задачи, если не рассматривать очень малые значения n , при любом выборе L для указания подавляющей части логических функций длина последовательностей должна быть близкой к $2^n / \log_2 |A|$ или даже превосходить это число. Заметим, что длина определений функций \max_n , \min_n , p_n , Maj_n , приводившихся ранее в качестве примеров, не зависит от n . Конечно, в этих определениях отсутствует запись переменных x_1, x_2, \dots, x_n и конкретного значения числа n , широко используется понятие четного числа и т. д. Главным же образом краткость этих определений связана с тем, что все указанные функции имеют простую структуру, обусловленную их симметричностью.

Определение симметричной функции

Если функция $f(x_1, x_2, \dots, x_n) \in \mathcal{B}_n$ такова, что при $1 \leq i < j \leq n$ $f(x_1, x_2, \dots, x_n) = f(x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_n)$, т. е. если значение функции не изменяется при замене местами переменных x_i и x_j , то f называется *симметричной относительно x_i, x_j* . Аналогично определяется симметричность функции относительно переменных $x_{i_1}, x_{i_2}, \dots, x_{i_r}$, $1 \leq i_1 < i_2 < \dots < i_r \leq n$. Функция, симметричная относительно всех переменных x_1, x_2, \dots, x_n , называется просто *симметричной*.

Если $f \in \mathcal{B}_n$ — симметричная функция, A, B — два произвольных элемента из V_n , имеющих одинаковый вес ($w(A) = w(B)$), то

$f(A)=f(B)$. Следовательно, для того чтобы определить область $D_1(f)$, достаточно указать

$$W_1(f) = \{i | f(A) = 1, \omega(A) = i, A \in V_n\}. \quad (7.24)$$

Наоборот, для любого подмножества $W \subset \{1, 2, \dots, n\}$ множество $D_1(f) = \{A | \omega(A) \in W, A \in V_n\}$ однозначно задает симметричную функцию. Функцию $f \in \mathcal{B}_n$, определяемую множеством $W_1(f) = \{i_1, i_2, \dots, i_r\}$, где $0 \leq i_1 < i_2 < \dots < i_r \leq n$, обозначают через $S_{i_1, i_2, \dots, i_r}^{(n)}$. Например, $x_1 \oplus x_2 \oplus x_3 = S_{1,3}^{(3)}$.

Имеется еще один способ задания симметричной функции от n переменных; каждому целому числу $0 \leq i \leq n$ сопоставляются символ ω_i , равный 1, если $i \in W_1(f)$, и 0 в противном случае. При этом симметричная функция задается двоичной последовательностью $\omega_0, \omega_1, \dots, \omega_n$ длины $n+1$ с компонентами 0, 1.

Задача 7.32. Общее число симметричных функций в \mathcal{B}_n равно 2^{n+1} .

Как мы видели, функции $\max_n, \min_n, \text{Maj}_n$ являются монотонными.

Задача 7.33. Покажите, что $\max_n, \min_n, \text{Maj}_n$ — симметричные монотонные функции.

Если функция $f \in \mathcal{B}_n$ является симметричной монотонной, то существует целое число t ($0 \leq t \leq n$) такое, что $W_1(f) = \{i | t \leq i \leq n\}$, и, следовательно, для задания такой функции достаточно указать лишь число t . Такие функции обозначают либо через $S_{t \leq i \leq n}^{(n)}$, либо через $T_t^{(n)}$. Например,

$$\max_n = T_1^{(n)}, \quad \min_n = T_n^{(n)}, \quad \text{Maj}_{2l+1} = T_{l+1}^{(2l+1)}.$$

Задача 7.34. Общее число симметричных монотонных функций в \mathcal{B}_n равно $n+1$.

Задача 7.35. Симметричные монотонные функции являются пороговыми функциями.

Указание. Для $T_t^{(n)}$ в качестве коэффициентов соотношения (7.23) следует взять $c_1 = c_2 = \dots = c_n = 1, c_0 = t$.

7.4. Разложение булевых функций [12,13]

Здесь будет описано, каким образом произвольную функцию из \mathcal{B}_n , зависящую от переменных x_1, x_2, \dots, x_n , можно представить с помощью функций 0, 1, x_i ($1 \leq i \leq n$) и операций $\vee, \cdot, -, \oplus$.

Упражнение 7.6. Для функции $f(x_1, x_2, \dots, x_n) \in \mathcal{B}_n$ справедливы соотношения:

$$\bar{f} = \bar{x}_1 \cdot f_{x_1=0} \vee x_1 \cdot f_{x_1=1} \quad (7.25)$$

$$= (\bar{x}_1 \vee f_{x_1=0}) \cdot (x_1 \vee f_{x_1=1}) \quad (7.26)$$

$$= \bar{x}_1 \cdot f_{x_1=0} \oplus x_1 \cdot f_{x_1=1}. \quad (7.27)$$

Решение. Как следует из упражнения (7.3), определения функции x_i , соотношений (1.6) и (1.7) группы I и соотношения (6.5) группы III на любом наборе значений переменных $(0, a_2, a_3, \dots, a_n) \in V_n$ правые части доказываемых

соотношений равны $f(0, a_2, a_3, \dots, a_n)$, т. е. совпадают с левой частью. Аналогично можно проверить, что левые и правые части также совпадают на наборах значений переменных вида $(1, a_2, a_3, \dots, a_n) \in V_n$.

Задача 7.36. Покажите, что если переменная x_i функции $f(x_1, x_2, \dots, x_n) \in \mathcal{B}_n$ является положительной, то

$$\bar{f} = f_{x_i=0} \vee x_i \cdot f_{x_i=1}; \quad (7.28)$$

$$f = f_{x_i=0} \cdot (x_i \vee f_{x_i=1}). \quad (7.29)$$

Решение. Согласно (7.6), $f_{x_i=0} \leq f_{x_i=1}$. Из соотношений (3.6) и (3.4) группы II получаем соответственно, что

$$x_i \cdot f_{x_i=0} \leq x_i f_{x_i=1} \text{ и } x_i \cdot f_{x_i=1} = x_i f_{x_i=0} \vee x_i \cdot f_{x_i=1}.$$

Из (7.25) имеем:

$$f = \bar{x}_i \cdot f_{x_i=0} \vee x_i \cdot f_{x_i=1} = \bar{x}_i \cdot f_{x_i=0} \vee x_i \cdot f_{x_i=0} \vee x_i \cdot f_{x_i=1};$$

$$f = (\bar{x}_i \vee x_i) f_{x_i=0} \vee x_i \cdot f_{x_i=1} \text{ (дистрибутивность);}$$

$$f = f_{x_i=0} \vee x_i \cdot f_{x_i=1} \text{ (соотношения (1.8), (1.6) группы I).}$$

Аналогично доказывается (7.29).

Задача 7.37. Если $f(x_1, x_2, \dots, x_n) = \bar{f}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$, то $f = x_1 \oplus \oplus f_{x_1=0}$.

Решение. По предположению, $f_{x_1=1} = \bar{f}_{x_1=0}$. Отсюда и (7.25) получаем $f = \bar{x}_1 \cdot f_{x_1=0} \vee x_1 \cdot f_{x_1=0} = x_1 \oplus f_{x_1=0}$ (здесь мы воспользовались также соотношением (4) группы III). Эта задача является обобщением задачи 7.23.

Обозначения

Для переменной x и $a \in V_1$ определим x^a , положив

$$x^0 = \bar{x}, \quad x^1 = x. \quad (7.30)$$

По определению,

$$\overline{x^a} = x^{\bar{a}}, \quad x^a = 1 \Leftrightarrow x = a. \quad (7.31), (7.32)$$

Определение элементарной конъюнкции и элементарной дизъюнкции. Функции вида $x_{i_1}^{a_1} x_{i_2}^{a_2} \dots x_{i_r}^{a_r}$, $x_{i_1}^{a_1} \vee x_{i_2}^{a_2} \vee \dots \vee x_{i_r}^{a_r}$, где $1 \leq i_1 < i_2 < \dots < i_r \leq n$, $a_i \in V_1$, называются соответственно *элементарной конъюнкцией* и *элементарной дизъюнкцией*. Число r называется *рангом* элементарной конъюнкции (дизъюнкции). Ранг элементарной конъюнкции (дизъюнкции) t обозначается через $\deg t$. Переменные и их отрицания называются при этом буквами. Например, x_i, \bar{x}_i — элементарные конъюнкции ранга 1, а $x_i x_j, \bar{x}_i x_j, x_i \bar{x}_j, \bar{x}_i \bar{x}_j$ ($i < j$) — элементарные конъюнкции ранга 2.

Задача 7.38. (1) $t = x_{i_1}^{a_1} x_{i_2}^{a_2} \dots x_{i_r}^{a_r} \Leftrightarrow x_{i_1} = a_1; x_{i_2} = a_2, \dots, x_{i_r} = a_r$. Следовательно, если $1 \leq i_1 < i_2 < \dots < i_r \leq n$ и t представляет функцию из \mathcal{B}_n , то $|D_1(t)| = 2^{n-t}$.

(2) $t < t' = x_{j_1}^{b_1} x_{j_2}^{b_2} \dots x_{j_l}^{b_l}$ ($1 \leq j_1 < j_2 < \dots < j_l \leq n$) тогда и только тогда, когда все буквы, входящие в t' , входят в t .

Решение. Первая часть утверждения (1) следует из (7.32). Для того чтобы элементарная конъюнкция t была равна 1 на наборе $C = (c_1, c_2, \dots, c_n) \in V_n$, необходимо выполнение равенств $c_{i_k} = a_{i_k}$, $1 \leq r \leq r$. Число таких наборов в V_n равно 2^{n-r} . Отсюда следует вторая часть утверждения (1). Если $x_{i_k} = a_k$ ($1 \leq k \leq r$), то, согласно первой части утверждения (1), $x_{i_k} = b_k$ ($1 \leq k \leq r$). Поэтому для каждого h существует такое k , что $i_k = j_h$, $a_k = b_h$, т. е. буквы j' входят в t . Очевидно, верно и обратное утверждение.

Задача 7.39. (1) Число элементарных конъюнкций в \mathcal{B}_n равно $3^n - 1$.

$$(2) |D_1(x_{i_1}^{a_1} \vee x_{i_2}^{a_2} \vee \dots \vee x_{i_r}^{a_r})| = 2^n - 2^{n-r}.$$

Указание (1). Число элементарных конъюнкций ранга r ($0 \leq r \leq n$) равно $\binom{n}{r} 2^r$. Далее заметим, что $\sum_{r=1}^n \binom{n}{r} 2^r = 3^n - 1$. (2). Отрицание $\bar{x}_{i_1}^{a_1} \times \dots \times \bar{x}_{i_r}^{a_r}$, как следует из задачи 7.16, равно $x_{i_1}^{a_1} \vee x_{i_2}^{a_2} \vee \dots \vee x_{i_r}^{a_r}$. Далее нужно воспользоваться утверждением (1) задачи 7.38.

Минимальный и максимальный термы. Элементарная конъюнкция $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ ранга n и элементарная дизъюнкция $x_1^{a_1} \vee x_2^{a_2} \vee \dots \vee x_n^{a_n}$ ранга n называются соответственно *минимальной конъюнкцией* и *максимальной дизъюнкцией*. Элементарная конъюнкция $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ равна 1 только в том случае, если $x_1 = a_1$, $x_2 = a_2$, ..., $x_n = a_n$; в остальных случаях она равна 0. Элементарная дизъюнкция $x_1^{a_1} \vee x_2^{a_2} \vee \dots \vee x_n^{a_n}$ равна 0 только в том случае, если $x_1 = \bar{a}_1$, $x_2 = \bar{a}_2$, ..., $x_n = \bar{a}_n$, и равна 1 в остальных случаях. Минимальная конъюнкция $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ и ее дополнение $\bar{x}_1^{a_1} \vee \bar{x}_2^{a_2} \vee \dots \vee \bar{x}_n^{a_n}$ обозначаются соответственно через $m_J(a_1, a_2, \dots, a_n)$ и $M_J(a_1, a_2, \dots, a_n)$. Поскольку $0 \leq m_J(a_1, a_2, \dots, a_n)$, то минимальная конъюнкция является минимальным элементом \mathcal{B}_n (см. § 6.4).

Дизъюнктивная и конъюнктивная нормальные формы. Пусть $t_1, t_2, \dots, t_m(s_1, s_2, \dots, s_m)$ — элементарные конъюнкции (дизъюнкции). Выражение $t_1 \vee t_2 \vee \dots \vee t_m(s_1 \cdot s_2 \cdot \dots \cdot s_m)$ называется *дизъюнктивной (конъюнктивной) нормальной формой* с элементарными конъюнкциями (дизъюнкциями). Дизъюнктивную (конъюнктивную) нормальную форму, не содержащую элементарных конъюнкций (дизъюнкций), условимся обозначать через 0 (1).

Задача 7.40. Для произвольной функции $f \in \mathcal{B}_n$ существуют дизъюнктивная и конъюнктивная нормальные формы, представляющие f .

Указание. Для доказательства этого утверждения следует воспользоваться индукцией по числу переменных, соотношениями (7.25), (7.26) и законом дистрибутивности.

Задача 7.41. Произвольную монотонную функцию f , отличную от константы, можно представить в дизъюнктивной (конъюнктивной) нормальной форме, состоящей из элементарных конъюнкций (дизъюнкций), не содержащих отрицаний переменных.

Указание. Для доказательства следует воспользоваться индукцией по числу переменных, соотношениями (7.28), (7.29) и законом дистрибутивности. При $n=1$ условием задачи удовлетворяет только тождественная функция, для которой справедливость доказываемого утверждения очевидна.

Далее вместо $f(A)$ используется запись $f(I)$, если целочисленное представление $J(A)$ элемента A равно I .

У п р а ж н е н и е 7.7. Функция $f \in \mathcal{B}_n$ может быть представлена в виде *

$$X(f) = \bigvee_{I=0}^{2^n-1} c_I \cdot m_I(X), \quad c_I \in V_1,$$

тогда и только тогда, когда для всех $0 \leq I < 2^n$ $c_I = f(I)$.

Решение. Утверждение очевидно, так как $m_I(I') = 1$, если $I = I'$, и $m_I(I') = 0$, если $I \neq I'$.

Согласно приведенному определению,

$$\begin{aligned} f(X) &= \bigvee_{I=0}^{2^n-1} f(I) m_I(X) = \bigvee_{I \in D_1(f)} m_I(X); \\ &= \bigvee_{(a_1, a_2, \dots, a_n) \in V_n} f(a_1, a_2, \dots, a_n) x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}; \\ &= \bigvee_{(a_1, a_2, \dots, a_n) \in D_1(f)} x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}. \end{aligned} \quad (7.33)$$

Правая часть этого выражения называется *совершенной дизъюнктивной нормальной формой* ** функции f . Условимся считать, что совершенной дизъюнктивной нормальной формой константы 0 является сама константа 0. Задание совершенной дизъюнктивной нормальной формы функции, по существу, не отличается от задания $D_1(f)$. Кроме того, заметим, что совершенная дизъюнктивная форма представляет собой частный случай теоремы 6.1.

Определение двойственной и самодвойственной функции. Пусть $f(x_1, x_2, \dots, x_n) \in \mathcal{B}_n$. Функция $\bar{f}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ называется двойственной к функции f и обозначается через f^d . Если $f = f^d$, то функция f называется *самодвойственной*. Например, как следует из задачи 7.16, двойственной к функции $\max_n(x_1, x_2, \dots, x_n) = x_1 \vee x_2 \vee \dots \vee x_n$ является функция $\bar{x}_1 \vee \bar{x}_2 \vee \dots \vee \bar{x}_n = x_1 x_2 \dots x_n = \min_n(x_1, x_2, \dots, x_n)$.

Для $A = (a_1, a_2, \dots, a_n) \in V_n$ положим $\bar{A} = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n)$. Очевидно, $w(\bar{A}) = n - w(A)$.

З а д а ч а 7.42. Если $f_1 \leq f_2$, то $f_2^d \leq f_1^d$. Следовательно, если f_1 и f_2 являются самодвойственными и $f_1 \leq f_2$, то $f_1 = f_2$.

* Запись $\bigvee_{I=0}^{2^n-1} \left(\bigwedge_{I=0}^{2^n-1} \right)$ означает, что берется логическая сумма (произведение) по параметру I , пробегающему все целочисленные значения от 0 до 2^n-1 .

** Авторы называют совершенную дизъюнктивную нормальную форму канонической суммой. (Прим. ред.)

Решение. Так как $f_2(\bar{A}) \leq f_1(\bar{A})$ для произвольного $\bar{A} \in V_n$, то $f_2^d \leq f_1^d$. Вторая часть утверждения, очевидно, следует из первой.

Задача 7.43 (1) Если функция $f \in \mathcal{B}_n$ является самодвойственной, то $|D_1(f)| = 2^{n-1}$. (2) Общее число самодвойственных функций в \mathcal{B}_n равно $2^{2^{n-1}}$

Решение. По определению, $f(\bar{A}) = f(A)$, т. е. если функция на одном из наборов A или \bar{A} принимает значение 0, то на другом ее значение равно 1. Число различных неупорядоченных пар (A, \bar{A}) равно 2^{n-1} . Отсюда следует утверждение (1). Утверждение (2) следует из того, что на каждой неупорядоченной паре (A, \bar{A}) самодвойственную функцию можно определить независимо от других пар двумя различными способами.

Задача 7.44. Самодвойственная функция $f \in \mathcal{B}_n$ может быть представлена в виде

$$f(x_1, x_2, \dots, x_n) = x_1 \cdot g(x_2, \dots, x_n) \vee \bar{x}_1 \cdot g^d(x_2, \dots, x_n),$$

где $g \in \mathcal{B}_n$ — функция, зависящая только от x_2, x_3, \dots, x_n .

Указание. Если положить $g(x_2, x_3, \dots, x_n) = f(0, x_2, x_3, \dots, x_n) = f(1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n)$, то $g^d = \bar{g}(\bar{x}_2, \bar{x}_3, \dots, \bar{x}_n) = f(1, x_2, x_3, \dots, x_n)$. Отсюда и из соотношения (7.25) получается указанное представление функции.

Задача 7.45. Функция четности p_n является самодвойственной тогда и только тогда, когда n нечетно.

Указание. Следует воспользоваться равенством $\bar{x}_i = 1 \oplus x_i$.

Задача 7.46. Самодвойственная монотонная симметрическая функция является пороговой и наоборот.

Решение. Прямое утверждение следует из того, что если $f = T_l^{(n)}$ и $|D_1(f)| = 2^{n-1}$, то $n = 2l + 1$, $t = l + 1$. Обратное утверждение следует из того, что $w(A) > \geq n/2 \Leftrightarrow w(\bar{A}) = n - w(A) < n/2$ для $A \in V_n$.

Упражнение 7.8. Если функция f не является константой, то

$$f(x_1, x_2, \dots, x_n) = \bigvee_{(a_1, a_2, \dots, a_n) \in V_n} (f(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) \vee$$

$$\vee x_1^{a_1} \vee x_2^{a_2} \vee \dots \vee x_n^{a_n}).$$

Решение. Возьмем отрицание от обеих частей совершенной дизъюнктивной нормальной формы функции f^d , двойственной для f ,

$$f(x_1, x_2, \dots, x_n) = \bigvee_{(a_1, a_2, \dots, a_n) \in V_n} (f(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) \vee x_1^{a_1} \vee x_2^{a_2} \vee \dots \vee x_n^{a_n})$$

Если каждую переменную x_i заменить на \bar{x}_i и далее воспользоваться соотношениями задачи 7.16 и равенством $(\bar{x})^a = x^a$, то можно получить требуемое представление функции. Это представление можно переписать следующим образом:

$$f(x) = \bigvee_{(a_1, a_2, \dots, a_n) \in V_n} (f(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) \vee x_1^{a_1} \vee x_2^{a_2} \vee \dots \vee x_n^{a_n});$$

$$f(x) = \bigvee_{(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) \in V_n} (f(a_1, a_2, \dots, a_n) \vee x_1^{\bar{a}_1} \vee x_2^{\bar{a}_2} \vee \dots \vee x_n^{\bar{a}_n});$$

$$f(x) = \prod_{I=0}^{2^n-1} (f(\bar{I}) \nabla M_I);$$

$$f(x) = \prod_{\bar{I} \in D_0(f)} M_I \quad (7.34)$$

(здесь $I = 2^n - 1 - I$).

Правая часть равенства (7.34) называется *совершенной конъюнктивной нормальной формой* *. Условимся считать, что совершенной конъюнктивной нормальной формой константы 1 является сама константа 1.

Задача 7.47. Найдите совершенные дизъюнктивную и конъюнктивную нормальные формы функции $x_1 \oplus x_2 \oplus x_3$.

Задача 7.48. Докажите справедливость при $k \geq 4$ следующих соотношений:

$$(1) \quad x_1 \vee x_2 \vee \dots \vee x_k \geq (x_1 \vee x_2 \vee y_1) (x_3 \vee \bar{y}_1 \vee y_2) (x_4 \vee \bar{y}_2 \vee y_3) \dots (x_{k-2} \vee \bar{y}_{k-4} \vee y_{k-3}) (x_{k-1} \vee x_k \vee \bar{y}_{k-3});$$

(2) если, по крайней мере, одна из переменных x_1, x_2, \dots, x_n равна 1, то значения y_1, y_2, \dots, y_{k-3} можно выбрать равными 0 или 1 таким образом, что правая часть приведенного соотношения будет равна 1.

Указание. (1) При $x_1 = x_2 = \dots = x_k = 0$ функция $(\bar{y}_2 \vee y_3) \dots (\bar{y}_{k-4} \vee y_{k-3}) \bar{y}_{k-3}$ тождественно равна 0.

(2) Если считать, что $x_i = 1$, то следует положить $y_j = 1$ ($1 \leq j \leq i-2$), $y_j = 0$ ($i-1 \leq j \leq k-3$).

Задача 7.49. Докажите равенства:

$$(1) \quad f(x_1, x_2, \dots, x_n) = \bigvee_{(a_1, a_2, \dots, a_m) \in V_m} f(a_1, a_2, \dots, a_m, x_{m+1}, x_{m+2}, \dots,$$

$$\dots, \nabla x_n) x_1^{a_1} x_2^{a_2} \dots x_m^{a_m};$$

$$(2) \quad f(x_1, x_2, \dots, x_n) = \prod_{(a_1, a_2, \dots, a_m) \in V_m} f(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_m, x_{m+1}, x_{m+2}, \dots,$$

$$\dots, \nabla x_n) \vee x_1^{a_1} \vee x_2^{a_2} \vee \dots \vee x_m^{a_m}.$$

Указание. Для произвольного набора $(b_{m+1}, b_{m+2}, \dots, b_n) \in V_{n-m}$ постройте совершенные дизъюнктивную и конъюнктивную нормальные формы функции $f(x_1, x_2, \dots, x_m, b_{m+1}, b_{m+2}, \dots, b_n)$ и с их помощью покажите, что правые и левые части доказываемых равенств совпадают.

Многочлены над Z_2 . В § 7.2 отмечалось, что V_1 можно рассматривать как поле Z_2 , а операции $\oplus, \cdot, -$ как сложение и умножение по модулю 2 (задача 7.8). Пусть x_1, x_2, \dots, x_n — переменные, принимающие значение в множестве V_1 , которое рассматривается как поле Z_2 . Функции вида

$$g = c_0 \oplus \sum_i \oplus c_{i_1 i_2 \dots i_r} x_{i_1} x_{i_2} \dots x_{i_r}, \quad (7.35)$$

* Авторы называют совершенную конъюнктивную нормальную форму каноническим произведением. (Прим. ред.)

где $c_0 \in V_1$, $c_{i_1 i_2 \dots i_r} \in V_1$, называются *многочленами* от переменных x_1, x_2, \dots, x_n . (Знак $\Sigma \oplus$ означает, что сумма берется по всем совокупностям целых чисел i_1, i_2, \dots, i_r таким, что $1 \leq i_1 < i_2 < \dots < i_r \leq n$). Максимальное значение индекса r у ненулевых коэффициентов $c_{i_1 i_2 \dots i_r}$ называется степенью g и обозначается через $\deg g$. Если все коэффициенты равны 0, то будем считать, что многочлен имеет степень 0.

У п р а ж н е н и е 7.9. Для произвольной функции $f \in \mathcal{B}_n$ существует многочлен, представляющий f . Его коэффициенты однозначно определяются функцией f .

Решение. Вначале индукцией по n покажем, что для произвольной функции $f \in \mathcal{B}_n$ существует, по крайней мере, один многочлен, представляющий f . При $n=1$ имеется четыре функции, 0, 1, x_1 , $\bar{x}_1 = 1 \oplus x_1$, и для каждой из них утверждение верно. Предположим, что доказываемое утверждение верно при $n-1$. Из (7.27) имеем

$$f = f_{x_1=0} \oplus x_1 \cdot (f_{x_1=0} \oplus f_{x_1=1}).$$

По предположению индукции, существуют многочлены g_1, g_2 от переменных x_2, x_3, \dots, x_n такие, что $f = g_1 \oplus x_1 g_2$. Правая часть последнего равенства является многочленом от переменных x_1, x_2, \dots, x_n . Покажем далее, что коэффициенты многочлена определяются единственным образом. Общее число различных совокупностей коэффициентов $c_0, c_{i_1 i_2, \dots, i_r}$ ($1 \leq i_1 < i_2 < \dots < i_r \leq n$) много-

членов от переменных x_1, x_2, \dots, x_n равно $\sum_{r=0}^n \binom{n}{r} = 2^n$. Следовательно, число

различных многочленов от переменных x_1, x_2, \dots, x_n равно 2^{2^n} (многочлены считаются различными, если они отличаются хотя бы одним коэффициентом). Так как $|\mathcal{B}_n| = 2^{2^n}$ то ясно, что существует ровно один многочлен, представляющий функцию f .

З а д а ч а 7.50. Покажите существование многочлена, представляющего f , используя совершенную дизъюнктивную нормальную форму функции f .

Указание. Используя равенство $\bar{x}_i = 1 \oplus x_i$, каждую конъюнкцию m_I можно представить с помощью многочлена. Если $I \neq I'$, то $m_I m_{I'} = 0$, и, следовательно, $m_I \vee m_{I'} = m_I \oplus m_{I'}$. Далее следует воспользоваться этим равенством.

Если в (7.35) все коэффициенты при членах степени 2 и выше равны 0, то функции, представляемые такими многочленами, называются *линейными* *.

Задачи

7.1. Пусть P, Q, R — высказывания. Следующее высказывание является истинным (транзитивность импликации): $[(P \rightarrow Q) \cdot (Q \rightarrow R)] \rightarrow (P \rightarrow R)$.

7.2. Найдите функцию $f \in \mathcal{B}_2$, которая существенно зависит от обеих переменных, x_1, x_2 , и удовлетворяет следующим двум условиям:

* Это определение совпадает с определением линейной функции, приведенным ранее (см. пример 7.3). (Прим. ред.)

1) закону коммутативности: $f(x_1, x_2) = f(x_2, x_1)$;

2) закону ассоциативности: $f(x_1, f(x_2, x_3)) = f(f(x_1, x_2), x_3)$.

7.3. Пусть $f \in \mathcal{B}_n$, $X = (x_{i_1}, x_{i_2}, \dots, x_{i_r})$ — совокупность переменных и $A = (a_1, a_2, \dots, a_r) \in V_r$. Обозначим через $f_{X,A}$ и $f_{X,\bar{A}}$ функции, получающиеся при подстановке в $f(x_1, x_2, \dots, x_n)$ соответственно $x_{i_1} = a_1, \dots, x_{i_r} = a_r$ и $x_{i_1} = \bar{a}_1, \dots, x_{i_r} = \bar{a}_r$. Если для произвольных $X \subset \{x_1, x_2, \dots, x_n\}$ и A выполняется одно из неравенств $f_{X,\bar{A}} \leq f_{X,A}$ или $f_{X,A} \leq f_{X,\bar{A}}$, то f называется вполне смешанной монотонной функцией. По определению, если функция является вполне смешанной монотонной, то она является смешанной монотонной функцией (см. § 7.1). Покажите, что: 1) пороговые функции являются вполне смешанными монотонными функциями; 2) монотонная функция $f = x_1 x_2 \vee x_3 x_4$ пороговой не является.

7.4. k -пороговая функция f может быть представлена в виде $f = g_1 g_2 \vee g_3 g_4 \vee \dots \vee g_{k-1} g_k$, если k — четное, и $g_1 g_2 \vee g_3 g_4 \vee \dots \vee g_{k-2} g_{k-1} \vee g_k$, если k — нечетное (g_1, g_2, \dots, g_k — пороговые функции).

7.5. Пусть $f \in \mathcal{B}_n$. Функция $f^{sd}(x_1, x_2, \dots, x_n, y) = yf(x_1, x_2, \dots, x_n) \vee \bar{y}f^d(x_1, x_2, \dots, x_n)$ является самодвойственной.

7.6. Предположим, что функции $f, g \in \mathcal{B}_n$ являются самодвойственными. Если $f_{x_1=1} = g_{x_1=1}$, то $f = g$.

7.7. Если f — это многочлен степени r ($0 \leq r \leq n$) от n переменных, не равный тождественно нулю, то $|D_1(f)| = 2^{n-r}$. Существуют многочлены, для которых последнее неравенство переходит в равенство.

7.8. Пусть $f(x_1, x_2, \dots, x_n)$ — многочлен и $c_{1,2,\dots,n}$ — его коэффициент при произведении x_1, x_2, \dots, x_n .

Тогда

$$\sum_{(a_1, a_2, \dots, a_n) \in V_n} \oplus f(a_1, a_2, \dots, a_n) = c_{1,2,\dots,n},$$

где $\sum \oplus$ означает суммирование в смысле операции \oplus .

Глава 8

Применение теории булевых функций

8.1. Схемы из функциональных элементов

Булевы функции могут служить для алгебраического описания связи входов с выходами в рассматриваемых далее схемах. При анализе цифровых устройств удобно считать, что имеющаяся информация задается с помощью двух состояний некоторого числа физических объектов. При этом каждый объект может находиться в одном из этих двух состояний, если не рассматривается переходный режим (состояния мы будем обозначать через 0 и 1).

Определение схемы из функциональных элементов. Рассмотрим схему D с n входами и m выходами, изображенную на рис. 8.1.

Предположим, что в некоторый момент времени t каждый вход и каждый выход схемы D находятся в одном из состояний 0 или 1. Схема D называется схемой из функциональных элементов с n входами и m выходами, если:

1) существует некоторое положительное число δ такое, что в момент $t + \delta$ каждый из выходов схемы также оказывается в некотором фиксированном состоянии;

2) эти фиксированные состояния выходов зависят только от совокупности состояний входов в момент t . Число δ называется временной задержкой схемы.

Схемы, удовлетворяющие условию (1), называются *устойчивыми*; если же это условие нарушается, то схемы называются *неустойчивыми*. Устойчивые цифровые схемы называют также *логическими* схемами.

Приведенное условие (2) означает, что схема «помнит» состояния, в которых находились входы до момента t , только в течение интервала времени длительностью δ . В схемах, не удовлетворяющих условию (2), прежние состояния входов оказывают то или иное влияние на текущие состояния выходов, т. е. схема обладает в некотором смысле «памятью». Такие схемы называются *последовательными* (см. § 10.3).

Функции, реализуемые схемами из функциональных элементов. Предположим, что для схемы из функциональных элементов D с n входами и m выходами определены булевы функции f_1, f_2, \dots, f_m от n переменных, удовлетворяющие следующим условиям. Если x_1, x_2, \dots, x_n — это величины, представляющие собой фиксированные на время δ или более состояния входов (входные переменные), а y_1, y_2, \dots, y_m — величины, представляющие состояния выходов схемы, определяемые условиями (1) и (2), то $y_i = f_i(x_1, x_2, \dots, x_n)$, $1 \leq i \leq m$. Такие функции f_1, f_2, \dots, f_m называются *выходными функциями* схемы D . При этом также говорят, что схема D реализует (с задержкой δ) совокупность функций f_1, f_2, \dots, f_m . Построением схем, реализующих заданные булевы функции, из стандартных выпускаемых промышленностью схем (интегральных схем определенного типа) с известными параметрами (например, известной временной задержкой δ и др.), а также проектированием стандартных схем на уровне «вентилей» (см. гл. 9) занимается инженерная дисциплина, называемая схемотехникой.

В действительности при проектировании схемы с n входами, как правило, не требуется, чтобы выходные функции были определены на каждом из 2^n возможных наборов состояний входов. Рассмотрим, например, цифровое устройство, в котором цифры от 0 до 9 представляются четырехразрядным двоичным кодом 8—4—2—1 (ДКДЦ — двоичный код десятичных цифр). На вход устройства, которое вырабатывает сигнал управления для формирования по двоичному представлению соответствующего ему де-

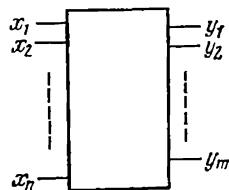


Рис. 8.1. Цифровая схема с n входами и m выходами

сятичного числа (например, для его индикации), при отсутствии ошибок не поступают двоичные представления цифр от 10 до 15. При этом часто не определяются и соответствующие им выходные значения. Такие совокупности состояний входов, на которых выходные значения не определяются, называются запрещенными, избыточными или несущественными наборами. Подбирая определенным образом выходные значения для таких запрещенных наборов, можно упростить схему. Эта идея лежит в основе решения одной из важных задач схемотехники, рассматривающейся в гл. 9.

Базисные схемы. На практике широко используют простые схемы, реализующие такие логические функции, как отрицание N , \max_n (ИЛИ) от n переменных, \min_n (И) от n переменных, ИЛИ—НЕ от n переменных, И—НЕ от n переменных*. Максимальное значение параметра n определяется технологией изготовления схем, но в области допустимых значений сложность схемы почти не зависит от n . Схем, реализующих функцию четности p_n и имеющих ту же сложность, что и схемы для реализации переносных функций, пока не существует. Такие схемы имеют сложность, почти пропорциональную n (упражнение 8.1). На рис. 8.2 приведены графические обозначения, краткие названия и реализуемые функции для всех упомянутых схем.

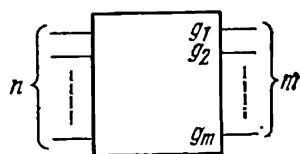
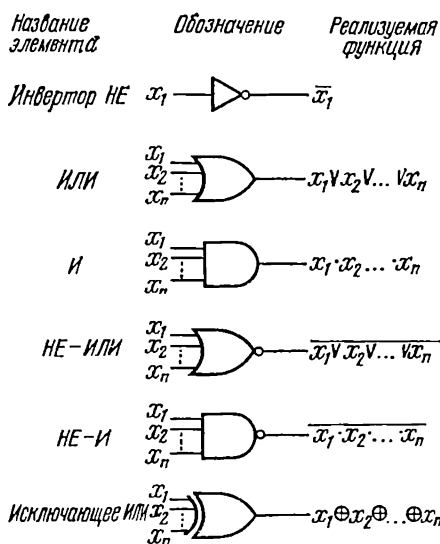


Рис. 8.3. Символическое представление схемы

Рис. 8.2. Графические обозначения и краткие названия базисных функций

Интегральные схемы подразделяются обычно на схемы малой степени интеграции (до 20 вентилей), схемы средней степени интеграции (20—100 вентилей) и большие интегральные схемы (свыше 100 вентилей). Приведенные в скобках цифры, характеризую-

* Эти и другие аналогичные им по сложности схемы называются далее вентилями.

щие степень интеграции схем, изменяются по мере развития технологии изготовления интегральных схем и поэтому являются не более чем временными критериями [20].

Задача проектирования устройств из функциональных элементов. При разработке обычно считается известным:

1) можно ли использовать в качестве внешних входов кроме входных переменных x_1, x_2, \dots, x_n также константы, отрицания входных переменных $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$ и другие функции;

2) типы стандартных схем, которые можно использовать в качестве компонент (совокупности реализуемых ими функций и другие параметры);

3) совокупности выходных функций, которые должны быть реализованы и предназначены для внешних устройств;

4) ограничения на временную задержку и другие системные параметры устройства.

Задача разработчика состоит в том, чтобы указать способ соединения внешних входов и выходов некоторого числа стандартных схем, указанных в п. 2), который позволил бы получить наиболее разумную с экономической точки зрения систему, реализующую указанные в п. 3) функции и удовлетворяющую ограничениям п. 4).

Что касается соединения внешних входов устройств, то обычно имеются следующие ограничения:

а) к каждому внешнему входу и выходу компоненты может подключаться лишь некоторое ограниченное число входов компонента*;

б) каждый внешний выход или вход компоненты может подключаться только к одному внешнему входу или к одному выходу некоторой другой компоненты схемы**.

Для схемы (сети) N граф соединений $C(N)$ определяется следующим образом. Каждому внешнему входу, внешнему выходу и каждой компоненте системы сопоставляется взаимно однозначно вершина графа (каждой вершине графа присваивается имя соответствующего ей внешнего входа, выхода или компоненты). Из вершины v_1 графа проводится ориентированное ребро в вершину v_2 , если вершине v_1 соответствует внешний вход или выход компоненты, соединенные с внешним выходом или входом, или компоненты системы, соответствующими вершине v_2 .

При построении устройств из функциональных элементов обычно имеется также следующее ограничение:

в) каждая компонента является схемой из функциональных элементов, и граф соединений $C(N)$ не содержит ориентированных циклов (наличие этого ограничения будет предполагаться и в гл. 8, 9 данной книги). Для каждой вершины v графа $C(N)$ (или соответствующей ей компоненты схемы) определим ее глубину.

* Кратность ответвления выхода определяется технологией изготовления схем; это ограничение известно как ограничение ветвления выхода.

** Можно соединять выходы схем типа И и ИЛИ, но функционально такие случаи следует рассматривать как операции, выполняемые элементами И и ИЛИ.

1) Если в вершину v не входит ни одного ориентированного ребра, то положим $l(v)=0$. Вершинам, соответствующим внешним входам, также припишем глубину 0.

2) Предположим, что определены все вершины глубины $\leq i$ и в графе еще имеются вершины, глубина которых не определена. Среди этих вершин выделим те, для которых все входящие в них ориентированные ребра начинаются в вершинах, глубина которых уже определена. Эти вершины отнесем к глубине $i+1$. Если допустить, что вершин, удовлетворяющих этому условию в графе нет, то это будет означать, что в каждую из вершин с неопределенной глубиной входит, по крайней мере, одно ребро, также начинающееся в вершине с неопределенной глубиной, т. е. в графе существует ориентированный цикл. Это противоречит предположению 2). Следовательно, описанная процедура через конечное число шагов закончится, причем для каждой вершины глубина будет определена однозначно.

Максимальная глубина вершин, соответствующих компонентам схемы, называется глубиной схемы N . Если схема N задана, то, как показывают приведенные далее примеры, сравнительно нетрудно определить реализуемые ею выходные функции.

Пример 8.1. Рассмотрим изображенную на рис. 8.3 систему с n входами и m выходами. Покажем, что функции g_1, g_2, \dots, g_m , реализуемые этой системой, являются функциями (выходных) переменных, соответствующих входам системы, перенумерованным сверху вниз целыми числами $1, 2, \dots, n$ (функция g_i не обязательно зависит от всех входных переменных; в данном примере переменные, от которых зависят функции, не указываются). Определим выходные функции системы, изображенной на рис. 8.4. Так

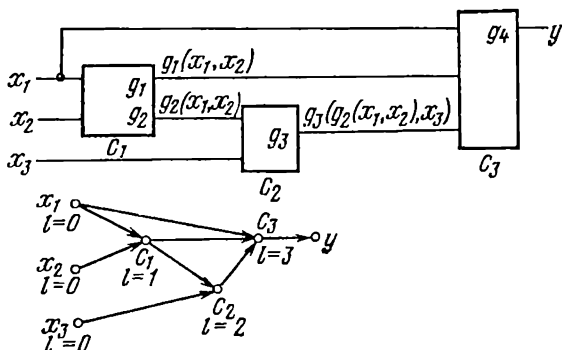


Рис. 8.4. Схема и граф соединений из примера 8.1

как входами компонент глубины 1 являются внешние входы, то их выходные функции определены. Предположим, что выходные функции всех компонент системы глубины до i включительно уже заданы как функции входных переменных. Так как входами ком-

понсит глубины $i+1$ являются выходы компонент глубины i и не-
 нсе, или внешние входы, то выходные функции всех компонент
 глубины $i+1$ также можно выразить в виде функций входных пе-
 ременных. Так, выходная функция, реализуемая системой, приве-
 денной на рис. 8.4, является следующей суперпозицией функций,
 реализуемых компонентами системы: $g_4(x_1, g_1(x_1, x_2), g_3(g_2(x_1, x_2),$
 $x_3))$.

З а д а ч а 8.1. Система, удовлетворяющая условию в), является
 схемой из функциональных элементов. Если ее глубина l , а макси-
 мальная временная задержка ее компонент равна δ_m , то времен-
 ная задержка всей системы не превосходит $l\delta_m$.

Указание. Следует воспользоваться индукцией по глубине l и определением
 глубины.

З а д а ч а 8.2. На рис. 8.5 приведен пример средней интеграль-
 ной схемы [26] со входами I_1, I_2, \dots, I_9 . Найдите выходные функции,
 реализуемые на внешних выходах и E . Эта схема используется для
 обнаружения однопочных ошибок, как это будет описано в приме-
 ре 8.6, и ряда других целей.

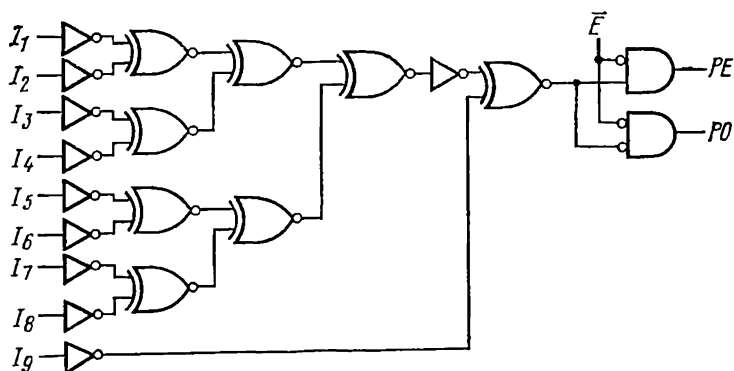


Рис. 8.5. Схема вычисления символа четности с девятью вхо-
 дами (TTL/MSI 93562)

Указание. При $E=0$ обе выходные функции PE и PO равны 0. Если $E=1$
 на выходе PE реализуется функция четности p_9 , а на выходе PO — ее отрица-
 ние \bar{p}_9 . Включение в схему некоторого числа инверторов связано с удобством
 технологии их изготовления.

У п р а ж н е н и е 8.1. Ориентированный граф, который получа-
 ется из ориентированного дерева (§ 5.3) изменением ориентации
 всех его ребер на противоположную, называется далее r -ориен-
 тированным деревом. Если граф соединений схемы является r -ориен-
 тированным деревом, то такая схема называется *древовидной*.
 Рассмотрим древовидную схему N , каждая компонента которой
 имеет m входов и один выход. В r -ориентированном дереве $C'(N)$
 (которое получается из $C(N)$ исключением вершины, отвечающей
 внешнему выходу и входящего в нее ребра) из каждой вершины,

отличной от концевой точки (корня), выходит ровно одно ребро и в каждую вершину, отличную от концевой точки (соответствующей внешнему входу), входит m ребер. Такое r -ориентированное дерево называется m -деревом. Покажите, что:

1) если в m -дереве имеется n концевых точек (n равно числу внешних входов), то число n' остальных вершин (число компонент в схеме) равно $n' = (n-1)/(m-1)$;

2) схема N имеет глубину не менее чем $\lceil \log_m n \rceil$ *;

3) существует древовидная схема из $n-1$ элементов исключающие ИЛИ с двумя входами, имеющая глубину $\lceil \log_2 n \rceil$ и реализующая функцию четности p_n .

Решение. (1) Число ребер $|E|$ в m -дереве по теореме 4.11 равно $n+n'-1$. В то же время, согласно упражнению 5.1, $|E| = mn'$ (напомним, что ориентация ребер меняется на противоположную). Поэтому $n' = (n-1)/(m-1)$.

(2) Пусть l — глубина схемы N . Тогда, по предположению, найдется, по крайней мере, одна вершина глубины l . Число вершин глубины $l-i$ не может превосходить m^i . Следовательно, $n \leq m^l$, т. е. $l \geq \lceil \log_m n \rceil$.

(3) Пусть $2^{l-1} < n < 2^l$ (l — неотрицательное целое число). Воспользуемся индукцией по l . Поскольку $n=1$ при $l=0$, то $p_1(x_1) = x_1$. При этом входы можно отождествить с выходами, и число элементов исключающее ИЛИ с двумя входами равно 0. Предположим, что утверждение справедливо при $0 \leq l < j$. Рассмотрим случай $l=j$. Положим $n_1 = 2^{j-1}$, $n_2 = n - 2^{j-1}$. По предположению индукции, существует древовидная схема $N_1(N_2)$ из n_1-1 (n_2-1) элементов исключающее ИЛИ с двумя входами, имеющая глубину $j-1$ (не более $j-1$) и реализующая функцию $p_{n_1}(x_1, \dots, x_{n_1})$ ($p_{n_2}(x_{n_1+1}, \dots, x_n)$). Пусть G — это двухходовый элемент исключающее ИЛИ и пусть N -схема, получающаяся в результате подключения выходов схем N_1 и N_2 ко входам элемента G . Схема N является древовидной, имеет глубину j и состоит из $n-1$ элементов исключающее ИЛИ с двумя входами. Ее внешней выходной функцией, согласно (7.17), является $p_{n_1}(x_1, x_2, \dots, x_n) \oplus p_{n_2}(x_{n_1+1}, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$. Если $n \geq 2$, то $C'(N)$ является 2-деревом.

Задача 8.3. Докажите п. (3) упражнения 8.1, если каждая компонента реализует функцию \bar{p}_2 , а выходной функцией является p_n или \bar{p}_n .

Указание. Следует воспользоваться индукцией по глубине схемы.

Пример 8.2. Комбинационная схема, реализующая все операции по вычислению одного разряда суммы двух двоичных чисел, называется *полным сумматором* и обозначается через FA . Пусть A, B, C — три входа и S, C' — выходы, определяемые соотношениями $S = A \oplus B \oplus C$, $C' = \text{Maj}_3(A, B, C)$. Комбинационная схема с выходами S и C , которые определяются соотношениями $S = A \oplus B$, $C = A \cdot B$, где A, B — два входа, называется *полусумматором* и обозначается через HA . Если вход C схемы FA фиксировать, положив его равным 0, то получится схема HA .

Задача 8.4. На рис. 8.6 приведена схема (на уровне вентилей) полного сумматора, осуществляющая сложение двоичных четырехразрядных чисел и реализующая описанную в упражнении 7.5 процедуру вычислений. Эта схема является типичным при-

* $[a]$ — наименьшее целое число, не меньшее действительного числа a .

мером схемы средней степени интеграции [27]. На этом рисунке A_i , B_i — входы схемы, соответствующие i -м разрядам a_i и b_i складываемых чисел; C_0 — вход, соответствующий переносу из младших разрядов; S_i ($1 \leq i \leq 3$) и C_4 — выходы схемы.

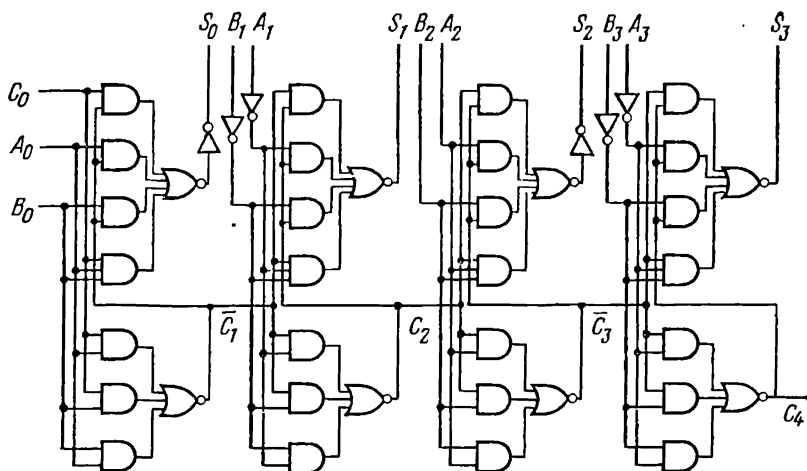


Рис. 8.6. 4-разрядный двоичный полный сумматор

Пусть $\bar{C}_1, C_2, \bar{C}_3$ — выходы трех внутренних элементов НЕ-ИЛИ-схемы, как это показано на рисунке. Вначале следует по схеме проверить следующие соотношения:

- (1) $\bar{C}_1 = \overline{C_{i-1} \cdot A_{i-1} \vee C_{i-1} \cdot B_{i-1} \vee A_{i-1} \cdot B_{i-1}}, \quad i = 1, 3;$
- (2) $C_i = \overline{\bar{C}_{i-1} \cdot \bar{A}_{i-1} \vee \bar{C}_{i-1} \cdot \bar{B}_{i-1} \vee \bar{A}_{i-1} \cdot \bar{B}_{i-1}}, \quad i = 2, 4;$
- (3) $S_i = (A_i \vee B_i \vee C_i) \cdot \bar{C}_{i+1} \vee A_i \cdot B_i \cdot C_i, \quad i = 0, 2;$
- (4) $S_i = (\bar{A}_i \vee \bar{B}_i \vee \bar{C}_i) \cdot C_{i+1} \vee A_i \cdot B_i \cdot \bar{C}_i, \quad i = 1, 3.$

После этого показать, что

- (а) $C_i = \text{Maj}_3(A_{i-1}, B_{i-1}, C_{i-1}), \quad 0 \leq i \leq 4;$
- (б) $S_i = A_i \oplus B_i \oplus C_i, \quad 0 \leq i \leq 3.$

Решение. (а) Поскольку функция Maj_3 является самодвойственной (задача 7.46), то соотношение (а) выполняется. (б) Используя (а), в (3) получаем, что

$$S_i = (A_i \vee B_i \vee C_i) \cdot \overline{\text{Maj}_3(A_i, B_i, C_i)} \vee A_i \cdot B_i \cdot C_i.$$

Выражение $A_i \vee B_i \vee C_i$ принимает значение 1, если равна 1 хотя бы одна из величин A_i, B_i, C_i . Значение $\text{Maj}_3(A_i, B_i, C_i)$ равно 1, если среди величин A_i, B_i, C_i не более чем одна принимает значение 1. Поэтому значение $(A_i \vee B_i \vee C_i) \cdot \overline{\text{Maj}_3(A_i, B_i, C_i)}$ равно 1 только в том случае, если ровно одно из значений A_i, B_i, C_i равно 1. Следовательно, правая часть приведенной формулы равна $A_i \oplus B_i \oplus C_i$. Для того чтобы доказать (4), следует воспользоваться самодвойственностью функции p_3 (задача 7.45).

Полный $4m$ -разрядный сумматор можно построить, например, из m схем средней степени интеграции, если их соединить так, как показано на рис. 8.7. Недостаток такой схемы — большое время прохождения сигналов переноса от младших разрядов к старшим. Известно несколько способов уменьшения задержки, связанной с прохождением сигнала переноса [20].

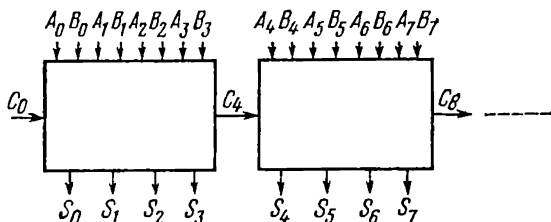


Рис. 8.7. $4m$ -разрядный двоичный полный сумматор

Пример 8.3. Если функция f представлена в виде суперпозиции функций g_1, g_2, \dots, g_l , то соединением компонент, реализующих функции g_1, g_2, \dots, g_l , можно построить схему, реализующую функцию f (утверждение, обратное 8.1).

Например, пусть $f(x_1, x_2, x_3) = g_1(x_1, g_2(x_1, x_2), g_1(g_2(x_1, x_2), x_3, x_2))$ и G_1, G_2 — компоненты, реализующие функции $g_1 \in \mathcal{B}_3$ и $g_2 \in \mathcal{B}_2$ соответственно. Схемой, выход которой будет непосредственно выходом искомой схемы, является G_1 . Входы этой схемы: внешний вход x_1 , выход схемы G_2 и выход второй схемы G_1 . Входами схемы G_2 являются внешние входы x_1 и x_2 , а входами второй схемы G_1 — выход схемы G_2 и внешние входы x_2 и x_3 . Получающаяся в результате схема, реализующая функцию f , показана на рис. 8.8.

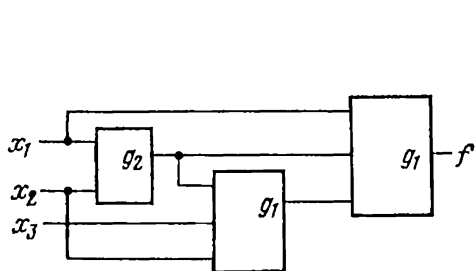


Рис. 8.8. Схема из упражнения 8.3

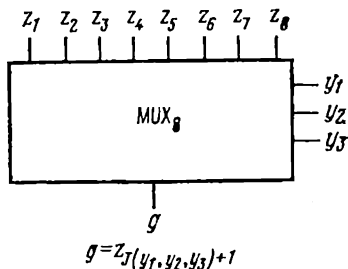


Рис. 8.9. Мультиплексор на восемь входов

Как видно из изложенного, проблема построения комбинационных схем состоит в представлении заданной функции f в виде суперпозиции функций из заданного набора g_1, g_2, \dots, g_l . Вообще говоря, функция может быть представлена в виде суперпозиции заданных функций многими различными способами (в дальнейшем мы будем иметь возможность это неоднократно наблюдать), а поэтому важной задачей является выбор представления, которое

дает в том или ином смысле «оптимальную» реализацию. В дальнейшем мы должны будем определить понятие оптимальности формально, но сразу же оговоримся, что эта задача всегда связана с текущим уровнем развития технологии и видоизменяется по мере ее развития. Критерием оптимальности не может стать простая сумма стоимостей компонент, так как она не учитывает стоимость платы, необходимой для соединения нескольких модулей, стоимость соединения плат и ряд других факторов. Обычно использование модулей, выполняющих сложные функции (с высокой степенью интеграции), позволяет уменьшить общее число модулей в устройстве, и это считается желательным.

Другая сложная проблема — выбор самой совокупности стандартных схем (базиса). Эта совокупность, в частности, должна позволять реализовать произвольную булеву функцию в виде суперпозиции функций, принадлежащих множеству $G = (g_1, g_2, \dots, g_l)$ функций, реализуемых компонентами базиса; т. е. множество G должно быть полным [см. § 8.3]. Поскольку стоимость интегральной схемы почти обратно пропорциональна спросу на нее, то стандартные схемы должны позволять: а) просто реализовывать широко используемые функции путем выбора в качестве их входов констант 0, 1, переменных, отрицаний переменных и т. д. и б) реализовывать после относительно простой предварительной обработки булевой функции, необходимые пользователю в данный момент, подобно тому, как это делают программируемые логические матрицы (см. пример 9.2).

Задача 8.5. Комбинационная схема с $m+2^m$ входами $y_1, y_2, \dots, y_m, z_1, z_2, \dots, z_{2^m}$ и одним выходом g , которая при $y_i = b_i (1 \leq i \leq m)$ выбирает из z_1, z_2, \dots, z_{2^m} вход $z_{J(b_1, b_2, \dots, b_m)+1}$ и подает его значение на выход в качестве выходной функции g , называется *мультиплексором (селектором) 2^m каналов* и обозначается MUX_{2^m} (рис. 8.9). Например, при $m=3$ и $y_1=1, y_2=0, y_3=1$ выход $g = z_{J(1, 0, 1)+1} = z_6$.

1) Покажите, что произвольную функцию от трех переменных можно реализовать с помощью одного элемента MUX_8 и постоянных входов 0 и 1.

2) Используя этот результат, покажите, что произвольную логическую функцию от четырех переменных можно реализовать с помощью одного элемента MUX_8 , одного инвертора и постоянных входов 0 и 1.

3) Если в качестве входов разрешается использовать только x_1, x_2, x_3 , а постоянные входы использовать нельзя, то с помощью MUX_8 нельзя реализовать логическую функцию, для которой $f(0, 0, 0) = 1$ и $f(1, 1, 1) = 0$. Однако можно реализовать любую функцию от трех переменных, для которой $f(0, 0, 0) = 0, f(1, 1, 1) = 1$. Если допустить возможность использования инверторов, то можно реализовать все функции от трех переменных.

Указание. Для доказательства утверждения (1) следует положить $z_{J(a_1, a_2, a_3)+1} = f(a_1, a_2, a_3), (a_1, a_2, a_3) \in V_3$.

Из задачи 7.49 для доказательства (2) имеем

$$f(x_1, x_2, x_3, x_4) = \bigvee_{(a_1, a_2, a_3) \in V_3} f(a_1, a_2, a_3, x_4) x_1^{a_1} x_2^{a_2} x_3^{a_3}.$$

Далее следует положить $y_1 = x_1$, $y_2 = x_2$, $y_3 = x_3$ и заметить, что каждая из функций $f(a_1, a_2, a_3, x_4)$ представляет собой 0, 1, x_4 или \bar{x}_4 .

(3) Каждый из входов совпадает с одной из переменных x_1, x_2, x_3 . Если $x_1 = x_2 = x_3 = 0$, то все входы элемента MUX_8 будут равны 0. Следовательно, $g = z_1 = x_1 = 0$. Если $x_1 = x_2 = x_3 = 1$, то $g = z_8 = x_1 = 1$. Для функции, для которой $f(0, 0, 0) = 0$ и $f(1, 1, 1) = 1$, следует положить $y_1 = x_1$, $y_2 = x_2$, $y_3 = x_3$. Для каждого набора $(a_1, a_2, a_3) \in V_3$ выберем одну из его компонент a_i так, что $f(a_1, a_2, a_3) = a_i$. При этом $z_{f(a_1, a_2, a_3)+1} = x_i$. Вторая половина утверждения столь же очевидна.

Задача 8.6. Если в качестве входов разрешается использовать постоянные, то с помощью девяти элементов MUX_8 можно реализовать любую функцию от шести переменных.

Указание. Из задачи 7.49 следует, что для произвольной функции $f \in \mathcal{B}_6$

$$f(x_1, x_2, \dots, x_6) = \bigvee_{(a_1, a_2, a_3) \in V_3} f(a_1, a_2, a_3, x_4, x_5, x_6) x_1^{a_1} x_2^{a_2} x_3^{a_3}.$$

Согласно решению задачи 8.5, функцию $f(a_1, a_2, a_3, x_4, x_5, x_6)$ ($(a_1, a_2, a_3) \in V_3$) можно реализовать с помощью восьми элементов MUX_8 . Если выходы этих элементов подключить ко входам z_1, z_2, \dots, z_8 еще одного элемента MUX_8 , то, положив $y_i = x_i$ ($1 \leq i \leq 3$) и основываясь на приведенной формуле, можно реализовать функцию $f(x_1, x_2, \dots, x_6)$ (рис. 8.10).

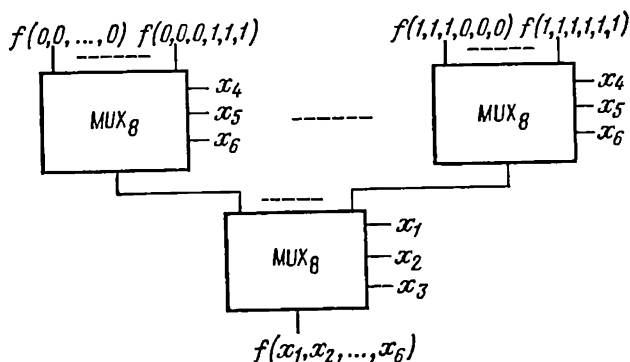


Рис. 8.10. Представление произвольной функции $f \in \mathcal{B}_6$ с помощью MUX_8

Задача 8.7. Покажите, что если в предыдущей задаче разрешить кроме постоянных входов использовать в качестве внешнего входа также \bar{x}_7 , то с помощью девяти элементов MUX_8 можно реализовать произвольную функцию от семи переменных $f(x_1, x_2, \dots, x_7)$.

Указание. Пользуясь утверждением (2) задачи 8.5, покажите, что с помощью восьми элементов MUX_8 первого каскада схемы из предыдущей задачи можно реализовать $f(a_1, a_2, a_3, x_4, x_5, x_6, x_7)$.

Задача 8.8. Обобщите задачи 8.5—8.7 на случай произвольных элементов MUX_{2m} .

Задача 8.9. Найдите дизъюнктивную нормальную форму, представляющую выходную функцию g элемента MUX_{2m} .

Решение. Используя разложение (1) из задачи 7.49, получаем

$$g = \bigvee_{(b_1, b_2, \dots, b_m) \in V_m} z_{j(b_1, b_2, \dots, b_m)+1} y_1^{b_1} y_2^{b_2} \dots y_m^{b_m}.$$

8.2. Суперпозиции и полные системы функций [13,28]

До настоящего времени понятие суперпозиции функций формально не было определено и использовалось в предположении, что оно интуитивно понятно. Однако здесь для рассмотрения полных систем функций необходимо его определить строго.

Пусть $g \in \mathcal{B}_m$, $g_i \in \mathcal{B}_n \cup \{x_1, x_2, \dots, x_n\}$, $1 \leq i \leq m$.

Через

$$g(g_1, g_2, \dots, g_m) \quad (8.1)$$

будем обозначать функцию из \mathcal{B}_n , которая на наборе $A \in V_n$ принимает значение $g(g_1(A), g_2(A), \dots, g_m(A))$. В случае, когда g_i является переменной x_j , будем рассматривать x_j как функцию на V_n в соответствии с соглашением в § 7.3, т. е. считать, что $g_i(A)$ представляет j -ю компоненту A . Из соотношений (7.15) — (7.17) и упражнения 7.4 для $g_i \in \mathcal{B}_n$ ($1 \leq i \leq m$) получаем

$$\max_m(g_1, g_2, \dots, g_m) = g_1 \vee g_2 \vee \dots \vee g_m; \quad (8.2)$$

$$\min_m(g_1, g_2, \dots, g_m) = g_1 \cdot g_2 \cdot \dots \cdot g_m; \quad (8.3)$$

$$p_m(g_1, g_2, \dots, g_m) = g_1 \oplus g_2 \oplus \dots \oplus g_m, \quad (8.4)$$

т. е. операции \vee , \cdot , \oplus на множестве функций задают представление соответственно функций \max_2 , \min_2 и p_2 . Выражение (8.1), рассматриваемое как последовательность из символов g_1, g_2, \dots, g_m , запятых и скобок, называется *формулой*. Для того чтобы ввести общее понятие формулы и функции, которую она представляет, удобно воспользоваться представлением формулы в виде ориентированного дерева.

Пусть \mathcal{L} — подмножество множества \mathcal{B} (не обязательно конечное). Пусть $f \in \mathcal{B}$. При рассмотрении f как символа, обозначающего функцию, будем называть его *именем функции*. До тех пор пока это не приводит к недоразумению, один и тот же символ далее используется и как название функции, и как функция, которую он представляет. При этом множество названий функций из \mathcal{L} также обозначается через \mathcal{L} (эти соглашения относятся также к переменным и их названиям). Положим $VX = \{x_i | i = 1, 2, \dots\}$ и $VX_m = \{x_1, x_2, \dots, x_m\}$. Если функция $f \in \mathcal{B}_n$ существенно зависит от n переменных, то положим, по определению, $v(f) = n$. Условимся считать, что $v(0) = v(1) = 0$ для констант 0 и

1. Конечное ориентированное дерево T называется \mathcal{LX} -деревом*, если оно удовлетворяет двум условиям:

1) именем концевой вершины является либо имя переменной, принадлежащей VX , либо имя константы g , содержащейся в \mathcal{L} ;

2) именем любой вершины v , отличной от конечной вершины, является имя функции из \mathcal{L} . Из вершин v с именем g выходит $v(g)$ ребер; номера этих ребер — целые числа $1, 2, \dots, v(g)$.

Если имя корня \mathcal{LX} -дерева T (конечной вершины T) принадлежит \mathcal{L} , то такое дерево называют также \mathcal{L} -деревом.

На рис. 8.11 приведен пример \mathcal{L} -дерева. В данном случае $\mathcal{L} = \{0, g_1, g_2, g_3\}$, $v(g_1) = 2$, $v(g_2) = 3$, $v(g_3) = 1$.

Вершины \mathcal{LX} -дерева T , именами которых являются символы множества \mathcal{L} , называются \mathcal{L} -точками; при этом вершины, именами которых являются символы VX , называются X -точками. Максимальное значение длины ориентированного пути, идущего из корня в конечную точку дерева T , обозначим через $L(T)$ и назовем глубиной T . Если дерево T состоит лишь из одного корня, то положим $L(T) = 0$. Максимальное значение индекса имени X -точки дерева T обозначим через $n(T)$ (если дерево не содержит X -точек, то положим $n(T) = 1$). Для дерева, изображенного на рис. 8.11, $L(T) = 3$, $n(T) = 6$.

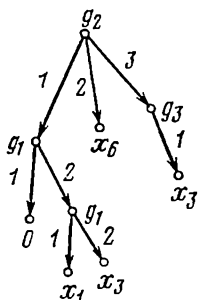


Рис. 8.11. Пример \mathcal{L} -дерева

Задача 8.10. Рассмотрим \mathcal{L} -дерево T .

1) Ориентированное поддерево T' дерева T , корнем которого является вершина v , отличная от корня дерева T , является \mathcal{LX} -деревом; если v является \mathcal{L} -вершиной, то T' — \mathcal{L} -деревом и $L(T') < L(T)$.

2) Ориентированное дерево, которое получается при замене T' любым другим \mathcal{LX} -деревом, является \mathcal{L} -деревом.

Указание. Утверждения следуют непосредственно из определений \mathcal{L} -дерева, \mathcal{LX} -дерева и ориентированного поддерева (§ 5.3).

Определения формулы и функции, представляемых \mathcal{LX} -деревом. Формулой $\varphi(T)$, представляемой \mathcal{LX} -деревом T , является последовательность символов, которыми могут быть имена функций из \mathcal{L} , имена переменных из VX , запятые и скобки. Формула $\varphi(T)$ представляет некоторую функцию из \mathcal{B}_n для произвольного целого числа $n \geq n(T)$. Эта функция $\rho_n(\varphi(T))$ (кратко обозначаемая через $\rho_n(T)$, индекс n может опускаться, если он понятен из контекста) определяется по индукции следующим образом:

1) Если $L(T) = 0$, то имя корня l представляет собой либо одну из констант $0, 1$, либо одну из переменных x_i . По определе-

* Для удобства граф, состоящий из одной вершины и не содержащий ребер, будет также называться ориентированным деревом. Эта вершина является корнем и концевой вершиной дерева одновременно.

нию, $\varphi(T)=l$, $\rho_n(T)=l$. В последнем равенстве константа или переменная l рассматривается как функция (§ 7.3).

2) Пусть $1 \leq L(T)=k$. Предположим, что φ , ρ определены до глубины $k-1$ включительно. Пусть g — имя корня и T_i — ориентированное поддерево, корнем которого является концевая точка ориентированного ребра, выходящего из корня и имеющего имя i ($1 \leq i \leq v(g)$). Тогда

$$\varphi(T) = g(\varphi(T_1), \varphi(T_2), \dots, \varphi(T_{v(g)})); \quad (8.5)$$

$$\rho_n(T) = g(\rho_n(T_1), \rho_n(T_2), \dots, \rho_n(T_{v(g)})). \quad (8.6)$$

Формула, которую представляет \mathcal{L} -дерево, изображенное на рис. 8.11, имеет вид $g_2(g_1(0, g_1(x_1, x_3)), x_6, g_3(x_3))$.

Для \mathcal{L} -дерева T формула $\varphi(T)$ называется \mathcal{L} -формулой; по определению,

$$\tilde{\mathcal{L}} = \left\{ f \mid \begin{array}{l} \text{существуют } n \text{ и } \mathcal{L}\text{-дерево } T \\ \text{такие, что } f = \rho_n(T) \end{array} \right\}. \quad (8.7)$$

Если $f = \rho_n(T)$, то говорят, что f может быть получена как суперпозиция функций из \mathcal{L} глубины $L(T)$.

Задача 8.11. 1) $\mathcal{L} \subset \tilde{\mathcal{L}}$. 2) Если $\mathcal{L}_1 \subset \mathcal{L}_2$, то $\mathcal{L}_1 \subset \mathcal{L}_2$.

Указание. Решение следует непосредственно из определений.

Приведенное определение ρ_n дает способ нахождения значения функции, представляемой \mathcal{L} -деревом T , на произвольном наборе $(a_1, a_2, \dots, a_n) \in V_n$.

Этот способ заключается в следующем:

(1) X -вершине T с именем x_i сопоставляется значение a_i .

(2) \mathcal{L} -вершинам, именами которых являются константы, сопоставляются значения соответствующих констант.

(3) Пусть v — произвольная \mathcal{L} -вершина с именем g . Если конечной точке v_i ориентированного ребра с именем i , выходящего из v ($1 \leq i \leq v(g)$), сопоставлено значение b_i , то вершине v при этом сопоставляется значение $g(b_1, b_2, \dots, b_{v(g)})$.

(4) Описанные выше операции повторяются до тех пор, пока не будет достигнут корень дерева. Значение, сопоставляемое корню дерева, является искомым значением функции.

Задача 8.12. Пусть $f = \rho_n(T)$ и T является \mathcal{L} -деревом. Тогда функцию f можно реализовать с помощью схемы N глубины $L(T)$, компонентами которой являются схемы, реализующие функции с именами \mathcal{L} -вершин, отличных от конечных вершин T .

Указание. Следует воспользоваться индукцией по $L(T)$. Для \mathcal{L} -дерева, изображенного на рис. 8.11, соответствующая схема показана на рис. 8.12.

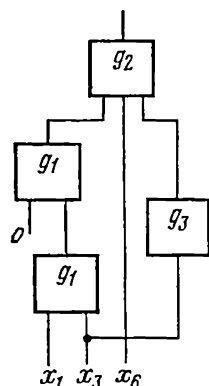


Рис. 8.12. Схема, реализующая функцию, представляемую \mathcal{L} -деревом, изображенным на рис. 8.11

Упражнение 8.2. Если $g \in \tilde{\mathcal{L}} \cap \mathcal{B}_m$, $g_i \in (\tilde{\mathcal{L}} \cap \mathcal{B}_n) \cup \mathcal{V}X_n$ ($1 \leq i \leq m$), то $f = g(g_1, g_2, \dots, g_m) \in \tilde{\mathcal{L}}$.

Решение. По предположению, существуют \mathcal{L} -дерево T_0 такое, что $g = \rho_m(T_0)$, и $\mathcal{L}X$ -дерево T'_i такое, что $g_i = \rho_m(T'_i)$. Если $L(T_0) = 0$, то g является константой и $f = g \in \tilde{\mathcal{L}}$. Предположим, что $L(T_0) \geq 1$. Пусть g_0 — имя корня T_0 , $q = v(g_0)$ и T_{0h} — ориентированное поддерево T_0 , корнем которого является концевая точка v_h ориентированного ребра, выходящего из корня и имеющего имя h ($1 \leq h \leq g$). Рассмотрим ориентированные деревья T и T_{0h} , получающиеся соответственно из T_0 и T_{0h} в результате замены каждой их X -вершины с именем x_i на дерево T'_i (рис. 8.13). Индукцией по $L(T_0)$ покажем, что T является \mathcal{L} -де-

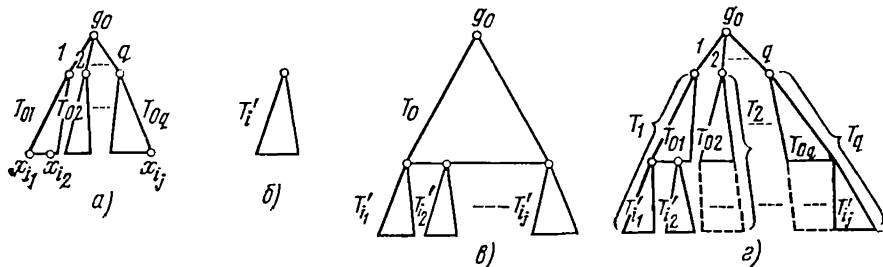


Рис. 8.13. Иллюстрации к упражнению 8.2:

a — ориентированные поддерева T_{0h} ($1 \leq h \leq q$) \mathcal{L} -дерева T_0 , $g_{0h} = \rho_m(T_{0h})$, $g = \rho_m(T_0)$; b — $\mathcal{L}X$ -дерево T'_i , $g_i = \rho_m(T'_i)$; c — \mathcal{L} -дерево T ; z — ориентированные поддерева T_h ($1 \leq h \leq q$) дерева T , $g_{0h}(g_1, g_2, \dots, g_m) = \rho_n(T_h)$

ревом и $f = \rho_n(T)$. Предположим, что доказываемое утверждение верно при $0 \leq L(T_0) < k$, и пусть $L(T_0) = k$. Если положить $g_{0h} = \rho_m(T_{0h})$, то из (8.6) будет следовать, что

$$g = g_0(g_{01}, g_{02}, \dots, g_{0q}). \quad (8.8)$$

Предположим, что T_h является \mathcal{L} -деревом. Так как $L(T_{0h}) < k$, то, по предположению индукции

$$g_{0h}(g_1, g_2, \dots, g_m) = \rho_n(T_h). \quad (8.9)$$

Если v_h является X -вершиной с именем x_i , т. е. если $g_{0h} = x_i$, то $T_h = T'_i$ и, следовательно,

$$g_i = \rho_n(T_h). \quad (8.10)$$

Из равенств $f = g(g_1, g_2, \dots, g_m)$, (8.8) — (8.10) получаем

$$f = g_0(\rho_n(T_1), \rho_n(T_2), \dots, \rho_n(T_q)).$$

Согласно (8.6), правая часть последнего равенства равна $\rho_n(T)$.

Определение полной системы. Если множество булевых функций \mathcal{L} таково, что $\tilde{\mathcal{L}} = \mathcal{B}$, то \mathcal{L} называется *полной системой*. Другими словами, множество булевых функций является полным, если произвольная булева функция может быть представлена в виде суперпозиции функций из \mathcal{L} .

Упражнение 8.3. Если \mathcal{L}_1 — полная система и $\mathcal{L}_1 \subset \tilde{\mathcal{L}}_2$, то \mathcal{L}_2 также является полной.

Решение. Для произвольной функции $f \in \mathcal{B}$, по предположению, существует такое \mathcal{L}_1 -дерево T , что $f = \rho_n(T)$. Пусть $n = v(f)$. Если $L(T) = 0$, то $f \in \mathcal{L}_1 \subset \tilde{\mathcal{L}}_2$.

Предположим, что $0 \leq L(T) = k$, $f \in \tilde{\mathcal{L}}_2$, и рассмотрим случай $L(T) = k$. Пусть g — имя корня, T_i — ориентированное поддерево дерева T , корень которого — концевая вершина v_i ориентированного ребра с именем i , выходящего из корня. Если v_i является \mathcal{L} -вершиной, то $L(T_i) < L(T)$ и $\rho_n(T_i) \in \tilde{\mathcal{L}}_2$. Если же v_i имеет имя x_j , то $\rho_n(T_i) = x_j$. Кроме того, из (8.6) имеем $f = g(\rho_n(T_1), \rho_n(T_2), \dots, \rho_n(T_{v(q)}))$. Так как $g \in \mathcal{L}_1 \subset \tilde{\mathcal{L}}_2$, то, согласно упражнению 8.2, $f \in \tilde{\mathcal{L}}_2$. Следовательно, $\mathcal{B} \subset \tilde{\mathcal{L}}_2$.

У п р а ж н е н и е 8.4. Покажите, что следующие системы функций являются полными: (1) $\mathcal{L}_A = \{N, \min_2\}$; (2) $\mathcal{L}_O = \{N, \max_2\}$; (3) $\mathcal{L}_{NA} = \{\text{НЕ} - \text{И от двух переменных}\}$; (4) $\mathcal{L}_{NO} = \{\text{НЕ} - \text{ИЛИ от двух переменных}\}$; (5) $\mathcal{L}_G = \{1, \min_2, p_2\}$.

Решение. Вначале докажем полноту системы $\mathcal{L}_L = \{N, \min_2, \max_2\}$. Поскольку $0 = \min_2(N(x_1), x_1)$, $1 = \max_2(N(x_1), x_1)$, $x_1 = \min_2(x_1, x_1)$, то предположим, что $\mathcal{B}_1 \subset \tilde{\mathcal{L}}_L$, $\mathcal{B}_{n-1} \subset \tilde{\mathcal{L}}_L$. Для произвольной функции $f \in \mathcal{B}_n$ из (7.25) имеем $f = \max_2(\min_2(N(x_1), f_{x_1=0}), \min_2(x_1, f_{x_1=1}))$. Поскольку $f_{x_1=1}, f_{x_1=0} \in \tilde{\mathcal{L}}_L$, $N \in \mathcal{L}_L$, то из упражнения (8.2) получаем $f \in \tilde{\mathcal{L}}_L$.

(1) Согласно правилу Де Моргана, $\max_2(x_1, x_2) = N(\min_2(N(x_1), N(x_2)))$. Из упражнения 8.2 получаем $\mathcal{L}_L \subset \tilde{\mathcal{L}}_A$. Отсюда и упражнения 8.3 следует полнота системы функций (1). Аналогично доказывается полнота системы функций (2). Так как $N(x_1) = \min_2(x_1, x_1)$, $\min_2(x_1, x_2) = N(\min_2(x_1, x_2))$, то $\mathcal{L}_A \in \tilde{\mathcal{L}}_{NA}$ (3). Полнота системы функций (4) доказывается аналогично (5). Поскольку $\min_2(x_1, x_2) = p_2(1, \min_2(x_1, x_2))$, то $\mathcal{L}_{NA} \subset \tilde{\mathcal{L}}_G$.

Одним из достаточных условий для того, чтобы система функций \mathcal{L} не была полной, является следующее:

$$\tilde{\mathcal{L}} = \mathcal{L}, \quad \mathcal{L} \subset \mathcal{B}, \quad L \neq \mathcal{B}. \quad (8.11)$$

Л е м м а 8.1. Пусть M — множество всех монотонных функций, L — множество всех линейных функций, $a \in V_1$ и U_a — множество всех булевых функций таких, что $f(a, a, \dots, a) = a$ (число переменных может быть любым), D — множество всех самодвойственных функций. Тогда $M = \tilde{M}$, $L = \tilde{L}$, $U_a = \tilde{O}_a (a \in V_1)$, $D = \tilde{D}$, и ни один из этих классов функций не является полным.

Доказательство. Пусть \mathcal{L} — это один из перечисленных классов функций. Для произвольной функции $f \in \tilde{\mathcal{L}} \cap \mathcal{B}_n$ существует \mathcal{L} -дерево T такое, что $f = \rho_n(T)$. Пусть g — имя корня этого дерева и $m = v(g)$. Тогда, если $L(T) = 0$, то $f = g \in \mathcal{L}$. Предположим, что $0 < L(T) < k$, $f \in \mathcal{L}$, и пусть $L(T) = k$. Согласно соотношению (8.6), существует $\mathcal{L}X$ -дерево $T_i (1 \leq i \leq m)$ такое, что

$$f = g(\rho_n(T_1), \rho_n(T_2), \dots, \rho_n(T_m)),$$

где $L(T_i) < L(T)$. Положим $g_i = \rho_n(T_i)$. Если T_i является \mathcal{L} -деревом, то, по предположению, $g_i \in \mathcal{L}$. При этом если $g_i = x_j$, то функция x_j не принадлежит ни одному из семейств функций, перечисленных в лемме, т. е. $g_i \in \mathcal{L}$. Следовательно, для завершения доказательства леммы достаточно показать, что для произвольных $g \in \mathcal{L} \cap \mathcal{B}_m$, $g_1, g_2, \dots, g_m \in \mathcal{L} \cap \mathcal{B}_n$ $f \in g(g_1, g_2, \dots, g_m) \in \mathcal{L}$.

(1) Пусть $\mathcal{L} = M$. Поскольку $g_i(A) \leq g_i(B)$ ($1 \leq i \leq m$) для произвольных наборов $A \leq B$, $A, B \in V_n$, то $f(A) = g(g_1(A), g_2(A), \dots, g_m(A)) \leq g(g_1(B), g_2(B), \dots, g_m(B)) = f(B)$, т. е. $f \in M$.

(2) — (4) Доказательство аналогичное (проверьте).

(5) Пусть $\mathcal{L} = D$. Поскольку $\bar{g}(\bar{A}) = g(A)$, $\bar{g}_i(\bar{A}) = g_i(A)$ ($1 \leq i \leq m$) на произвольном наборе $A \in V_n$, то $f(A) = \bar{g}(g_1(\bar{A}), \dots, g_m(\bar{A})) = \bar{g}(\bar{g}_1(A), \dots, \bar{g}_m(A)) = g(g_1(A), \dots, g_m(A)) = f(A)$, т. е. $f \in D$.

С другой стороны, (1) $N, p_2 \notin M$ (задача 7.22), (2) $\min_2 \notin L$ (упражнение 7.9), (3) $1 \notin U_0$, (4) $0 \notin U_1$, (5) $p_2 \notin D$ (задача 7.45). Из (8.11) следует справедливость леммы.

Теорема 8.1 (теорема Поста). Система функций $\mathcal{L} \subset \mathcal{F}$ является полной тогда и только тогда, когда она удовлетворяет следующим условиям (1) — (5):

(1) $\mathcal{L} \not\subset M$ (\mathcal{L} содержит, по крайней мере, одну функцию, не являющуюся монотонной);

(2) $\mathcal{L} \not\subset L$ (\mathcal{L} содержит, по крайней мере, одну функцию, не являющуюся линейной);

(3) $\mathcal{L} \not\subset U_0$ (существует функция $g \in \mathcal{L}$ такая, что $g(0, 0, \dots, 0) = 1$);

(4) $\mathcal{L} \not\subset U_1$ (существует функция $g \in \mathcal{L}$ такая, что $g(1, 1, \dots, 1) = 0$);

(5) $\mathcal{L} \not\subset D$ (\mathcal{L} содержит, по крайней мере, одну функцию, не являющуюся самодвойственной).

Доказательство. Необходимость. Если $\mathcal{L} \subset M$, то, по определению, $\tilde{\mathcal{L}} \subset \tilde{M}$. По лемме 8.1 $\tilde{\mathcal{L}} \not\subset \mathcal{F}$, т. е. система \mathcal{L} полной не является. Точно так же проверяется необходимость условий (2) — (5).

Достаточность. Случай А: $\tilde{\mathcal{L}}$ содержит константы 0 и 1. Покажем, что $N(x_1) \in \tilde{\mathcal{L}}$. По условию (1) существует немонотонная функция $g_1 \in \mathcal{L}$. Положим $v(g_1) = n_1$. Так как по определению существует неположительная переменная x_i , то $a_j \in V_1$ ($j \neq i$, $1 \leq j \leq n_1$) и

$$g_1(a_1, \dots, a_{i-1}, 1, a_{i+1}, \dots, a_{n_1}) = 0;$$

$$g_1(a_1, \dots, a_{i-1}, 0, a_{i+1}, \dots, a_{n_1}) = 1.$$

Следовательно, $N(x_1) = g_1(a_1, \dots, a_{i-1}, x_1, a_{i+1}, \dots, a_{n_1}) \in \tilde{\mathcal{L}}$ (упражнение 8.2). Далее покажем, что $\min_2 \in \tilde{\mathcal{L}}$. По условию (2) существует нелинейная функция $g_2 \in \mathcal{L}$. Положим $v(g_2) = n_2$. Из упражнения (7.9) получаем равенство

$$g_2(x_1 \oplus \dots \oplus x_{n_2}) = c_0 \oplus c_1 x_1 \oplus \dots \oplus c_{n_2} x_{n_2} \oplus c_{12} x_1 x_2 \oplus \dots \oplus c_{i_1 i_2 \dots i_k} x_{i_1} x_{i_2} \dots x_{i_k} \oplus \dots \oplus c_{12 \dots n_2} x_1 x_2 \dots x_{n_2}, \quad (8.12)$$

в котором константы равны 0 или 1. Так как функция не является линейной, то некоторые коэффициенты степени 2 или выше отличны от нуля. Выберем один из ненулевых членов степени 2 и выше, который имеет минимальную степень; пусть соответствующий

щий коэффициент есть $c_{i_1 i_2 \dots i_k} (k \geq 2)$. Если в функцию $g_2(x_1, \dots, x_{n_2})$ подставить $x_{i_1} = x_1, x_{i_2} = x_2, x_{i_j} = 1 (3 \leq j \leq k)$, а остальные переменные положить равными 0, то полученная в результате функция $f(x_1, x_2)$ лежит в $\tilde{\mathcal{L}}$. В результате такой подстановки член $c_{i_1 i_2 \dots i_k} x_{i_1} x_{i_2} \dots x_{i_k}$ в (8.12) перейдет в $x_1 x_2$, а остальные члены степени k и выше будут равны 0 (так как, по крайней мере, одна из входящих в них переменных будет равна 0). Далее, так как, по предположению, $k \geq 2$, то коэффициенты при членах степени от 2 до $k-1$ включительно равны 0 и $f(x_1, x_2) = d_0 \oplus c_{i_1} x_1 \oplus c_{i_2} x_2 \oplus x_1 x_2$, где $d_0 = c_0 \oplus c_{i_3} \oplus \dots \oplus c_{i_k}$. Следовательно, $f(x_1 \oplus c_{i_1}, x_2 \oplus c_{i_2}) = d_0 \oplus c_{i_1} c_{i_1} \oplus x_1 x_2$. Положим $h_1(x_1) = x_1$ при $c_{i_1} = 0$, $h_1(x_1) = N(x_1)$ при $c_{i_1} = 1$, $h_2(x_2) = x_2$ при $c_{i_2} = 0$ и $h_2(x_2) = N(x_2)$ при $c_{i_2} = 1$. Тогда $N, f \in \tilde{\mathcal{L}}$, и функция $f(h_1(x_1), h_2(x_2))$ также будет принадлежать $\tilde{\mathcal{L}}$ (упражнение 8.2). Эта функция совпадает с \min_2 , если $d_0 \oplus c_{i_1} c_{i_2} = 0$, и с $N(\min_2)$, если $d_0 \oplus c_{i_1} c_{i_2} = 1$. Отсюда и упражнений 8.3 и 8.4 следует, что система функций \mathcal{L} является полной.

Случай Б: $\tilde{\mathcal{L}}$ содержит константу 1, но не содержит константу 0 (если $\tilde{\mathcal{L}}$ содержит 0 и не содержит 1, теорема доказывается аналогично).

По условию (4) существует функция $g_4 \in \mathcal{L}$ такая, что $g_4(1, 1, \dots, 1) = 0$. Так как $1 \in \tilde{\mathcal{L}}$, то $g_4(1, 1, \dots, 1) = 0 \in \mathcal{L}$, получили противоречие.

Случай В: $\tilde{\mathcal{L}}$ не содержит констант.

По условию (3) существует функция $g_3 \in \mathcal{L}$ такая, что $g_3(0, 0, \dots, 0) = 1$. Функция $f_1(x_1) = g_3(x_1, x_1, \dots, x_1)$ может быть либо константой 1, либо функцией $N(x_1)$. Однако так как $f_1 \in \tilde{\mathcal{L}}$, то, по предположению, эта функция совпадает с $N(x_1)$. По условию (5) в \mathcal{L} существует функция g_5 , не являющаяся самодвойственной. Пусть $v(g_5) = n_5$. Так как эта функция не является самодвойственной, то найдется такой набор $(a_1, a_2, \dots, a_{n_5}) \in V_{n_5}$, что

$$g_5(a_1, a_2, \dots, a_{n_5}) = g_5(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_{n_5}). \quad (8.13)$$

Пусть $f_2(x_1)$ — функция, которая получается при замене переменных x_i функции $g_5(x_1, x_2, \dots, x_{n_5})$ таких, что $a_i = 0$, на x_1 , а остальных переменных на $N(x_1)$. Имеем $f_2 \in \tilde{\mathcal{L}}$. Из (8.13) получаем $f_2(0) = g_5(a_1, a_2, \dots, a_{n_5}) = g_5(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_{n_5}) = f_2(1)$.

Следовательно, f_2 является константой. Получили противоречие.

Следствие 8.1. Функция $g \in \mathcal{B}$ образует полную систему $\{g\}$ тогда и только тогда, когда она удовлетворяет следующим условиям теоремы 8.1: (3) $g(0, 0, \dots, 0) = 1$, (4) $g(1, 1, \dots, 1) = 0$ и (5) g не является самодвойственной.

Доказательство. Необходимость указанных условий следует из теоремы 8.1. Покажем их достаточность. По условиям (3) и (4)

функция g не является константой. Пусть $v(g) = n$. Если предположить, что g монотонная функция, то $g_1(a_1, a_2, \dots, a_n) \leq g(1, 1, \dots, 1) = 0$ для произвольного набора $(a_1, a_2, \dots, a_n) \in V_n$. Получили противоречие. Следовательно, условие (1) выполняется. Допустим, что g — одна из линейных функций p_n или \bar{p}_n . Тогда, как следует из задачи 7.45, число n оказывается четным. Однако при этом p_n не удовлетворяет условию (3), а \bar{p}_n — условию (4). Вновь получили противоречие.

Задача 8.13. В \mathcal{B}_2 выделите функции, каждая из которых образует полную систему.

Решение. Функции $f(x_1, x_2)$, удовлетворяющие условиям (3) и (4) следствия 8.1, имеют следующую совершенную дизъюнктивную нормальную форму: $f(x_1, x_2) = \bar{x}_1 \bar{x}_2 \vee f(0, 1) \bar{x}_1 x_2 \vee f(1, 0) x_1 \bar{x}_2$. Так как f не может быть самодвойственной, то $f(0, 1) = f(1, 0)$. Если последние два значения равны 0, то $f = \bar{x}_1 \bar{x}_2 = x_1 \vee x_2$ (функция НЕ—ИЛИ от двух переменных); если же эти значения равны 1, то $f = x_1 x_2$ (функция НЕ—И от двух переменных).

Задача 8.14. Найдите число функций в \mathcal{B}_n , каждая из которых образует полную систему.

Решение. Число функций в \mathcal{B}_n , удовлетворяющих условиям (3) и (4) следствия 8.1, равно 2^{2^n-2} . Среди них имеется $2^{2^{n-1}-1}$ самодвойственных функций. Это следует из задачи 7.43, так как на каждой из $2^{n-1}-1$ неупорядоченных пар наборов значений переменных самодвойственную функцию можно определить независимо от других наборов, положив ее равной либо 0, либо 1 (пара наборов $(0, 0, \dots, 0)$ и $(1, 1, \dots, 1)$ исключается). Следовательно, искомое число равно $2^{2^n-2} - 2^{2^{n-1}-1}$. При больших n приблизительно каждая четвертая функция в \mathcal{B}_n образует полную систему функций.

Задача 8.15. Функция MUX из задачи 8.5 не образует полную систему функций.

Указание. При решении задачи 8.5 было показано, что рассматриваемая функция не удовлетворяет условиям (3) и (4) следствия 8.1.

Полная система функций \mathcal{L} называется *безызбыточной*, если при исключении из нее любой одной функции она уже не является полной. Положим $\mathcal{L}_1 = M$, $\mathcal{L}_2 = L$, $\mathcal{L}_3 = U_0$, $\mathcal{L}_4 = U_1$, $\mathcal{L}_5 = D$. Для $g \in \mathcal{B}$ и $1 \leq j \leq 5$ положим $P_j(g) = 1$, если $g \notin \mathcal{L}_j$, $P_j(g) = 0$, если $g \in \mathcal{L}_j$. Далее для любого конечного подмножества $\mathcal{L} = (g_1, g_2, \dots, g_m)$ множества \mathcal{B} определим следующую матрицу из m строк и пяти столбцов:

$$P(\mathcal{L}) = (P_j(g_i) | 1 \leq i \leq m, 1 \leq j \leq 5).$$

Следующее утверждение является следствием из теоремы 8.1.

Следствие 8.2. Подмножество $\mathcal{L} = \{g_1, g_2, \dots, g_m\}$ множества \mathcal{B} является безызбыточной полной системой функции тогда и только тогда, когда каждый столбец матрицы $P(\mathcal{L})$ содержит, по крайней мере, один символ 1 и для каждой строки найдется, по крайней мере, один столбец, который имеет символ 1 только в этой строке.

Упражнение 8.5. Пусть $\mathcal{L} = (g_1, g_2, \dots, g_m)$ — безызбыточная полная система функций.

(1) Пусть $g_1=0$, $g_2=1$. Тогда m равно 3 или 4. Если $m=4$, то при естественном упорядочении $g_3=p_{2l+1}$ ($l \geq 1$), а g_4 является монотонной функцией с $v(g_4) \geq 2$.

(2) Если $g_1=0$ (или 1) и $1 \notin \mathcal{L}$ ($0 \notin \mathcal{L}$), то m равно 2 или 3.

(3) Если \mathcal{L} не содержит констант, то $1 \leq m \leq 3$.

Решение. (1) Первой и второй строками матрицы $P(\mathcal{L})$ в данном случае являются соответственно $(0, 0, 0, 1, 1)$ и $(0, 0, 1, 0, 1)$. Согласно следствию 8.2, существование строки (строк), которая (которые) имела бы символы 1 в первом и втором столбцах, означало бы, что $m \leq 4$. Если $m=4$, то без ограничения общности можно считать, что третья и четвертая строки имеют соответственно вид $(1, 0, -, -, -)$ и $(0, 1, -, -, -)$ (знак «—» показывает, что соответствующая компонента может быть произвольной). Следовательно, g_3 есть p_n или \bar{p}_n , а g_4 — нелинейная монотонная функция, т. е. монотонная функция с $v(g_4) \geq 2$. Если n четно, то третьей строкой может быть либо $(1, 0, 0, 1, 1)$, либо $(1, 0, 1, 0, 1)$, но в таком случае можно исключить либо первую, либо вторую строку. Получили противоречие.

Следовательно, n является нечетным и $n \geq 3$. Если g_3 представляет собой \bar{p}_3 , то третья строка $(1, 0, 1, 1, 0)$, и мы вновь приходим к противоречию. Если же g_3 представляет собой p_n , то третья строка $(1, 0, 0, 0, 0)$. При этом четвертой строкой является $(0, 1, 0, 0, -)$. Все условия следствия 8.2 для $P(\mathcal{L})$ выполняются.

(2) Если $g_1=0$, $1 \notin \mathcal{L}$ (для $g_1=1$, $0 \notin \mathcal{L}$ доказательство аналогично), то первой строкой является $(0, 0, 0, 1, 1)$. Функции \mathcal{L} , удовлетворяющие условию (3) теоремы 8.1, не являющиеся константами. Пусть g_2 — одна из таких функций, удовлетворяющая условию (1). Если g_2 удовлетворяет также условию (2), то $m=2$. Если она не удовлетворяет условию (2), то необходима еще функция, удовлетворяющая условию (2).

(3) Пусть g_1 — функция из \mathcal{L} , удовлетворяющая условию (3) или (4) теоремы 8.1; g_1 удовлетворяет также условию (1). Из следствия 8.2 $m \leq 4$. Предположим, например, что $m=4$. Каждая из функций g_1, g_2, g_3, g_4 удовлетворяет ровню одному из условий (2)–(5) и не удовлетворяет трем другим. Отличными от констант линейными функциями, удовлетворяющими условию (5) и не удовлетворяющими условию (2), являются лишь функции p_{2l} и \bar{p}_{2l} ($l \geq 1$). Однако каждая из этих функций удовлетворяет либо условию (3), либо условию (4). Получили противоречие.

Задача 8.16. Пара функций $\mathcal{L} = \{N, g\}$ образует безызбыточную полную систему функций в том и только в том случае, если функция g линейная, но не самодвойственная и принадлежит либо U_0 , либо U_1 .

Указание. Первой строкой матрицы $P(\mathcal{L})$ является $(1, 0, 1, 1, 0)$. Отсюда и следствия 8.2 следует решение задачи.

8.3. Булевы формулы [11–13]

Положим $\mathcal{L}_B = \{0, 1, N, \min_m, \max_m | m \geq 2\}$. Булева формула $\varphi_B(T)$, соответствующая $\mathcal{L}_B X$ -дереву T , определяется по индукции следующим образом. Пусть g — имя корня v дерева T . Если $g \in \mathcal{L}_B$ и $v(g)=0$, то пусть v_i ($1 \leq i \leq v(g)$) — это концевые точки ориентированных ребер, выходящих из v , и T_i — поддеревья дерева T , корнем которого является вершина v_i .

(1) Если $v(g)=0$ или $g \in VX$, то

$$\varphi_B(T) = g.$$

$$(8.14)$$

(2) Если $g=N$, то

$$\varphi_B(T) = \overline{(\varphi_B(T_1))}. \quad (8.15)$$

Если глубина T_1 равна 0, т. е. $\varphi_B(T_1)$ представляет собой 0, 1, x_j , то скобки опускаются и $\varphi_B(T) = \varphi_B(T_1)$.

(3) Если $g = \min_m$, то

$$\varphi_B(T) = (\varphi_B(T_1)) (\varphi_B(T_2)) \dots (\varphi_B(T_m)). \quad (8.16)$$

Так же, как и в случае (2), если $L(T_i)$ или $L(T_i)=1$ и именем корня T_i является N , т. е. если $\varphi_B(T_i)$ представляет собой 0, 1, x_j , $\bar{x}_j (1 \leq j)$, то внешние скобки в выражении $(\varphi_B(T_i))$ опускаются.

(4) Если $g = \max_m$, то

$$\varphi_B(T) = (\varphi_B(T_1)) \vee (\varphi_B(T_2)) \vee \dots \vee (\varphi_B(T_m)). \quad (8.17)$$

Условия исключения скобок такие же, как и в случае (3).

Выражения x_1 , \bar{x}_1 , (\bar{x}_1) , $(x_1 \vee \bar{x}_2) \cdot x_3$, $(x_1 \cdot \bar{x}_2) \vee x_3$ являются булевыми формулами. В булевых формулах вида (8.16) знак « \cdot » может опускаться. Выражения, полученные после исключения скобок в соответствии с соглашениями о порядке выполнения операций, принятыми в § 7.2, также называются булевыми формулами.

Если переписать булеву формулу в терминах операций \vee , \cdot , $—$, то она превратится в \mathcal{L}_B -формулу* в смысле определений, введенных в § 8.2. При $n(T) \leq n$ булева формула $F = \varphi_B(T)$, по определению, реализует функцию $\rho_n(T)$; при этом вместо $\rho_n(T)$ используется запись $\rho_n(F)$. Иногда символ ρ_n опускается, и функция, реализуемая формулой F , обозначается также через F . Введенные в § 7.4 элементарные конъюнкции, элементарные дизъюнкции, переменные x_i , дизъюнктивные нормальные формы, конъюнктивные нормальные формы — булевы формулы. Если $F = \varphi_B(T)$, то определим $n(F) = n(T)$. Если F_1, F_2 — булевы формулы, $n = \max\{n(F_1), n(F_2)\}$ и $\rho_n(F_1) = \rho_n(F_2)$ или $\rho_n(F_1) \leq \rho_n(F_2)$, то будем писать, что $F_1 = F_2$ или $F_1 \leq F_2$.

Далее рассмотрим обобщение правила Де Моргана.

Двойственная булева формула. Предположим, что в булевой формуле F написаны все знаки « \cdot » и все скобки (хотя, в принципе, их можно было бы опустить в соответствии с соглашениями о порядке выполнения операций \vee и \cdot). Если в булевой формуле F заменить все знаки операций \vee на \cdot и наоборот, а также поменять местами символы 0 и 1, то получится новая булева формула, которая называется *двойственной* к F и обозначается через F^d . Например, двойственной к булевой формуле $x_1 \vee (\bar{x}_2 \cdot x_3 \cdot 1)$ является $x_1 \cdot (\bar{x}_2 \vee x_3 \vee 0)$.

Теорема 8.2. (Обобщение правила Де Моргана). Функция h , реализуемая булевой формулой F^d , двойственной к булевой формуле F , двойственна функции f , реализуемой формулой F .

* Ситуация иная, если x_i также является булевой формулой.

Доказательство. Пусть $n(F) \leq n$ и $f = \rho_n(F)$, $h = \rho_n(F^d)$. По определению булевой формулы, существует $\mathcal{L}_B X$ -дерево T такое, что $F = \varphi_B(T)$. Пусть T^d — это $\mathcal{L}_B X$ -дерево, которое получается из T , если поменять местами соответственно имена вершин 0 и 1, и функции \min_m и \max_m ($m \geq 2$). Индукцией по глубине l дерева T покажем, что $F^d = \varphi(T^d)$ и $h = f^d$. При $l=0$ булева формула F имеет вид 0, 1 или x_i , и справедливость доказываемого утверждения очевидна. Предположим, что утверждение справедливо при всех $l < j$, и рассмотрим случай $l=j$. Пусть g (или g^d) — имя корня v (v^d) дерева T (T^d) и T_i — ориентированные поддеревья дерева T (T^d), корнями которых являются концевые точки ориентированных ребер, выходящих из корня v (v^d) ($1 \leq i \leq v(g)$). Положим $F_i = \varphi_B(T_i)$ и $f_i = \rho_n(F_i)$. По предположению индукции $F_i^d = \varphi_B(T_i^d)$, $f_i^d = \rho_n(F_i^d)$.

(1) Если $g=N$, то из (8.15) получаем $F = \overline{(F_1)}$, $f = \overline{f_1}$, $F^d = \overline{(F_1^d)}$, $h = \overline{f_1^d} = (\overline{f_1})^d = f^d$.

(2) Если $g = \min_m$, то из (8.16) получаем $F = (F_1) \cdot (F_2) \cdot \dots \cdot (F_m)$, $f = f_1 \cdot f_2 \cdot \dots \cdot f_m$. При этом из предположений индукции следует, что

$$\begin{aligned} F_i^d &= (F_1^d) \vee (F_2^d) \vee \dots \vee (F_m^d), \quad h(X_n) = f_1^d(X_n) \vee f_2^d(X_n) \vee \dots \vee f_m^d(X_n) = \\ &= \overline{f_1(\overline{X}_n)} \vee \overline{f_2(\overline{X}_n)} \vee \dots \vee \overline{f_m(\overline{X}_n)} = \\ &= \overline{f_1(\overline{X}_n) \cdot f_2(\overline{X}_n) \cdot \dots \cdot f_m(\overline{X}_n)} = \\ &= \overline{f(\overline{X}_n)} = f^d(X_n), \quad \text{где } X_n = (x_1, x_2, \dots, x_n), \quad \overline{X}_n = (\overline{x}_1, \overline{x}_2, \dots, \overline{x}_n). \end{aligned}$$

(3) Если $g = \max_m$, то доказательство аналогично (2).

Следствие 8.3. Пусть F_1^d, F_2^d двойственны соответственно булевым формулам F_1, F_2 . Если $F_1 \leq F_2$, то $F_2^d \leq F_1^d$. Следовательно, если $F_1 = F_2$, то $F_1^d = F_2^d$.

Доказательство следует из теоремы 8.2 и задачи 7.42.

Следствие 8.3 является законом двойственности.

Задача 8.17. Пусть F^d двойственна булевой формуле F . Функция $f = \rho_n(F)$ является самодвойственной тогда и только тогда, когда $F = F^d$.

Решение следует из теоремы 8.2.

Задача 8.18. $\text{Maj}_3(x_1, x_2, x_3) = (x_1 \cdot x_2) \vee (x_2 \cdot x_3) \vee (x_3 \cdot x_1) =$
 $= (x_1 \vee x_2) \cdot (x_2 \vee x_3) \cdot (x_3 \vee x_1).$

Указание. Следует воспользоваться тем, что функция Maj_3 самодвойственна (задача 7.46), а также утверждениями задач 7.24 и 8.17.

Схемы из элементов \mathcal{L}_B и двойственные схемы. Схемы из функциональных элементов, компонентами которых являются инверторы, элементы И и ИЛИ, называются \mathcal{L}_B -схемами. Если в \mathcal{L}_B -схеме N каждый из элементов И и ИЛИ заменить соответственно элементом ИЛИ и И с теми же числами входов и выходов, оставив конфигурацию соединений без изменений, то получится новая схема, которая называется *двойственной* к исходной и обозначается через N^d .

Упражнение 8.6. Если \mathcal{L}_B -схема N реализует функции f_1, f_2, \dots, f_m , то двойственная схема N^d реализует функции $f_1^d, f_2^d, \dots, f_m^d$.

Решение. Воспользуемся индукцией по числу k компонент схемы. При $k=0$ утверждение очевидно. Предположим, что оно верно при всех $k < j$, и рассмотрим случай $k=j$. Так как в графе соединений $C(N)$ схемы N нет ориентированных циклов, то в нем найдется вершина v , соответствующая компоненте схемы, такая, что все выходящие из нее ориентированные ребра входят в вершины, соответствующие внешним выходам. Выходную функцию этой компоненты обозначим через f . Пусть v_1, v_2, \dots, v_h — вершины графа $C(N)$, из которых выходят ориентированные ребра, входящие в вершину v , и g_i — выходная функция компоненты, соответствующей вершине v_i . В $C(N^d)$ вершины обозначаются теми же символами, что и соответствующие им вершины $C(N)$, но с индексом d сверху. По предположению индукции, выходной функцией компоненты v_i^d является g_i^d . Если v — вершина, соответствующая инвертору, то выходной функцией f вершины v будет \bar{g}_1 , а выходной функцией вершины v^d — функция $\bar{g}_1^d = f^d$. Если v — вершина, соответствующая элементу И, то $f = g_1 \cdot g_2 \dots g_h$. При этом выходной функцией вершины v^d будет $g_1^d \vee g_2^d \vee \dots \vee g_h^d = f^d$ (см. пункт (2) доказательства теоремы 8.2). Если v соответствует элементу ИЛИ, то доказательство аналогично.

К настоящему времени достаточно подробно исследован случай, когда в качестве совокупности базисных функций используется \mathcal{L}_B . Это обусловлено следующими причинами. 1) Существуют простые стандартные схемы, реализующие эти функции (§ 8.1). 2) \mathcal{L}_B образует полную систему функций (упражнение 8.4). 3) Как было показано в § 7.2, эта совокупность имеет структуру булевой алгебры. 4) Как показывает рассмотренный далее пример, реализация возникающих на практике логических функций с помощью операций «—», « \vee » и « \cdot », является довольно естественной.

Пример 8.4. Рассмотрим задачу проектирования блока управления автомата для продажи «кока-колы» [23].

Предположим, что одна порция стоит 50 иен, и поэтому покупатель, опустив в автомат монету в 100 иен, должен получить сдачу 50 иен. Предположим также, что покупателю разрешается опустить одну монету в 50 иен или 5 монет по 10 иен. Если напиток не нужно разбавлять водой, то прежде, чем опустить монеты на сумму 50 и более иен, следует нажать кнопку «вода не нужна». Кроме того, предположим, что до того, как кончится кока-кола, покупатель берет стаканы только после того, как закончится наполнение. Для простоты другие функции автомата здесь рассматриваться не будут. В табл. 8.1 приведены различные события, которые могут произойти при продаже кока-колы, и соответствующие им логические функции, принимающие значение 1, если событие происходит, и значение 0, если это событие не происходит. В действительности мы рассмотрим лишь те события, которые могут случиться после завершения обслуживания предыдущего покупателя и связаны с выдачей всех команд на исполнительный механизм, необходимых для обслуживания только очередного покупателя. Исходя из смысла переменных, можно записать следующие формулы:

$$(1) C = C_s \cdot (I_s \vee N_{01}) \cdot K_s \cdot H \cdot (F_s \vee T_{55}) \vee F \vee T_1 \cdot T_2 \cdot T_3 \cdot T_4 \cdot T_5. \quad (8.18)$$

А именно, стакан выдается, если имеются в наличии стаканы, кока-кола, имеется вода или нажата кнопка «вода не нужна», опущена монета в 100 иен и имеется сдача для нее или опущено ровно 50 иен;

$$(2) I = C_r \cdot \overline{N_{01}}. \quad (8.19)$$

Таблица 8.1

Событие	Логическая переменная
События на входе:	
опущена монета в 100 иен	H
опущена монета в 50 иен	F
опущено i монет по 10 иен ($1 \leq i \leq 5$)	T_i
нажата кнопка «вода не нужна»	N_{0i}
Выходные команды:	
выдать команду подачи стакана	C
выдать команду наполнения стакана водой	I
выдать команду подачи кока-колы	K
выдать команду оповещения покупателя об отсутствии возможности его обслужить и возврата всех опущенных им монет;	
выдать команду возврата монеты в 50 иен	F_r
выдать команду возврата 5 монет по 10 иен	T_{5r}
Внутренние состояния:	
для выдачи сдачи осталась только одна монета в 50 иен	F_s
для выдачи сдачи осталось только 5 монет по 10 иен	T_{5s}
осталась только одна порция кока-колы	K_s
вода поступает в автомат	I_s
стаканы имеются в наличии	C_s
стакан уже подан на выход	C_r
стакан уже наполнен водой	I_r

Здесь предполагается, что вода наливается раньше, чем кока-кола,

$$(3) K = I_r \vee C_r \cdot N_{0i}. \quad (8.20)$$

Подача кока-колы начинается, если либо вода уже налита в стакан, либо нажата кнопка «вода не нужна» и стакан подан к месту наполнения;

$$(4) R = \bar{C}_s \vee \bar{K}_s \vee I_s \cdot \bar{N}_{0i} \vee H \cdot \bar{F}_s \cdot \bar{T}_{5s}. \quad (8.21)$$

Покупатель извещается об отсутствии возможности его обслужить, и ему возвращаются все опущенные монеты, если в автомате нет стаканов, нет кока-колы, или покупателю требуется вода, а ее нет;

$$(5) T_{br} = K \cdot H \cdot T_{5s}. \quad (8.22)$$

Возвращается сдача в 50 иен, если выдана команда подачи кока-колы, покупатель опустил в автомат монету в 100 иен, и в автомате имеются 5 и более монет по 10 иен для выдачи сдачи;

$$(6) F_r = K \cdot H \cdot \bar{T}_{5s} \cdot F_s. \quad (8.23)$$

В данном случае, в отличие от (5), сдача выдается одной монетой в 50 иен, если ее невозможно выдать монетами по 10 иен. Конечно, в первую очередь следовало бы возвращать сдачу одной монетой в 50 иен, но с точки зрения экономии места в автомате лучше это сделать так, как было указано. Таким образом, все выходные функции определены. Эта операция оказалась значительно более простой, чем если бы мы строили таблицу значений функций для схемы с десятью входами. При этом все выходные функции имеют вполне понятный смысл, и поэтому с ними легче обращаться. Обычно непосредственно по булевым формулам (8.18)—(8.23) реализующую их схему не строят. К построению схемы приступают после того, как эти булевы формулы будут преобразованы к более «простому» виду (см. гл. 9).

Пример 8.5. Кодирование приоритета. Это устройство представляет собой схему с 2^m+1 входными $E, x_1, x_2, \dots, x_{2^m}$ и $m+2$ выходными переменными $GS, EO, y_1, y_2, \dots, y_m$, связанными между собой следующим образом:

1) если $E=0$, то все выходные переменные равны 1;

2) если $E=1$ и $x_1=x_2=\dots=x_{2^m}=0$, то $EO=1$, а остальные выходные переменные равны 0;

3) если $E=1$ и, по крайней мере, одна из входных переменных x_1, x_2, \dots, x_{2^m} принимает значение 1, то $GS=1, EO=0$ и $y_i=b_i (1 \leq i \leq m)$, где (b_1, b_2, \dots, b_m) — двоичное представление числа $J_{\max}-1$, а J_{\max} — номер последней, равной 1, входной переменной.

Найдем булевы формулы, реализующие каждую из выходных функций, если $m=3$.

1) Поскольку переменная GS равна 1, когда $E=1$ и, по крайней мере, одно из значений x_1, \dots, x_8 равно 1, и равна 0 в остальных случаях, то

$$GS = E \cdot (x_1 \vee x_2 \vee \dots \vee x_8). \quad (8.24)$$

2) Поскольку $EO=1$, только когда $E=1$ и все переменные x_1, \dots, x_8 равны 0, то

$$EO = E \cdot \bar{x}_1 \bar{x}_2 \dots \bar{x}_8. \quad (8.25)$$

3) Переменная y_1 равна 1, когда $E=1$ и $5 \leq J_{\max}$; в остальных случаях она равна 0. Поскольку, по определению, высказывание $(5 \leq J_{\max})$ представляется выражением $x_5 \vee x_6 \vee x_7 \vee x_8$, то

$$y_1 = E \cdot (x_5 \vee x_6 \vee x_7 \vee x_8). \quad (8.26)$$

4) Переменная y_2 равна 1, если $E=1$, а кроме того, $3 \leq J_{\max} \leq 4$ или $7 \leq J_{\max} \leq 8$; в остальных случаях она равна 0. Поскольку высказывания $(3 \leq J_{\max} \leq 4)$ и $(7 \leq J_{\max} \leq 8)$ представляются соответственно выражениями $(x_3 \vee x_4) \bar{x}_5 \bar{x}_6 \bar{x}_7 \bar{x}_8$ и $x_7 \vee x_8$, то $y_2 = E \cdot ((x_3 \vee x_4) \bar{x}_5 \bar{x}_6 \bar{x}_7 \bar{x}_8 \vee x_7 \vee x_8)$. Это выражение можно упростить следующим образом (можно было бы также воспользоваться методами упрощения булевых формул, описанными в гл. 9). Высказывание $(3 \leq J_{\max} \leq 4) \vee (7 \leq J_{\max} \leq 8)$ принимает значение 1, если либо $x_7 \vee x_8 = 1$, либо равны 0 обе величины x_5, x_6 , и равна 1 одна из величин x_3 или x_4 . В остальных случаях это высказывание равно 0. Следовательно,

$$y_2 = E \cdot ((x_3 \vee x_4) \bar{x}_5 \bar{x}_6 \vee x_7 \vee x_8). \quad (8.27)$$

5) Имеем $y_3 = (E=1) \cdot (J_{\max}-1 = \text{четному числу})$. Так как, по определению, высказывание $(J_{\max} = \text{четному числу})$ совпадает с выражением

$$\bar{x}_2 \bar{x}_3 \bar{x}_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \bar{x}_8 \vee x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \bar{x}_8 \vee x_6 \bar{x}_7 \bar{x}_8 \vee x_8, \quad (8.28)$$

то y_3 является произведением E и приведенной ранее дизъюнктивной нормальной формы. Воспользовавшись упражнением 8.3, выражение (8.28) можно упростить, при этом

$$y_3 = E \cdot (\bar{x}_2 \bar{x}_3 \bar{x}_5 \bar{x}_7 \vee x_4 \bar{x}_5 \bar{x}_7 \vee x_6 \bar{x}_7 \vee x_8). \quad (8.29)$$

Кодирование приоритета используется для выбора максимального числа из заданной совокупности и других целей.

* Это выражение является функцией от J_{\max} , принимающей значение 1 при $J_{\max} \geq 5$, и 0 в остальных случаях.

8.4. Коды с обнаружением и исправлением ошибок

Одним из путей повышения надежности систем связи и вычислительной техники является использование кодов, обнаруживающих и исправляющих ошибки [29, 30].

Пример 8.6. Коды, обнаруживающие одиночные ошибки. Предположим, что нужно передать по каналу связи или хранить в памяти информационную последовательность $(a_1, a_2, \dots, a_n) \in V_n$. Пусть вместо этой последовательности передается или хранится последовательность $A = (a_1, a_2, \dots, a_{n+1})$, которая получается добавлением к исходной еще одного символа $a_{n+1} = a_1 \oplus a_2 \oplus \dots \oplus a_n$. Если последовательность A была передана по каналу или записана в память, а получена на выходе канала или считана из памяти последовательность $B = (b_1, b_2, \dots, b_{n+1})$ также с $n+1$ компонентами, то равенство $a_i \oplus b_i = e_i = 1$ означает, что в i -й компоненте произошла ошибка; если же эта сумма равна 0, то в i -й компоненте при передаче (хранении) ошибки не произошло. Поскольку $a_1 \oplus a_2 \oplus \dots \oplus a_{n+1} = 0$, то

$$r = b_1 \oplus b_2 \oplus \dots \oplus b_{n+1} = e_1 \oplus e_2 \oplus \dots \oplus e_{n+1}. \quad (8.30)$$

Таким образом, если $r=1$, то в нечетном числе компонент произошли ошибки; если же $r=0$, то ошибок не было или ошибки произошли в четном числе компонент. Если вероятность возникновения двух или более ошибок очень мала, то этим событием можно пренебречь. При этом описанный метод, требующий введения одного избыточного двоичного символа, позволяет обнаружить, возникла ли ошибка при передаче или хранении информации. Благодаря своей простоте метод нашел очень широкое применение. Добавляемый к исходной последовательности символ a_{n+1} называется символом четности. Вместо присоединения символа a_{n+1} к исходной последовательности можно также присоединять его отрицание, что также позволяет обнаруживать нечетное число ошибок.

Далее в этом параграфе мы будем рассматривать V_n как n -мерное векторное пространство над полем Z_2 (см. § 4.5). При этом элементы v из V_n записываются в виде векторов-строк и для обозначения сложения в Z_2 вместо символа \oplus используется $+$. Сложение векторов также будем обозначать знаком $+$.

Пример 8.7. Рассмотрим векторы $X = (x_1, x_2, \dots, x_7) \in V_7$, удовлетворяющие следующей системе уравнений:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = 0 \Leftrightarrow \begin{cases} x_1 = x_3 + x_5 + x_7 \\ x_2 = x_3 + x_6 + x_7 \\ x_4 = x_5 + x_6 + x_7 \end{cases} \quad (8.31)$$

Соотношение в левой части (8.31) записано в матричной форме, и символ 0 в его правой части представляет собой нулевой

вектор-столбец размерности 3. Произвольному вектору $Y = (y_1, y_2, y_3, y_4)$ из V_4 сопоставим вектор $\Phi(y) = X = (x_1, x_2, \dots, x_7)$, где $x_3 = y_1$, $x_5 = y_2$, $x_6 = y_3$, $x_7 = y_4$, а x_1 , x_2 и x_4 выбираем так, чтобы выполнялось соотношение (8.31). Предположим, что вместо Y по каналу связи передается (или хранится в памяти) вектор $\Phi(Y)$. При передаче или хранении слова могут возникнуть (а могут не возникнуть) ошибки, т. е. некоторые компоненты исходного слова могут измениться: 0 может перейти в 1, а 1 — в 0. Поэтому полученное в результате передачи или хранения слово может отличаться от исходного, и мы обозначим его через $X' = (x'_1, x'_2, \dots, x'_7)$. Задача состоит в том, чтобы по известному слову $(x'_1, x'_2, \dots, x'_7)$ найти правильное исходное слово (x_1, x_2, \dots, x_7) . Предположим, что вероятность возникновения ошибок в двух или более компонентах достаточно мала, и этим событием можно пренебречь.

Положим $x_i + x'_i = e_i$. Поскольку знак «+» означает здесь сложение по модулю 2, то равенство $e_i = 1$ означает, что в i -й компоненте произошла ошибка; если же $e_i = 0$, то в i -й компоненте ошибки нет. Используя матрицу из левого соотношения в (8.31), введем величины s_1, s_2, s_3 , положив

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \\ x'_5 \\ x'_6 \\ x'_7 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}. \quad (8.32)$$

Из (8.31) и (8.32) имеем

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}. \quad (8.33)$$

По предположению, только одна из величин e_1, e_2, \dots, e_7 может быть равной 1. При отсутствии ошибок, когда $e_1 = e_2 = \dots = e_7 = 0$, имеем $s_1 = s_2 = s_3 = 0$.

Далее рассмотрим случай, когда $e_i = 1$ и $e_j = 0$ ($i \neq j$). Заметим, что матрица из левой части (8.33)

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

состоит из различных столбцов, причем k -й из них является двоичным представлением числа k (младшими разрядами двоичного представления являются верхние символы столбца). По предполо-

жению, $\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}$ совпадает с i -м столбцом матрицы H . Следова-

тельно, ошибочной компонентой является компонента с номером $i = s_1 + 2s_2 + 4s_3$. Так как $x_i = x'_i + e_i$, то легко найти правильные символы x_1, \dots, x_7 . Множество векторов, удовлетворяющих условию (8.31), называемых кодовыми векторами, называется *кодом Хэмминга длины 7*.

В общем случае для представления K различных объектов необходимо $\lceil \log K \rceil$ двоичных символов ($\lceil A \rceil$ — введено ранее). Однако если для представления этих объектов использовать некоторое число избыточных двоичных символов, то, как показывает приведенный пример, можно исправлять и обнаруживать некоторую часть ошибок или упростить схемы, предназначенные для обработки информации. Если, в действительности, вместо $\lceil \log_2 K \rceil$ используется n двоичных символов, то число $n - \lceil \log_2 K \rceil$ называется числом *проверочных символов*. Число проверочных символов всегда желательно выбирать возможно меньшим, конечно, при соблюдении всех остальных требований к системе, например обеспечении возможности исправления всех одиночных ошибок.

Задача 8.19. Для кода из примера 8.7, исправляющего одиночные ошибки, в границе Хэмминга (2.26) имеет место равенство, т. е. этот код содержит максимальное возможное число кодовых слов среди всех кодов длины 7, исправляющих одиночные ошибки.

Указание. Число кодовых слов в коде из примера 8.7 равно 2^4 . Далее следует положить $l=7, r=1$.

Некоторые определения, относящиеся к кодам. Непустое подмножество C множества V_n называется *кодом длины n* . Элементы C называются *кодowymi словами* или *кодowymi векторами*. Минимальное расстояние d кода C определяется равенством

$$d = \min \{d(v, u) \mid v \neq u, v, u \in C\},$$

где $d(v, u)$ — расстояние Хэмминга между u и v (см. § 2.2).

Предположим, что по каналу передается или хранится в памяти кодовый вектор v , а принимается на выходе канала или считывается из памяти вектор u . Вектор $e = v + u$ называется *вектором ошибок*. Если i -я компонента вектора ошибок e равна 1, то это означает, что при передаче или при записи, хранении или считывании i -го двоичного символа кодового слова возникла ошибка; если же $e_i = 0$, то в этом символе ошибки нет. В задачах связи обычно известен лишь вектор $u = v + e$. Обозначим через $w(e)$ вес вектора e . Пусть $0 \leq t_1 \leq t_2$. Код C называется *кодом, исправляющим ошибки кратности (веса) t_1 и обнаруживающим ошибки кратности (веса) t_2* , если для произвольного кодового вектора $v \in C$ и произвольного вектора ошибок $e \in V_n$ веса t_2 или менее по вектору $v + e$: 1) при $w(e) \leq t_1$ можно определить v ; 2) при $t_1 <$

$\leq \omega(e) \leq t_2$ можно установить, что вес e равен t_1+1 или больше. В частности, при $t_1=t_2$ такой код называют кодом, исправляющим ошибки кратности t_1 , а при $t_1=0$ — кодом, обнаруживающим ошибки кратности t_2 . Если t_1 или t_2 равно 1, то ошибки называют одиночными. Операции (1) и (2) называют декодированием.

З а д а ч а 8.20. Пусть $0 \leq t_1 < t_2$. Код C исправляет ошибки кратности t_1 и обнаруживает ошибки кратности t_2 тогда и только тогда, когда минимальное расстояние d кода C удовлетворяет неравенству

$$d \geq t_1 + t_2 + 1. \quad (8.34)$$

Указание. К решению этой задачи можно подойти так же, как в § 2.2. Необходимое и достаточное условие, указанное в задаче, эквивалентно тому, что для двух различных кодовых векторов u, u' , множество $\{u | d(u, v) \leq t_1, u \in V_n\} \cap \{u | d(u, v') \leq t_2, u \in V_n\}$ является пустым.

Заметим, что для одного и того же кода C значения t_1 и t_2 , удовлетворяющие условию (8.34), можно выбрать различными способами.

Определения линейного кода и проверочной матрицы. Если код C является k -мерным линейным подпространством n -го линейного пространства V_n над Z_2 , то он называется (двоичным) *линейным* (n, k) -кодом; при этом k называется *числом информационных символов*, а $(n-k)$ — *числом проверочных символов*. Коды, описанные в примерах 8.6 и 8.7, являются линейными.

Подпространство, двойственное линейному коду C , называется *кодом, двойственным коду C* . Как хорошо известно, нулевое подпространство, двойственное k -мерному линейному подпространству пространства V_n , является $(n-k)$ -мерным линейным подпространством, а следовательно, код, двойственный линейному (n, k) -коду, — линейным $(n, n-k)$ -кодом. Пусть h_1, h_2, \dots, h_{n-k} — произвольный базис двойственного кода. Матрица H размера $(n-k) \times n$, i -й строкой которой является h_i , называется *проверочной матрицей* кода C . По определению,

$$Hu^T = 0 \Leftrightarrow u \in C, \quad (8.35)$$

где u^T — вектор, транспонированный к u , и 0 — нулевой вектор размерности $n-k$.

В примере 8.7 был определен код с помощью матрицы H . Поскольку матрица H имеет ранг $n-k$, она содержит невырожденную подматрицу H' размера $(n-k) \times (n-k)$. После перестановки столбцов матрицы H , т. е. перенумерации компонент векторов, можно считать, что H' состоит из последних $n-k$ столбцов матрицы H . Хорошо известно, что с помощью элементарных операций над строками (одна строка прибавляется к другой строке) матрицу H можно привести к следующему виду:

$$H_0 = [P, I_{n-k}], \quad (8.36)$$

где P — матрица размера $(n-k) \times k$, а I_{n-k} — единичная матрица размера $(n-k) \times (n-k)$.

Поскольку H_0 получена из H с помощью элементарных операций над строками, то

$$Hv^T = 0 \Leftrightarrow H_0 v^T = 0, \quad (8.37)$$

т. е. H_0 также является проверочной матрицей кода C . Обозначим через p_{ij} элемент (i, j) матрицы P . Как следует из (8.37), необходимым и достаточным условием принадлежности вектора $v = (v_1, v_2, \dots, v_n)$ коду C является выполнение равенств

$$v_{k+i} = \sum_{j=1}^k p_{ij} v_j \quad (1 \leq i \leq n-k). \quad (8.38)$$

Эти равенства показывают, что для произвольных v_1, \dots, v_k существует ровно один кодовый вектор, первыми k компонентами которого являются v_1, \dots, v_k . Они дают также способ определения остальных $n-k$ символов кодового слова. Другими словами, если первые k компонент считать информационными символами и рассматривать их как блоки длины k , формируемые для передачи источником информации, то остальные $n-k$ компонент кодового слова можно считать дополнительными и вычислять с помощью равенств (8.38). Эта операция называется *кодированием*.

Пусть e — вектор ошибок. Из (8.35) получаем

$$s = H(v + e)^T = He^T. \quad (8.39)$$

Вектор s называется *синдромом*. Синдром s представляет собой $(n-k)$ -мерный вектор-столбец над Z_2 .

Пример 8.8. Предположим, что задано число k информационных символов. Пусть m — целое положительное число такое, что

$$2^{m-1} - m < k \leq 2^m - m - 1 \quad (8.40)$$

и $n = k + m$. Зададим матрицу H в виде

$$H = [P, I_m], \quad (8.41)$$

где P — такая матрица размера $m \times k$ над Z_2 , что никакие два столбца матрицы H не совпадают и все столбцы являются ненулевыми. При выполнении условий (8.40) такую матрицу H всегда можно найти. Точно так же, как и в примере 8.7, можно показать, что линейный (n, k) -код C с описанной выше проверочной матрицей является кодом, исправляющим одиночные ошибки. При $k = 2^m - m - 1$ код C называется *кодом Хэмминга*; при других значениях k этот код называется *укороченным кодом Хэмминга*.

Кодов, имеющих то же число информационных символов k , что и (укороченный) код Хэмминга, но меньшее число m' проверочных символов, не существует. Действительно, допустим, что такой код существует. Тогда из границы Хэмминга (2.26) будет следовать, что

$$2^k (1 + m' + k) < 2^{m'+k}.$$

Так как $m' < m$ и $2^x - x$ является монотонно возрастающей функцией x , то последнее неравенство противоречит (8.40). Схема из функциональных элементов, реализующая операции (8.38) и (8.39), необходимые для осуществления кодирования и декодирования, содержит схемы, реализующие функции четности p_i . Так как сложность этой схемы почти пропорциональна l , а задержка пропорциональна $\log_2 l$ (утверждение (3) упражнения 8.1), то желательно, чтобы каждая строка P содержала как можно меньше символов 1. Если $k < 2^m - m - 1$, то в качестве столбцов P , в первую очередь, целесообразно брать столбцы минимального веса, т. е. содержащие меньше символов 1.

Далее положим

$$P_a = \begin{bmatrix} P \\ u \end{bmatrix},$$

где u — вектор из V_h , который выберем таким образом, чтобы все столбцы P_a имели нечетный вес. Пусть C_a — линейный $(n+1, k)$ -код, проверочной матрицей которого является матрица

$$H_a = [P_a, I_{m+1}].$$

Предположим, что при передаче не возникает ошибок кратности 3 и выше. Пусть $e = (e_1, e_2, \dots, e_{n+1})$ — вектор ошибок, $s = H_a e^T$ — синдром и $s^T = (s_1, s_2, \dots, s_{m+1})$.

(1) Если $e = 0$, то $s^T = 0$.

(2) Если $e_i = 1, e_j = 0$ ($i \neq j$), то s совпадает с i -м столбцом матрицы H_a и $\sum_{l=1}^{m+1} s_l = 1$. Так как столбцы матрицы H_a попарно различны, то по s можно определить i .

3) Предположим, что $e_i = e_j = 1$ ($i \neq j$) и остальные компоненты e равны 0. Тогда s будет представлять собой сумму i -го и j -го столбцов матрицы H_a , $s^T \neq 0$ и $\sum_{l=1}^{m+1} s_l = 0$, т. е. синдром в данном случае отличается от синдромов, которые получаются в случаях (1) и (2).

Из вышесказанного следует, что код C исправляет все одиночные ошибки и обнаруживает двойные ошибки. Этот код используется для повышения надежности запоминающих устройств и для других целей [29]. При использовании кодов в полупроводниковых запоминающих устройствах для реализации операций кодирования и декодирования требуются высокоскоростные схемы.

Пример 8.9. Рассмотрим линейный (7.3) код с проверочной матрицей

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

Пусть $e = (e_1, e_2, \dots, e_7)$ — вектор ошибок и s_1, s_2, s_3, s_4 — компоненты синдрома. Для вышеприведенной матрицы

$$s_1 = e_1 + e_2 + e_4;$$

$$s_2 = e_2 + e_3 + e_5;$$

$$s_3 = e_3 + e_4 + e_6;$$

$$s_4 = e_4 + e_5 + e_7;$$

Полагая $s'_1 = s_1$, $s'_2 = s_1 + s_2 + s_3$, $s'_3 = s_1 + s_2 + s_4$, имеем

$$s'_1 = e_1 + e_2 + e_4;$$

$$s'_2 = e_1 + e_5 + e_6;$$

$$s'_3 = e_1 + e_3 + e_7.$$

Заметим, что каждый из символов e_i , кроме e_1 , входит только в одну из сумм s'_1, s'_2, s'_3 . Если предположить, что число ошибок не превосходит 1, то при $e_1 = 0$ только одна из сумм s'_1, s'_2, s'_3 может быть равна 1, а остальные будут равны 0. Если же $e_1 = 1$, то $s'_1 = s'_2 = s'_3 = 1$. Следовательно,

$$e_1 = \text{Maj}_3(s'_1, s'_2, s'_3) = \text{Maj}_3(s_1, s_1 + s_2 + s_3, s_1 + s_2 + s_4).$$

Аналогично при $1 \leq i \leq 7$

$$e_i = \text{Maj}_3(s_i, s_i + s_{i+1} + s_{i+2}, s_i + s_{i+1} + s_{i+3})$$

(индексы вычисляются здесь по модулю 7). Этот метод декодирования называется *мажоритарным декодированием*.

Задачи

8.1. Так как значение симметрической функции определяется числом равных 1 значений входных переменных, то ее можно реализовать путем подсчета таких единичных значений [24]. Покажите, что приведенная на рис. 8.14 схема реализует $S_3^{(5)}$. На этом рисунке через FA_i обозначены одноразрядные полные сумматоры, не показаны входы, фиксированные в состоянии 0, и выходы, которые не используются.

8.2. Пусть $d_0 d_1$ — это двухразрядное десятичное число (d_0 и d_1 показывают соответственно число единиц и число десятков) и $a_3 a_2 a_1 a_0, b_3 b_2 b_1 b_0$ — двоичные представления чисел d_1 и d_0 (a_0, b_0 — младшие разряды). На рис. 8.15 приведена схема [23], входами которой являются $a_3, a_2, a_1, a_0, b_3, b_2, b_1, b_0$, а выходом — двоичное представление $c_6 c_5 c_4 \dots c_0$ числа $d_1 d_0$. Покажите, что эта схема правильно выполняет указанные выше функции.

8.3. Покажите, что заданную функцию $f \in \mathcal{B}_n$ можно реализовать, соединив несколько схем, указанных на рис. 8.16,а, так, как это указано на рис. 8.16,б (такая схема называется одномерной каскадной схемой). При этом следует выбрать соответствующим образом постоянные входы c_i и использовать одни и те же входные переменные несколько раз.

Рассмотрите другой способ реализации функции f с помощью одномерной каскадной схемы в случае, когда в качестве входов можно использовать не только переменные x_i , но и их отрицания \bar{x}_i .

8.4. Рассмотрим показанную на рис. 8.17 одномерную каскадную схему, на входы которой подаются двоичные представления

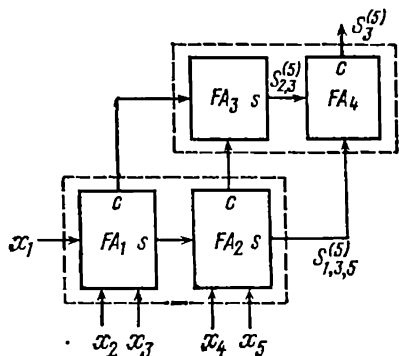


Рис. 8.14. Схема из задачи 8.1

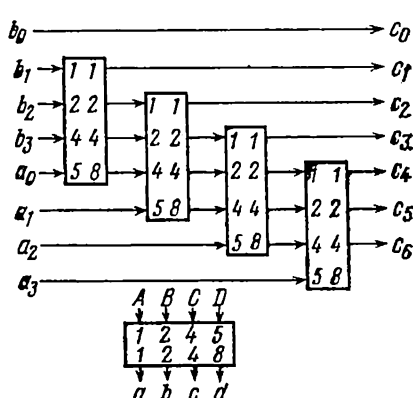


Рис. 8.15. Схема из задачи 8.2.

Нижний элемент представляет собой схему, удовлетворяющую условию $A + 2B + 4C + 5D = a + 2b + 4c + 8d$, если символы 0 и 1 рассматривать как целые числа

неотрицательных целых чисел $A = \sum_{i=0}^{n-1} a_i 2^{n-i-1}$ и $B = \sum_{i=0}^{n-1} b_i 2^{n-i-1}$,

не превосходящих $2^n - 1$. Найдите выходную функцию компонент этой схемы; левые выходы схемы принимают следующие значения: $x_n = 1, y_n = 0$, если $A > B$; $x_n = y_n = 1$, если $A = B$; $x_n = 0, y_n = 1$, если $A < B$. При решении этой задачи все компоненты следует считать одинаковыми, не зависящими от n .

8.5. На рис. 8.18 изображена двумерная каскадная схема, компоненты которой имеют по три входа и реализуют на выходе мажоритарную функцию. (1) Пусть функция $f(x_1, x_2, \dots, x_n)$ является самодвойственной. Из задачи 7.44 следует, что $f(x_1, x_2, \dots$

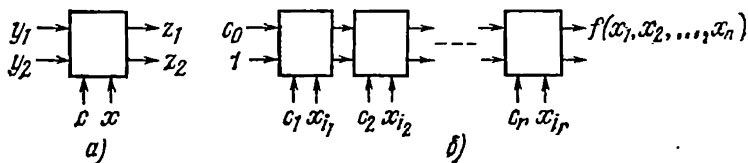


Рис. 8.16. Одномерная каскадная схема Шорта

$\dots, x_n) = \bar{x}_1 g(x_2, \dots, x_n) \vee x_1 g^d(x_2, \dots, x_n)$. Представим g, g^d в совершенной дизъюнктивной нормальной форме: $g = t_1 \vee t_2 \vee \dots \vee t_m$ и $g^d = u_1 \vee u_2 \vee \dots \vee u_l$. Вначале покажите, что для произвольных $i, j (1 \leq i \leq m, 1 \leq j \leq l)$ t_i и u_j имеют, по крайней мере, одну общую переменную. Далее покажите, что если в качестве z_{ij} брать одну

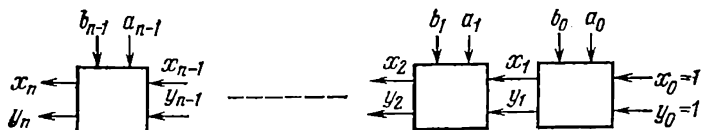


Рис. 8.17. Схема сравнения

из таких переменных, то на самом правом выходе h схемы, изображенной на рис. 8.18, будет реализована функция f . (2) Если f не является самодвойственной, то ее можно реализовать, построив описанную выше схему для функции f^{sd} из упражнения 7.5 и используя вместо y постоянный вход 1.

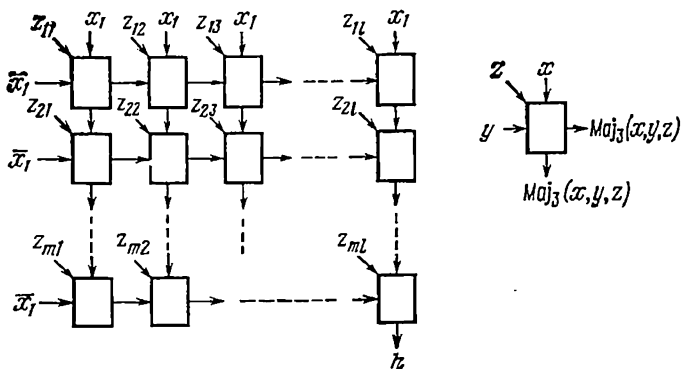


Рис. 8.18. Двумерная каскадная схема Канади

8.6. Покажите, что если $\{g_1, g_2, \dots, g_m\}$ — полная система функций, то $\{g^d_1, g^d_2, \dots, g^d_m\}$ также является полной системой.

Глава 9

Методы минимизации

9.1. Задачи минимизации булевых формул [17—19]

Из задачи 7.17 следует, например, что

$$(x_1 \vee x_2) (\bar{x}_1 \vee x_3) (x_2 \vee x_3) = (x_1 \vee x_2) (\bar{x}_1 \vee x_3).$$

Этот пример показывает, что несколько булевых формул может

представлять одну и ту же функцию. Рассмотрим вначале следующую задачу.

Задача о тождественности. Даны две булевы формулы, F и G . Спрашивается: «Верно ли, что $F=G$?» Основными методами решения этой задачи являются следующие.

1) Выделяется некоторая переменная, например x_1 , в F и G вместо переменной x_1 подставляются ее значения $a \in V_1$ и далее с помощью равенств $\overline{0}=1$, $\overline{1}=0$ и закона дополнения из табл. 7.1 полученные в результате подстановки булевы формулы преобразуются так, чтобы в них не входили символы 0 и 1. Полученные после преобразования булевы формулы обозначим через $F_{x_1=a}$ и $G_{x_1=a}$. Задача свелась, таким образом, к проверке справедливости равенств $F_{x_1=0}=G_{x_1=0}$ и $F_{x_1=1}=G_{x_1=1}$. Если описанную операцию проделать достаточное число раз, то через конечное число шагов мы получим ответ (этот метод применялся раньше для доказательства обычных равенств). В худшем случае этот метод требует последовательного построения $2^{n+2}-4$ булевых формул и перебора всех значений переменных x_1, x_2, \dots, x_n *

2) С помощью тождеств и теорем булевой алгебры булевы формулы F и G последовательно преобразуются шаг за шагом до тех пор, пока они обе не совпадут. Трудность применения этого метода заключается в отсутствии простых правил, которые указывали бы, какие преобразования и в какой последовательности необходимо выполнять.

3) Первые два метода можно комбинировать.

Далее в некоторых частных случаях будет изучаться сложность указанных методов решения задачи о тождественности.

Булева формула, которая представляет функцию, не равную тождественно нулю, называется «удовлетворимой».

Задача о тождественности (Задача об «удовлетворимости» конъюнктивной нормальной формы). Дана произвольная конъюнктивная нормальная форма. Спрашивается, равна ли она тождественно нулю?

Конъюнктивную и дизъюнктивную нормальные формы можно рассматривать просто как последовательность составляющих их букв (переменных и их отрицаний), скобок и знаков \vee , \cdot . Длина такой последовательности называется *длиной конъюнктивной или дизъюнктивной нормальной формы*. Как следует из задачи 7.48, увеличив длину конъюнктивной нормальной формы не более чем в 7 раз, сформулированную задачу о тождественности можно свести к следующей.

Задача о 3-удовлетворимости. Дана произвольная конъюнктивная нормальная форма $s_1 \cdot s_2 \cdot \dots \cdot s_m$, каждая элементарная

* Если предположить, что описанная операция выполняется для всех переменных, то число формул равно $2 \sum_{i=1}^n 2^i = 2^{n+2} - 4$.

дизъюнкция s_i которой содержит не более трех букв. Спрашивается, равна ли она тождественно нулю?

Рассмотрим задачу о тождественности D , например подобную сформулированным, содержащую бесконечно много конкретных вопросов некоторого определенного типа (в ранее приведенном примере вопросы относятся к различным конъюнктивным нормальным формам). Если для такой задачи D существует «программа» (процедура), которая дает правильный ответ за время не более чем $C_1 l^{C_2}$ после того, как в качестве входных данных дана запись конкретной задачи (в приведенном ранее примере это конъюнктивная нормальная форма, рассматриваемая как последовательность символов, а l — длина этой последовательности), то говорят, что задача D разрешима за полиномиальное время. Величины c_1, c_2 являются здесь константами, зависящими только от программы и вычислительного устройства*. Задача о 3-удовлетворимости и задача из примера 9.1 называются *NP-полными*. Доказано [25], что если хотя бы одна из этих задач разрешима за полиномиальное время, то целый класс задач**, включающий все *NP*-полные задачи, также разрешим за полиномиальное время. Так как этот класс задач известен и решается давно, а кроме того, содержит много задач, разрешимость которых за полиномиальное время до сих пор не доказана, то многие считают, что задача о *NP*-полноте не является разрешимой за полиномиальное время (авторам, однако, не известны доказательства этого факта).

Пример 9.1. Примеры *NP*-полных задач [25].

1) Для заданных неориентированного графа G и положительного целого числа k требуется определить:

а) имеется ли в G полный подграф из k вершин?

б) можно ли выбрать множество S из k вершин так, чтобы, по крайней мере, одна из концевых точек каждого ребра графа принадлежала S ?

в) существует ли в графе гамильтонов цикл?

г) верно ли, что хроматическое число графа (§ 4.6) не больше k ?

2) Для заданных ориентированного графа G и целого положительного числа k определить:

а) можно ли уничтожить все ориентированные циклы, удалив из графа k вершин (ребер)?

* Понятия программы, вычислительного устройства и времени вычислений, вообще говоря, должны быть определены строго [25], но здесь нам достаточно понимать их интуитивно.

** Например, если в задаче об удовлетворимости можно было бы доказать (также за полиномиальное время) возможность выбора для каждой переменной x_i значения 0 или 1, при котором каждая элементарная конъюнкция оказалась равной 1, то это означало, что соответствующая конъюнктивная нормальная форма является удовлетворимой. Таким образом, при попытке получить ответ можно опереться на «интуицию» и за полиномиальное время проверить правильность интуиции. Упомянутый класс задач является именно таковым [25].

б) существует ли в графе G ориентированный гамильтонов цикл?

3) Заданы совокупность S_1, S_2, \dots, S_n конечных множеств и целое положительное число k . Существуют ли k множеств $S_{i_1}, S_{j_1}, \dots, S_{i_k}$, обладающих одним из следующих двух свойств:

$$а) \bigcup_{j=1}^k S_{i_j} = \bigcup_{j=1}^n S_j; \quad б) \bigcup_{j=1}^k S_{i_j} = \bigcap_{j=1}^n S_j,$$

и различные множества S_{i_j} и S_{i_l} не имеют общих элементов.

Задача о 3-удовлетворимости эквивалентна следующей двойственной задаче: дана произвольная дизъюнктивная нормальная форма, состоящая из элементарных конъюнкций ранга 3 и менее. Спрашивается, равна ли она тождественно единице.

В § 9.2 и 9.3 рассматривается проблема минимизации дизъюнктивных и конъюнктивных нормальных форм. Какой бы критерий простоты формы не использовался, константы 0 и 1 (далее они рассматриваются как разновидности конъюнктивных и дизъюнктивных нормальных форм) должны рассматриваться как самые простые формы. Если бы для задачи минимизации существовала процедура определения простейшей формы за полиномиальное время (относительно длины произвольной дизъюнктивной или конъюнктивной нормальной формы), то задача о 3-удовлетворимости также была бы разрешима за полиномиальное время. Это означает, что задача минимизации не проще, чем NP -полные задачи.

К настоящему времени в теории проектирования логических схем наиболее полно исследована задача минимизации дизъюнктивных нормальных форм. Поскольку рассматриваемые далее методы и идеи применимы во многих подобных ситуациях, то в данной главе мы будем заниматься дизъюнктивными нормальными формами.

И—ИЛИ-схемы глубины 2. До § 9.3 будет предполагаться, что в качестве внешних входов разрешается использовать переменные x_1, x_2, \dots, x_n и их отрицания $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$. Согласно задаче 7.24,

$$\text{Maj}_3(x_1, x_2, x_3) = x_1 x_2 \vee x_2 x_3 \vee x_1 x_3.$$

Дизъюнктивная нормальная форма, представляющая собой правую часть последнего равенства, реализуется схемой, изображенной на рис. 9.1, и состоит из элементов И и ИЛИ (эта схема используется в качестве элемента полного сумматора, изображенного на рис. 8.6). Схемы, подобные только что описанной, а точнее, удовлетворяющие следующим двум условиям: 1) глубина схемы не превосходит 2; компоненты (далее просто вентили) глубины 2 все — элементы ИЛИ; 2) среди вентилях глубины 1, выходы которых соединяются непосредственно с внешними выходами схемы, могут быть как элементы И, так и элементы ИЛИ; все остальные вентили глубины 1 — элементы И, — называются И—

ИЛИ-схемами глубины 2. Схема называется ИЛИ—И-схемой глубины 2, если двойственная к ней схема является И—ИЛИ-схемой глубины 2. Приближенной мерой сложности таких схем обычно может служить неубывающая линейная функция

$$c_G N_G + c_I N_I, \quad (9.1)$$

где N_G — общее число вентилях в схеме; N_I — общее число входов у всех вентилях; c_G, c_I — постоянные, представляющие собой неотрицательные целые числа.

Как уже указывалось в примере 8.1, для заданной И—ИЛИ-схемы (ИЛИ—И-схемы) глубины 2 можно легко найти дизъюнктивную (конъюнктивную) нормальную форму, представляющую функцию, которая реализуется этой схемой. Наоборот, для произвольной заданной совокупности дизъюнктивных нормальных форм

$$F_i = \bigvee_{j=1}^{l_i} t_{ij}, \quad 1 \leq i \leq m, \quad (9.2)$$

построим следующие схемы.

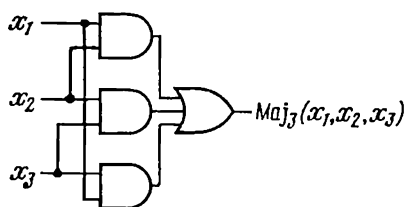


Рис. 9.1. Пример И—ИЛИ-схемы глубины 2

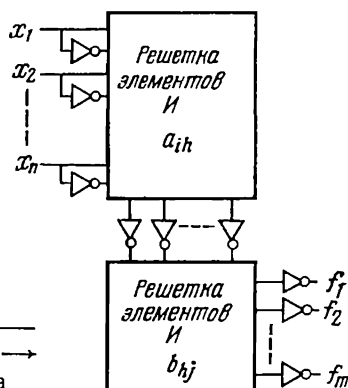


Рис. 9.2. Программируемая логическая матрица

1) В качестве вентилях глубины 1 возьмем элементы И, реализующие различные элементарные конъюнкции t_{ij} ($1 \leq i \leq m, 1 \leq j \leq l_i$) ранга 2 и выше (пусть N_A — число таких элементарных конъюнкций, а D_A — сумма их рангов). Входами элементов являются внешние входы, соответствующие буквам, входящим в элементарную конъюнкцию, реализуемую этим элементом.

2) Каждой дизъюнктивной нормальной форме F_i с $l_i \geq 2$ сопоставим элемент ИЛИ (пусть N_0 — число таких форм F_i). В качестве входов элемента ИЛИ возьмем: а) выходы элементов И, реализующие те из элементарных конъюнкций $t_{i1}, t_{i2}, \dots, t_{il_i}$, которые имеют ранг 2 или более; б) внешние входы, соответствующие тем элементарным конъюнкциям $t_{i1}, t_{i2}, \dots, t_{il_i}$, которые имеют ранг 1, т. е. являются буквами. На выходе этого элемента ИЛИ реализуется F_i (а точнее, функция, представляемая формой F_i).

3) Для каждой формы F_i с $l_i=1$ внешним выходом, реализующим F_i , будем считать: а) выход элемента И, реализующего элементарную конъюнкцию t_{i1} , если ранг t_{i1} больше или равен 2; б) внешний вход соответствующий конъюнкции t_{i1} , если она имеет ранг 1, т. е. является буквой.

Таким образом строится И—ИЛИ-схема глубины 2, реализующая совокупность F_1, F_2, \dots, F_m . По определению,

$$N_G = N_A + N_0, \quad N_I = D_A + \sum_{l_i \geq 2} l_i, \quad (9.3), (9.4)$$

и описанная схема имеет сложность

$$c_G(N_A + N_0) + c_I(D_A + \sum_{l_i \geq 2} l_i). \quad (9.5)$$

Точно так же можно построить ИЛИ—И-схему глубины 2, реализующую произвольную заданную совокупность конъюнктивных нормальных форм.

Пример 9.2. Программируемая логическая матрица (ПЛМ) [21]. Решетка с p входами и q выходами из элементов И определяется матрицей соединений $c_{ih} \in V_1 (1 \leq i \leq p, 1 \leq h \leq q)$ в узлах решетки; она реализует следующие выходные функции:

$$g_h = (z_1 \vee \bar{c}_{1h}) (z_2 \vee \bar{c}_{2h}) \dots (z_p \vee \bar{c}_{ph}),$$

где z_1, z_2, \dots, z_p — входные переменные. Программируемой логической матрицей (ПЛМ) называется изображенная на рис. 9.2 схема, получающаяся соединением решетки с $2n$ входами и k выходами из элементов И (пусть a_{ih} — матрица ее соединений) и решетки с k входами и m выходами из элементов И (с матрицей соединений b_{hj}). С помощью правила Де Моргана получаем, что при всех $j (1 \leq j \leq m)$

$$f_j = \bigvee_{h=1}^k b_{hj} (x_1 \vee \bar{a}_{1h}) (\bar{x}_1 \vee \bar{a}_{2h}) \dots (x_n \vee \bar{a}_{2n-1, h}) (\bar{x}_n \vee \bar{a}_{2n, h}).$$

Таким образом при соответствующем выборе матриц соединений (a_{ih}) и (b_{hj}) одновременно можно реализовать m произвольных дизъюнктивных нормальных форм, содержащих не больше k различных элементарных конъюнкций переменных x_1, x_2, \dots, x_n . Сложность схемы зависит только от n, k и m . Если функцию не удастся реализовать с помощью k элементарных конъюнкций, то можно попытаться использовать несколько ПЛМ.

Таким образом, сложность схемы, реализующей совокупность дизъюнктивных нормальных форм, зависит от схем, которые используются для реализации.

9.2. Метод Квайна-Мак-Класски [12, 17—19]

Определение $f^{(0)}, f^{(1)}$. Для частичной функции $f \in \mathcal{B}'_n$ определим две функции, $f^{(0)}$ и $f^{(1)}$, из \mathcal{B}_n , которые совпадают с функцией f в области определения D функции f и принимают соответственно значения 0 и 1 везде в области неопределенности функции f .

Проблема минимизации дизъюнктивных (конъюнктивных) нормальных форм заключается в нахождении для заданных частичных функций $f_1, f_2, \dots, f_m \in \mathcal{B}'_n$ так называемых *простейших* дизъюнктивных (конъюнктивных) нормальных форм F_1, F_2, \dots, F_m , которые имеют минимальную сложность $c(F_1, F_2, \dots, F_m)$ и удовлетворяют условиям

$$f_i^{(0)} \leq F_i \leq f_i^{(1)}. \quad (9.6)$$

Вообще говоря, при этом существуют входы, которыми можно пренебречь.

Упражнение 9.1. Будем считать, что сложность определяется таким образом, что любая \mathcal{L}_B -схема из функциональных элементов и двойственная ей схема имеют одну и ту же сложность. F_1, F_2, \dots, F_m являются конъюнктивными нормальными формами минимальной сложности, удовлетворяющими условиям (9.6), тогда и только тогда, когда их двойственные дизъюнктивные нормальные формы $F_1^d, F_2^d, \dots, F_m^d$ имеют минимальную сложность и удовлетворяют условиям *

$$(f_i^{(1)})^d \leq F_i^d \leq (f_i^{(0)})^d. \quad (9.7)$$

Решение. ИЛИ—И-схема N глубины 2, двойственная для И—ИЛИ-схемы N^d глубины 2 (или ПЛМ), реализующей $F_1^d, F_2^d, \dots, F_m^d$, реализует F_1, F_2, \dots, F_m (см. доказательство упражнения 8.6). По предположению, обе эти схемы имеют одинаковую сложность. Из теоремы 8.2 и задачи 7.42 получаются неравенства (9.7). С другой стороны, из минимальности сложности одной схемы следует минимальность сложности другой.

Таким образом, достаточно рассмотреть задачу минимизации дизъюнктивных нормальных форм. Эту задачу можно разделить на следующие две части: S1) до тех пор, пока это возможно, исследование конъюнкций, входящих в F_i , и отбрасывания тех из них, которые покрываются другими конъюнкциями **;

S2) выбор из оставшихся кандидатов тех, которые минимизируют сложность.

В этом параграфе рассматривается первая из этих задач. Относительно сложности дизъюнктивных нормальных форм будем предполагать следующее.

Предположения о сложности. Пусть F_1, F_2, \dots, F_m и F'_1, F'_2, \dots, F'_m — две совокупности дизъюнктивных нормальных форм. Будем предполагать, что для каждого i

$$1) F_i = \bigvee_{j=1}^{l_i} t_{ij}, \quad F'_i = \bigvee_{j=1}^{l'_i} t'_{ij};$$

$$2) \text{ если } l'_i \leq l_i \text{ и } t_{ij} \leq t'_{ij} (1 \leq j \leq l'_i), \text{ то } c(F'_1, F'_2, \dots, F'_m) \leq c(F_1, F_2, \dots, F_m).$$

Определение простой импликанты. Конъюнкция t называется

* $(f_i^{(1)})^d, (f_i^{(0)})^d$ — формулы двойственные $f_i^{(1)}, f_i^{(0)}$.

** Точный смысл (см. далее).

простой импликантой функции f , или простой f -импликантой*, если 1) $t < f$; 2) не существует конъюнкции t' такой, что $t < t' < f$.

Конъюнкция, удовлетворяющая условию (1), называется импликантой функции f , или f -импликантой. Множество всех импликант функции f обозначается через $T(f)$. Минимальная импликанта в $T(f)$ называется минимальной f -импликантой. Простая f -импликанта является максимальным элементом $T(f)$.

Лемма 9.1. Предположим, что дизъюнктивные нормальные формы F_1, F_2, \dots, F_m удовлетворяют условию (9.6), конъюнкция t входит в $F_{i_1}, F_{i_2}, \dots, F_{i_l}$ и не входит в остальные дизъюнктивные нормальные формы. Если t не является простой $f^{(1)}_{i_1} \cdot f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}$ импликантой, то существует простая $f^{(1)}_{i_1} \cdot f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}$ -импликанта t' такая, что $t < t'$. Пусть F'_1, F'_2, \dots, F'_m — дизъюнктивные нормальные формы, которые получаются из F_1, F_2, \dots, F_m заменой конъюнкции t на t' . Тогда

$$f^{(0)}_i \leq F'_i \leq f^{(1)}_i, \quad 1 \leq i \leq m; \quad (9.8)$$

$$c(F'_1, F'_2, \dots, F'_m) \leq c(F_1, F_2, \dots, F_m). \quad (9.9)$$

Доказательство. Поскольку t не является простой импликантой, то, по определению, существует конъюнкция t' такая, что

$$t < t' \leq f^{(1)}_{i_1} \cdot f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}. \quad (9.10)$$

Так как существует не более чем конечное число таких конъюнкций, то в качестве t' можно взять максимальную из конъюнкций, удовлетворяющих (9.10), т. е. простую $f^{(1)}_{i_1} \cdot f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}$ -импликанту. Из (9.10) и соотношения (3.2) табл. 7.2 имеем $t < t' \leq f^{(1)}_{i_j}$, $1 \leq j \leq l$.

Отсюда, из (9.6) и соотношения (3.2) табл. 7.2 получаем

$$f^{(0)}_{i_j} \leq F_{i_j} \leq F_{i_j} \vee t' \leq F'_{i_j}, \quad 1 \leq j \leq l.$$

Из неравенства $t < t'$ и соотношения (3.4) табл. 7.2 следует, что $t \vee t' = t'$ и $F'_{i_j} = F_{i_j} \vee t'$. Так как $F'_i = F_i$ для F_i , не содержащих t , то неравенства (9.8) доказаны. Из неравенства $t < t'$, определения F'_i и предположений о сложности следует (9.9).

Простую $f^{(1)}_{i_1} f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}$ -импликанту, не являющуюся $f^{(1)}_j$ -импликантой ни для одного $j \neq i_h$ ($1 \leq h \leq l$), далее будем называть p -простой $f^{(1)}_{i_1} f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}$ -импликантой. Из приведенной леммы следует, что в качестве конъюнкций достаточно рассматривать только p -простые $f^{(1)}_{i_1}$ -импликанты, p -простые $f^{(1)}_{i_2}$ -импликанты, ..., p -простые $f^{(1)}_{i_1} f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_l}$ -импликанты, ..., p -простые $f^{(1)}_{i_1} \cdot f^{(1)}_{i_2} \cdot \dots \cdot f^{(1)}_{i_m}$ -импликанты.

* В терминах теории высказываний неравенство $f \geq g$ означает, что g предполагает f . Английский глагол *imply* (подразумевать, предполагать) положен в основу этого термина.

Задача 9.1. В задаче S1 в качестве кандидатов конъюнкций нет необходимости рассматривать p -простые $f^{(1)}_{i_1} f^{(1)}_{i_2} \dots f^{(1)}_{i_l}$ - конъюнкции t , если $f^{(0)}_{i_j} = 0 (1 \leq j \leq l)$ всегда, когда $t = 1$.

Решение. Конъюнкцию t можно рассматривать в качестве кандидата, входящего в дизъюнктивные нормальные формы $F_{i_j} (1 \leq i \leq l)$ из (9.6), но если t исключить из F_{i_j} , то неравенства (9.6) по-прежнему останутся справедливыми, и рассматривать t в качестве кандидата не имеет смысла. Если $t = 1$, то $f^{(1)}_{i_j} = 1$, и такие простые импликанты являются конъюнкциями, принимающими значение 1 только на наборах значений входов, которые для выходов с номерами i_1, i_2, \dots, i_j несущественны.

9.2.1. Метод, основанный на согласователях

Рассмотрим метод нахождения всех простых импликант функции f , заданной одной из ее дизъюнктивных нормальных форм (если функция задана таблицей ее значений, то в качестве дизъюнктивных нормальных форм можно взять ее совершенную д.н.ф. На практике по смыслу задачи часто удается построить сравнительно просто булеву формулу без построения таблицы значений функции и далее преобразовать ее к дизъюнктивной нормальной форме). Множество всех простых импликант функции f обозначим через $PI(f)$.

Определение согласователей. Пусть $t_1 = x_{i_1}^{a_1} x_{i_2}^{a_2} \dots x_{i_r}^{a_r}$ и $t_2 = x_{j_1}^{b_1} x_{j_2}^{b_2} \dots x_{j_l}^{b_l}$ — элементарные конъюнкции такие, что а) для одной из букв t_1 , например $x_{i_1}^{a_1}$, в t_2 имеется ее отрицание (пусть это будет $x_{i_1}^{b_1}$; при этом $i_1 = j_1, a_1 = \bar{b}_1$; б) никакое из отрицаний $\bar{x}_{i_h}^{a_h}$ букв $x_{i_h}^{a_h}$ из t_1 , отличных от $x_{i_1}^{a_1} (1 < h \leq r)$, не входит в t_2 (буквы $x_{i_h}^{a_h}$ могут как входить, так и не входить в t_2). Произведение t всех различных букв, кроме $x_{i_1}^{a_1}, x_{j_1}^{b_1}$, входящих в t_1 или t_2 , называется *согласователем конъюнкций t_1 и t_2* и обозначается через $t = t_1 \not\subset t_2$.

Если указанные условия не выполняются, то говорят, что согласователь не существует. Например,

$$x_1 x_2 \bar{x}_3 x_4 \not\subset x_1 \bar{x}_2 x_5 = x_1 \bar{x}_3 x_4 x_5.$$

С другой стороны, согласователь $x_1 x_2 \bar{x}_3 x_4 \not\subset \bar{x}_1 \bar{x}_2 x_5$ не существует.

Если элементарные конъюнкции t_1 и t_2 одного и того же ранга r содержат по $r-1$ одинаковых букв и по одной букве, которые являются отрицанием друг друга, то говорят, что расстояние между t_1 и t_2 равно 1.

Задача 9.2. Пусть t_1 и t_2 — две минимальные элементарные конъюнкции. Согласователь между t_1 и t_2 существует тогда и только тогда, когда расстояние между t_1 и t_2 равно 1.

Решение. Очевидно из определений.

Задача 9.3. Расстояние между t_1 и t_2 равно 1 тогда и только тогда, когда существует согласователь $t_1 \not\leq t_2 = t$ и, кроме того $t_1 < t$ и $t_2 < t$.

Решение. Достаточность очевидна из определений. Наоборот, пусть существует согласованность $t_1 \not\leq t_2$. Не теряя общности, можно предположить, что $t_1 = x_1 t'_1$, $t_2 = \bar{x}_1 t'_2$. Тогда $t_1 \not\leq t_2 = t'_1 t'_2$. Так как при этом $x_1 t'_1 < t'_1 t'_2$, $\bar{x}_1 t'_2 < t'_1 t'_2$, то из задачи 7.38 следует что $t'_1 \leq t'_2$, $t'_2 \leq t'_1$, т. е. $t'_1 = t'_2$.

Задача 9.4. Если для элементарных конъюнкций t_1 и t_2 существует согласователь $t_1 \not\leq t_2$, то

$$t_1 \not\leq t_2 < t_1 \vee t_2. \quad (9.11)$$

Решение. Поменяв, если нужно, местами t_1 и t_2 , положим $t_1 = x_1 t'_1$, $t_2 = \bar{x}_1 t'_2$, где t'_1 и t'_2 представляют собой либо элементарные конъюнкции, не содержащие букв x_1 и \bar{x}_1 , либо константу 1. Поскольку $t_1 \not\leq t_2 = t'_1 \cdot t'_2$, то (9.11) верно как при $x_1 = 0$, так и при $x_1 = 1$.

Задача 9.5. Согласователь $t_1 \not\leq (t_1 \not\leq t_2)$ не существует.

Решение. Предположим, что $t_1 = x_h t'_1$, $t_2 = x_h t'_2$. Поскольку $t_1 \not\leq t_2 = t'_1 \cdot t'_2$, то $t_1 \not\leq t_2$ не содержит x_h . Так как все буквы t'_1 входят в $t_1 \not\leq t_2$, то согласователь $t_1 \not\leq (t_1 \not\leq t_2)$ существовать не может.

Задача 9.6. Согласователь не удовлетворяет закону ассоциативности.

Решение. $(x_1 x_3 \not\leq \bar{x}_1 x_2) \not\leq \bar{x}_1 \bar{x}_2 \neq x_1 x_3 \not\leq (x_1 x_2 \not\leq \bar{x}_1 \bar{x}_2)$.

Справедлива следующая важная теорема.

Теорема 9.1. Совокупность элементарных конъюнкций $S = \{t_1, t_2, \dots, t_l\}$ совпадает с множеством всех простых импликант функции $f = t_1 \vee t_2 \vee \dots \vee t_l$ тогда и только тогда, когда: 1) для любых $i \neq j$ $t_i \not\leq t_j$; 2) для любых $i \neq j$ либо согласователь $t_i \not\leq t_j$ не существует, либо для некоторого k $t_i \not\leq t_j \leq t_k$.

Доказательство. (а) Пусть $S = PI(f)$. Условие (1) очевидно. Если существует $t_i \not\leq t_j$, то из задачи 9.4 следует, что $t_i \not\leq t_j \leq t_i \vee t_j < f$. Из определения простой импликанты и равенства $S = PI(f)$ следует выполнение условия (2).

(б) Предположим, что выполняются условия (1) и (2); рассмотрим следующее множество T импликант функции f :

$$T = \{t \mid t \leq f, t \not\leq t_i \text{ для всех } 1 \leq i \leq l\}.$$

Предположим, что в S имеется конъюнкция t_j , не являющаяся простой импликантой функции f . Тогда существует простая импликанта t функции f такая, что $t_j < t$. При этом $t \notin S$, так как выполнение неравенства $t_j < t_i$ (если $t \equiv t_i$) для некоторого t_i , противоречило бы условию (1). Заметим, что если простая импликанта t функции f не принадлежит S , то, по определению простой импликанты, $t \in T$, т. е. если $S \neq PI(f)$, то T непусто.

Пусть t_0 — одна из предельных минимальных конъюнкций в T . Если t_0 является минимальной конъюнкцией $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$, то $t_0 \leq t_1 \vee t_2 \vee \dots \vee t_l = f$ и, следовательно, в S существует некоторая конъюнкция t_j , принимающая значение 1 при $x_h = a_h$ ($1 \leq h \leq n$).

При этом $t_0 < t_j$, и мы получили противоречие с определением T . Таким образом, $\deg(t_0) < n$, и, следовательно, существует переменная x_h такая, что ни x_h , ни \bar{x}_h не входят в t_0 . Так как t_0 является минимальным элементом T , то $x_h t_0, \bar{x}_h t_0 \notin T$. Поскольку $x_h t_0 \leq t, \bar{x}_h t_0 \leq f$, то при некоторых i, j

$$x_h t_0 \leq t_i, \quad \bar{x}_h t_0 \leq t_j. \quad (9.12)$$

Если t_i не содержит букву x_h , то, согласно задаче 7.38, $t_0 \leq t_i$. Это противоречит тому, что $t_0 \in T$. Следовательно, t_i содержит x_h . Аналогично доказывается, что t_j содержит \bar{x}_h . Из задачи 7.38 и соотношений (9.12) следует, что все буквы в t_i и t_j , кроме x_h, \bar{x}_h , входят в t_0 . Таким образом, существует согласователь $t_i \not\leq t_j$ и $t_0 \leq (t_i \not\leq t_j)$. Далее воспользуемся условием (2). Существование конъюнкции $t_h \in S$ такой, что $t_0 \leq t_i \not\leq t_j \leq t_h$, противоречит тому, что $t_0 \in T$. Следовательно, $S = PI(f)$.

С помощью этой теоремы и описываемой далее процедуры можно найти все простые импликанты.

Метод нахождения всех простых импликант. Предположим, что задано множество попарно различных элементарных конъюнкций $S_0 = \{t_{10}, t_{20}, \dots, t_{i0}\}$. Положим $f = t_{10} \vee t_{20} \vee \dots \vee t_{i0}$.

(1) Вначале исключаются все те элементарные конъюнкции t_{j0} , для которых существуют элементарные конъюнкции t_{i0} , удовлетворяющие условию $t_{j0} < t_{i0}$.

(2) Если не существует ни одной неупорядоченной пары различных элементарных конъюнкций из S , к которой можно было бы применить описанную далее процедуру Q , то полагаем $S = PI(f)$. Если существует пара элементарных конъюнкций t_i, t_j , к которой процедура Q еще не применялась, то к ней применяется процедура Q .

Процедура Q . Если согласователь $t_i \not\leq t_j$ не существует, то переходим к (2). Если же он существует, то проводим сравнение $t_i \not\leq t_j$ с каждой элементарной конъюнкцией $t_h \in S$. Если при этом находится конъюнкция t_h такая, что $(t_i \not\leq t_j) \leq t_h$, то возвращаемся к (2). В противном случае из S исключаем все конъюнкции t_h такие, что $t_h < (t_i \not\leq t_j)$; присоединяем к S согласователь $t_i \not\leq t_j$ и возвращаемся к (2).

Пусть S_{i-1}, S_i — множество S соответственно до и после i -го применения процедуры Q . Согласно задаче 9.4 и соотношению 3.4 из табл. 7.2

$$\bigvee_{t \in S_{i-1}} t = \bigvee_{t \in S_i} t, \text{ так что оба эти выражения равны}$$

$$f = \bigvee_{t \in S_0} t.$$

После завершения выполнения (2), согласно теореме 9.1, $S = PI(f)$. Предположим, что при выполнении процедуры Q из S_{i-1} была исключена элементарная конъюнкция t . Так как в S_i , а следовательно, и во всех множествах $S_j (i \leq j)$ существует элементарная конъюнкция, которая больше t , то t не может быть присоединена к S на следующих этапах. Положим $T = \{t | t \leq f, t \not\leq t_{i0} (1 \leq i \leq l)\}$. Так как $S_i \subset T$, то число пар конъюнкций, к которым

применима процедура Q , не превышает $|T|(|T|-1)/2$. Время, необходимое для выполнения процедуры Q , пропорционально $c_1|T|$. Таблицы, которые при этом используются, содержат также не более чем $c_2n|T|$ двончных символов (c_1, c_2 — константы).

Пример 9.3. Рассмотрим дизъюнктивную нормальную форму (8.28). Построим для нее табл. 9.1 из трех столбцов, первый из которых указывает имя элементарной конъюнкции, второй — переменные, входящие в эту конъюнкцию, а третий — результат проверки. Если элементарная конъюнкция t содержит букву $x_i(\bar{x}_i)$,

Таблица 9.1. Таблица определения простых импликант

Конъюнкции		
t_1	ε 1 0 0 0 0 0 0	∇
t_2	ε ε 1 0 0 0 0	∇
t_3	ε ε ε ε ε 1 0 0	∇
t_4	ε ε ε ε ε ε ε 1	
$t_1 \nsubseteq t_2$	ε 1 0 ε 0 0 0 0	∇
$t_2 \nsubseteq t_3$	ε ε ε 1 0 ε 0 0	∇
$t_3 \nsubseteq t_4$	ε ε ε ε ε 1 0 ε	
$t_4 \nsubseteq (t_1 \nsubseteq t_2)$	ε 1 0 ε 0 0 0 ε	∇
$t_4 \nsubseteq (t_2 \nsubseteq t_3)$	ε ε ε 1 0 ε 0 ε	
$(t_3 \nsubseteq t_4) \nsubseteq (t_4 \nsubseteq (t_1 \nsubseteq t_2))$	ε 1 0 ε 0 ε 0 ε	

то в строке этой таблицы, соответствующей t (далее эта строка называется просто строкой t), на пересечении со столбцом ставится символ 1 (0); если ни одна из этих букв в t не входит, то в указанной клетке таблицы ставится символ ε . В столбце C ставится знак ∇ , если рассматриваемая строка соответствует элементарной конъюнкции, которая должна быть исключена при выполнении процедуры Q . Для элементарных конъюнкций, которые должны быть добавлены, используются первоначально пустые строки. Элементарным конъюнкциям, входящим в дизъюнктивные нормальные формы (8.28), в порядке их записи в дизъюнктивные нормальные формы присвоим имена t_1, t_2, t_3, t_4 . Для удобства читателей в столбце имен элементарных конъюнкций указывается, с помощью какого согласователя получена соответствующая конъюнкция.

(1) Условием существования согласователя $t \nsubseteq t'$ является наличие среди столбцов x_1, x_2, \dots ровно одного столбца x_i , который в строках t, t' имеет символы 0 и 1 или 1 и 0. Если воспользоваться утверждением задачи 9.5, то можно несколько облегчить процедуру определения согласователей.

(2) Для строк, не содержащих знака \vee в столбце C , проверяется существование согласователей, при этом вначале рассматриваются строки, расположенные выше. Если согласователь существует (пусть это будет t), то строится строка t , и далее она последовательно сравнивается с вышерасположенными строками t' , которые не отмечены знаком \vee в столбце C . Если при этом оказывается, что $t' < t$, то у строки t' в столбце C ставится знак \vee . Если же $t \leq t'$, то строка t стирается.

(3) После выполнения указанных операций строки таблицы, не имеющие знака \vee в столбце C , будут простыми импликантами. В данном примере простыми импликантами являются $x_2\bar{x}_3\bar{x}_5\bar{x}_7$, $x_4\bar{x}_5\bar{x}_7$, $x_6\bar{x}_7x_8$.

Задача 9.7. Покажите, что простыми импликантами функции f с $D_1(f) = \{0, 2, 4, 6, 10, 11, 14, 29\}$ являются $x_1x_2x_3\bar{x}_4x_5$, $\bar{x}_1x_2\bar{x}_3x_4$, $\bar{x}_1\bar{x}_2\bar{x}_5$, $\bar{x}_1x_4\bar{x}_5$. Получите эти простые импликанты с помощью таблиц Карно приводимых далее.

Задача 9.8. Импликанта t функции f является простой тогда и только тогда, когда не существует ни одной импликанты функции f , находящейся от t на расстоянии 1.

Решение. Если t не простая импликанта, то существует импликанта t_0 такая, что $t < t_0 \leq f$. Из задачи 7.38 вытекает, что в этом случае существует буква x_i^* , которая входит в t , но не входит в t_0 . Пусть t' — это конъюнкция, которая получается из t при замене в последней буквы x_i^* на ее отрицание. При этом $t' < t_0 \leq f$, т. е. является импликантой функции f . Наоборот, если t простая импликанта, то, согласно задаче 9.3, не может существовать импликанты функции f , находящейся на расстоянии 1 от t .

Задача 9.9. Функции f, \bar{f} не имеют общих простых импликант, кроме минимальных конъюнкций, тогда и только тогда, когда $f = c \oplus x_1 \oplus x_2 \oplus \dots \oplus x_n$ ($c = 0$ или 1).

Указание. Предположим, что функции f, \bar{f} не имеют общих простых импликант, кроме минимальных конъюнкций. Воспользовавшись утверждением задачи 9.8, покажите, что $f(B) = \bar{f}(A)$ для произвольного набора $A \in V_n$ и всех наборов $B \in V_n$ таких, что $d(A, B) = 1$. Отсюда будет следовать, что для всех наборов A , имеющих четный (нечетный) вес $w(A)$, значения функции будут равны (для доказательства воспользуйтесь индукцией по $w(A)$). Наоборот, так как у $p_n(\bar{p}_n)$ не может быть минимальных конъюнкций, находящихся друг от друга на расстоянии 1, то из задачи 9.2 и теоремы 9.1 следует, что все минимальные конъюнкции являются простыми импликантами и, кроме них, других простых импликант нет.

Упражнение 9.2. Для $f \in \mathcal{F}_n$ положим $mD_1(f) = \{A \mid A \in V_n, f(A) = 1 \text{ и } f(B) = 0 \text{ для всех } B \text{ таких, что } B < A\}$; $mD_1(f)$ является множеством предельных минимальных элементов $D_1(f)$. Пусть $A = (a_1, a_2, \dots, a_n) \in V_n$, $a_i = a_i, \dots = a_i, = 1, a_i = 0, i \neq i_j (1 \leq j \leq r)$.

Набору A сопоставим элементарную конъюнкцию $t_m(A) = x_{i_1}x_{i_2}\dots x_{i_r}$ ранга r . Докажите, что если f является монотонной функцией, то $PI(f) = \{t_m(A) \mid A \in mD_1(f)\}$.

(9.13)

Решение. Предположим, что $t_m(A) \leq t_m(B)$ для некоторых наборов $A, B \in mD_1(f)$. Из задачи 7.38 и определения t_m следует, что $B < A$. По определению $mD_1(f)$, имеем $f(A) = 1, f(B) = 0$. Однако это противоречит тому, что $B \in mD_1(f)$.

Следовательно, множество S в правой части (9.13) удовлетворяет условию (1) теоремы 9.1. Так как $t_m(A)$ не содержит отрицаний переменных, то никакие две элементарные конъюнкции из S не имеют согласователя, т. е. выполняется также условие (2) теоремы 9.1.

Метод нахождения простых импликант при нескольких выходах. Как следует из леммы 9.1, при переходе к задаче с многими выходами возникает необходимость определения p -простых импликант функций $f^{(1)}_i (1 \leq i \leq m)$, $f^{(1)}_{i_1} f^{(1)}_{i_2} (1 \leq i_1 < i_2 \leq m)$, ..., $f^{(1)}_1 f^{(1)}_2 \dots f^{(1)}_m$. Поскольку при независимом их построении эффективность оказывается обычно невысокой, то лучше их строить одновременно, например, методом таблиц Мак-Класки.

Пример 9.4. Пусть $f_1, f_2 \in \mathcal{B}_3$, $D_1(f_1) = \{1, 3, 7\}$, $D_1(f_2) = \{2, 6, 7\}$. Как видно из табл. 9.2, в отличие от табл. 9.1, в данном случае в таблицу дополнительно вводятся столбцы f_1, f_2 .

Таблица 9.2. Таблица определения простых импликант в случае нескольких выходов

Имя конъюнкции	x_1	x_2	x_3	t_1	t_2	C
1	0	0	1	—	0	\bigvee \bigvee \bigvee \bigvee \bigvee
2	0	1	0	0	—	
3	0	1	1	—	0	
6	1	1	0	0	—	
7	1	1	1	—	—	
$1 \not\leq 3$	0	ε	1	—	0	
$2 \not\leq 6$	ε	1	0	0	—	
$7 \not\leq 1 \not\leq 3$	ε	1	1	—	0	
$8 \not\leq 2 \not\leq 6$	1	1	ε	0	—	

(1) Вначале строится таблица для различных элементарных конъюнкции, входящих в дизъюнктивные нормальные формы функций $f_1, f_2, f_1 f_2$. Далее выполняются те же операции, что и при $m=1$. При этом имеется ряд особенностей (см. далее). В данном примере построение начинается с минимальных конъюнкции 1, 2, 3, 6, 7 функции $f_1 \vee f_2$.

(2) В строке t и столбце f_i ставится знак «—», если $t < f_i$, и 0 в противном случае. Если согласователь $t_1 \not\leq t_2$ существует, то в соответствующей ему строке и столбце f_i ставится знак «—», когда соответствующие t_1 и t_2 строки столбца f_i обе имеют знак «—», и 0 в остальных случаях. Не имеет смысла рассматривать конъюнкции, соответствующие которым строки имеют во всех столбцах f_i символ 0.

(3) Будем говорить, что строка t_2 покрывает строку t_1 , если эти строки отличаются хотя бы в одном из столбцов таблицы, отличных от столбца имен конъюнкции, $t_1 < t_2$, а кроме того, для всех i таких, что строка t_2 в столбце f_i имеет 0, строка t_1 в столбце f_i также имеет 0. Если для вновь построенной строки в таблице уже имеется строка, которая с ней совпадает (имя конъюнкции при сравнении во внимание не принимается) или которая ее покрывает, то строка t стирается. Наоборот, если t покрывает некоторую ранее построенную строку t' , то у последней в столбце C ставится знак \vee .

(4) После завершения выполнения всех операций строки t , не содержащие знака \vee , будут простыми импликантами для произведения всех функций, ко-

которые соответствуют столбцам, имеющим знак «—» в этой строке. В данном примере p -простыми импликантами являются: \bar{x}_1x_3 , $x_2\bar{x}_3$ для f_1 , x_2x_3 , x_1x_2 для f_2 , $x_1x_2x_3$ для f_1f_2 .

Если построить дизъюнктивные нормальные формы минимальной сложности для функций f_1 , f_2 независимо, то полученные в результате простые импликанты не будут совпадать с самими функциями f_1 , f_2 . Отсюда и из леммы 9.1 следует, что $f_1 = \bar{x}_1x_3 \vee x_2x_3$; $f_2 = x_1x_2 \vee x_2\bar{x}_3$.

На рис. 9.3,а приведена И—ИЛИ-схема глубины 2, реализующая эти функции. На рис. 9.3,б приведена другая И—ИЛИ-схема глубины 2, также реализующая функции f_1 , f_2 . Вторая схема по сравнению с первой более экономна как по числу элементов, так и по суммарному числу входов. Она использует p -простую импликанту $x_1x_2x_3$ функции f_1f_2 .

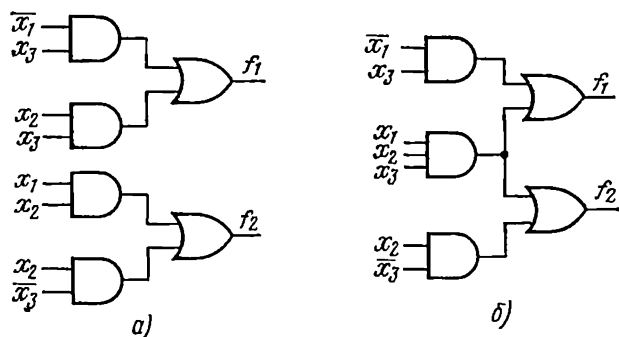


Рис. 9.3. Схемы из примера 9.4

Задача 9.10. Пусть $D_1(f_1) = \{0, 1, 3, 7\}$, $D_1(f_2) = \{1, 3, 4, 5, 7\}$. Найдите p -простые импликанты функций f_1 , f_2 , f_1f_2 .

Решение. В данном случае p -простыми импликантами являются: $\bar{x}_1\bar{x}_2$ для функции f_1 , $x_1\bar{x}_2$, x_3 для функции f_2 , \bar{x}_1x_3 , x_2x_3 для функции f_1f_2 .

Задача 9.11. Найдите простые импликанты для кодера приоритета из примера 8.5 при $m=3$.

Решение. (1) $GS \geq y_i (1 \leq i \leq 3)$.

(2) Произведение EO с остальными выходами GS , y_1 , y_2 , y_3 равно 0. При этом p -простой импликантой функции EO является $E\bar{x}_1\bar{x}_2 \dots \bar{x}_8$.

(3) $y_1 y_2 = E (5 \leq J_{\max}) [(3 \leq J_{\max} \leq 4) \vee (7 \leq J_{\max} \leq 8)] = E (7 \leq J_{\max} \leq 8) = E (x_7 \bar{x}_8 \vee x_8) = E (x_7 \vee x_8)$.

Аналогично $y_1 y_3 = E [(J_{\max} = 6) \vee (J_{\max} = 8)] = E (x_6 \bar{x}_7 \bar{x}_8 \vee x_8) = E (x_6 \bar{x}_7 \vee x_8)$, $y_2 y_3 = E [(J_{\max} = 4) \vee (J_{\max} = 8)] = E (x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \bar{x}_8 \vee x_8) = E (x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \vee x_8)$, $y_1 y_2 y_3 = E (J_{\max} = 8) = E x_8$.

(4) Из сказанного ранее, соотношений (8.26) — (8.29) и теоремы 9.1 следует, что p -простыми импликантами являются: $E x_1$, $E x_2$, $E x_3$, $E x_4$ для функции GS ; $E x_5$, $E x_6$ для $GS y_1$; $E x_3 \bar{x}_5 \bar{x}_6$, $E x_4 \bar{x}_5 \bar{x}_6$ для $GS y_2$; $E x_2 \bar{x}_3 \bar{x}_5 \bar{x}_7$,

$Ex_4x_5\bar{x}_7$ для GSy_3 ; Ex_7 для GSy_1y_2 ; $Ex_6\bar{x}_7$ для GSy_1y_3 ; $Ex_4\bar{x}_5\bar{x}_6\bar{x}_7$ для GSy_2y_3 ; Ex_8 для $GSy_1y_2y_3$.

9.2.2. Таблица Карно и коды с единичным расстоянием

При геометрическом представлении булевых функций (см. § 7.3) множество черных кружочков, соответствующих элементарным конъюнкциям $x_{i_1}^{a_1}x_{i_2}^{a_2}\dots x_{i_r}^{a_r}$ ранга r данной булевой функции, представляет собой некоторое подмножество вершин n -мерного единичного куба. Например, на рис. 9.4

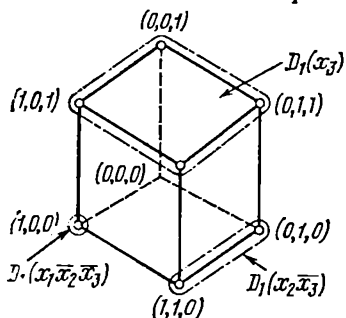


Рис. 9.4. Представление элементарных конъюнкций на единичном кубе

при $n=3$ на единичном кубе представлены элементарные конъюнкции x_3 , $x_2\bar{x}_3$, $x_1\bar{x}_2\bar{x}_3$. Множеству $D_1(x_3)$ соответствует плоскость, множеству $D_1(x_2, \bar{x}_3)$ — ребро, а $D(x_1x_2x_3)$ — вершина. Благодаря этому свойству можно записывать таблицы значений функции, используя понятие соседних элементов V_n (элементы являются соседними, если расстояние Хэмминга между ними равно 1). Типичным примером такой записи являются таблицы Карно, в которых ребра и плоскости, понимае-

Таблица 9.3. Таблицы Карно

а) $n=2$

x_1	x_2	
	0	1
0	$f(0,0)$	$f(0,1)$
1	$f(1,0)$	$f(1,1)$

б) $n=3$

x_1	$x_2 \ x_3$			
	0 0	0 1	1 1	1 0
0	$f(0,0,0)$	$f(0,0,1)$	$f(0,1,1)$	$f(0,1,0)$
1	$f(1,0,0)$	$f(1,0,1)$	$f(1,1,1)$	$f(1,1,0)$

в) $n=4$

$x_1 \ x_2$		$x_3 \ x_4$			
		0 0	0 1	1 1	1 0
0 0					
0 1					
1 1					
1 0					

г) $n=5$

$x_1 \ x_2$		$x_3 \ x_4 \ x_5$							
		0 0 0	0 0 1	0 1 1	0 1 0	1 1 0	1 1 1	1 0 1	1 0 0
0 0									
0 1									
1 1									
1 0									

мые в указанном смысле, помещаются в «легко обозримые» места таблицы. Таблица 9.3 представляет собой таблицы Карно при числе переменных от 2 до 5. Например, при $n=4$ в клетке таблицы, расположенной в строке 01 и столбце 10, записывается $f(0, 1, 1, 0)$. В трех последовательностях 0,1; 00, 01, 11, 10; 000, 001, 011, 010, 110, 111, 101, 100 элементов из V_1, V_2 или V_3 расстояние Хэмминга между соседними элементами равно 1; при этом расстояние Хэмминга между первым и последним элементами также равно 1. Следовательно, расстояние Хэмминга между элементом V_n , соответствующим некоторой клетке таблицы и клетке, расположенной слева, справа, сверху или снизу от нее, равно 1. При $n \leq 4$ верно также обратное утверждение, а именно, если двум клеткам таблицы соответствуют элементы V_n , расположенные на расстоянии 1, то эти клетки являются соседними. При $n \geq 5$ это утверждение неверно ни для одной таблицы (число элементов V_n , находящихся на расстоянии Хэмминга 1 от заданного элемента, равно n , а в таблице у каждой клетки имеется только четыре соседние клетки). Клетки, соответствующие наборам, которыми можно пренебречь, т. е. запрещенным набором, обозначаются либо буквой d , либо знаком \times .

Пример 9.5. Таблица 9.4 представляет собой таблицы Карно конъюнкций $x_1, x_3, \bar{x}_3, x_2x_3, \bar{x}_1\bar{x}_3$ от трех переменных. Используя описанные свойства, можно сравнительно просто определить простые импликанты функции f при $n < 5$.

Таблица 9.4. Примеры таблиц Карно, представляющих конъюнкции

x_1	x_2	x_3	\bar{x}_3
0	0	0	0
0	1	1	1
1	1	1	1
1	0	0	1
1	0	0	1
0	0	1	0
0	0	1	0

Пример 9.6. Таблица 9.5 является таблицей Карно функции $f = \bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4 \vee \bar{x}_1\bar{x}_2\bar{x}_3x_4 \vee \bar{x}_1x_2\bar{x}_3x_4 \vee x_1\bar{x}_2\bar{x}_3\bar{x}_4 \vee x_1\bar{x}_2x_3\bar{x}_4 \vee x_1x_2\bar{x}_3x_4 \vee x_1x_2x_3\bar{x}_4$. Простыми импликантами функции f являются $\bar{x}_1\bar{x}_2\bar{x}_3, \bar{x}_1\bar{x}_3x_4, \bar{x}_2\bar{x}_3x_4, x_2\bar{x}_3x_4, x_1\bar{x}_2\bar{x}_4, x_1x_3\bar{x}_4, x_1x_2x_4, x_1x_2x_3$.

Таблица 9.5.
К примеру 9.6.

	00	01	11	10
00	1	1	0	0
01	0	1	0	0
11	0	1	1	1
10	1	0	0	1

Таблица 9.6.
К примеру 9.7.

	00	01	11	10
00	1	d	d	1
01	0	0	0	0
11	0	d	d	1
10	1	1	d	1

Пример 9.7. Простые импликанты функции, заданной табл. 9.6 Карно (в данном случае имеется вход, которым можно пренебречь), как следует из леммы 9.1 и определения $f^{(1)}$, можно получить, положив $d=1$ (простую импликанту, состоящую только из одной буквы d , согласно задаче 9.1, необходимо исключить). Простыми импликантами являются $\bar{x}_2, x_1x_3, x_1x_4$.

Определение кода с единичным расстоянием. Взаимно однозначное отображение φ множества I_n целых чисел от 0 до 2^n-1 на V_n называется *кодом с единичным расстоянием*, если оно удовлетворяет условию $U1$: для всех целых положительных чисел $i \leq 2^n-1$ $d(\varphi(i-1), \varphi(i))=1$. Этот код называется *циклическим кодом с единичным расстоянием*, если он удовлетворяет также условию $U2$: $d(\varphi(0), \varphi(2^n-1))=1$.

Если V_n рассматривать как множество вершин n -мерного единичного куба, то в графе G_n , вершинами и ребрами которого являются соответственно вершины и ребра n -мерного куба, по циклическому коду с единичным расстоянием φ можно построить гамильтонов цикл с вершинами $\varphi(0), \varphi(1), \dots, \varphi(2^n-1), \varphi(0)$. Верно и обратное утверждение, а именно, если в G_n задан гамильтонов цикл, то по нему можно построить код с единичным расстоянием. В таблицах Карно по горизонтали и вертикали используются циклические коды с единичным расстоянием. Коды с единичным расстоянием используются также при преобразовании аналоговых величин в дискретные. В частности, циклические коды с единичным расстоянием используются при квантовании угла поворота [19].

Задача 9.12. Пусть $\varphi^{(n-1)}$ — отображение I_{n-1} на V_{n-1} , являющееся кодом с единичным расстоянием, и $\varphi_j^{(n-1)}(i)$ — j -я компонента вектора $\varphi^{(n-1)}(i)$. Определим $\varphi^{(n)}$, положив для каждого $0 \leq i \leq 2^n-1$:

$$\varphi^{(n)}(i) = (0, \varphi_1^{(n-1)}(i), \varphi_2^{(n-1)}(i), \dots, \varphi_{n-1}^{(n-1)}(i));$$

$$\varphi^{(n)}(2^n-1-i) = (1, \varphi_1^{(n-1)}(i), \varphi_2^{(n-1)}(i), \dots, \varphi_{n-1}^{(n-1)}(i)).$$

Покажите, что отображение $\varphi^{(n)}$ множества I_n на V_n является циклическим кодом с единичным расстоянием.

Указание. Заметьте, что $d(\varphi^{(n)}(2^n-1), \varphi^{(n)}(2^n-1))=d(\varphi^{(n)}(2^n-1), \varphi^{(n)}(0))=1$.

Горизонтальные и вертикальные последовательности в таблицах Карно 9.3 строятся последовательно описанным ранее способом из равенств $\varphi^{(1)}(0)=0, \varphi^{(1)}(1)=1$.

Задача 9.13. В V_5 имеется пять элементов, находящихся на расстоянии 1 от элемента, соответствующего заданной клетке таблицы Карно с $n=5$. Четырём из этих элементов соответствуют клетки таблицы Карно, расположенные слева, справа, сверху и снизу от исходной клетки. Какая клетка соответствует пятому из упомянутых элементов?

Указание. Рассмотрите метод построения кодов с единичным расстоянием, описанный в предыдущей задаче.

Среди кодов с единичным расстоянием наиболее широко известен код Грея*, который строится следующим образом. Каждое целое число i ($0 \leq i < 2^n$) представляется в двоичном коде:

$$i = \sum_{j=1}^n a_j 2^{n-j},$$

и далее полагается

$$\Phi(i) = (a_1, a_1 \oplus a_2, a_2 \oplus a_3, \dots, a_{n-1} \oplus a_n).$$

Это отображение Φ называется *кодом Грея*.

Задача 9.14. Покажите, что код Грея является циклическим кодом с единичным расстоянием.

Решение. Воспользуемся индукцией по n . При $n=1$ утверждение задачи верно. Предположим, что оно верно для $n-1$ при некотором n , и докажем, что оно верно для n . Если при $0 \leq i < 2^{n-1}$ положить

$$i = \sum_{j=1}^n a_j 2^{n-j}, \text{ то } a_1 = 0 \text{ и } 2^n - 1 - i = \sum_{j=1}^n \bar{a}_j 2^{n-j}.$$

Отсюда получаем

$$\Phi(i) = (0, a_2, a_2 \oplus a_3, \dots, a_{n-1} \oplus a_n);$$

$$\Phi(2^n - 1 - i) = (1, 1 \oplus \bar{a}_2, \bar{a}_2 \oplus \bar{a}_3, \dots, \bar{a}_{n-1} \oplus \bar{a}_n) = (1, a_2, a_2 \oplus a_3, \dots, a_{n-1} \oplus a_n).$$

Из задачи 9.12 следует, что это отображение является циклическим кодом с единичным расстоянием.

9.3. Задачи о покрытии множеств [17—19]

9.3.1. Постановка задачи

Пусть S — множество, состоящее из m элементов e_1, e_2, \dots, e_m и S_1, S_2, \dots, S_r — некоторые подмножества множества S . Предположим, что для каждого S_i определено неотрицательное целое число c_i .

Задача. Требуется найти подмножество I множества $\{1, 2, \dots, r\}$ такое, что

$$S = \bigcup_{i \in I} S_i \quad (9.14)$$

и сумма $\sum_{i \in I} c_i$ минимальна.

Эта задача называется задачей о минимальном покрытии и сводится к некоторым комбинаторным задачам. Иногда задачей о минимальном покрытии называется сформулированная ранее задача при $c_1 = c_2 = \dots = c_r = 1$. Семейство подмножеств, удовлетворяющих условию (9.14), называется *покрытием* множества S . Положим $a_{ij} = 1$, если $e_j \in S_i$, и $a_{ij} = 0$ в противном случае. Матрица

* Иногда кодом Грея называют любой код с единичным расстоянием. (Прим. ред.)

(a_{ij}) , имеющая r строк и m столбцов, называется *матрицей инцидентности* семейства S подмножеств S_1, S_2, \dots, S_r . Сформулированная задача эквивалентна следующей задаче целочисленного программирования [31]:

$$\min \sum_{i=1}^r c_i x_i; \quad (9.15)$$

$$\sum_{i=1}^r a_{ij} x_i \geq 1, \quad 1 \leq j \leq m \quad (9.16)$$

(здесь x_i — целочисленные переменные, принимающие значения 0 и 1).

Задача 9.15. Если существует процедура решения задачи о минимальном покрытии (случай $c_1 = c_2 = \dots = c_m = 1$) за время порядка некоторой степени p произведения $m \times r$, то NP -полные задачи разрешимы за полиномиальное время.

Решение. Если известно решение I задачи о минимальном покрытии, то в задаче (а) примера 9.1 (3) можно определить, выполняется ли неравенство $|I| \leq k$. Наоборот, если NP -полные задачи разрешимы за полиномиальное время, то, проверяя последовательно значения $k=1, 2, \dots$, можно найти минимальное значение $|I|$ за время порядка некоторой степени произведения $m \times r$.

Исходя из последней задачи, можно предположить, что не существует общего метода решения задачи о минимальном покрытии за полиномиальное по m, r время. Покажем, как свести задачу о минимальном покрытии к задаче с возможно меньшими значениями m и r . Если это не оговорено особо, в качестве меры сложности совокупности дизъюнктивных нормальных форм $c(F_1, F_2, \dots, F_m)$ далее берется функция (9.5).

Упражнение 9.3. Пусть $f^{(0)} \leq f^{(1)}$ и t_1, t_2, \dots, t_r — простые импликанты $f^{(1)}$. Предположим, что $f^{(0)} \neq t_i (1 \leq i \leq r)$. Пусть I — решение (вообще говоря, неединственное) следующей задачи о минимальном покрытии:

$$S = D_1(f^{(0)}); \quad S_i = D_1'(t_i) \cap S, \quad 1 \leq i \leq r; \quad (9.17)$$

$$c_i = c_0 + c_I (\deg(t_i) + 1) \text{ при } \deg(t_i) \geq 2;$$

$$c_i = c_I \text{ при } \deg(t_i) = 1.$$

$$\text{Тогда } F = \bigvee_{i \in I} t_i \quad (9.18)$$

является дизъюнктивной нормальной формы с минимальной сложностью, удовлетворяющей условию $f^{(0)} \leq F \leq f^{(1)}$.

Решение. Так как $t_i \leq f^{(1)}$, то $F \leq f^{(1)}$. Из леммы 9.1 следует, что среди логических сумм простых импликант функций $f^{(1)}$ существует логическая сумма с минимальной сложностью. Если

$$S = D_1(f^{(0)}) = \bigcup_{i \in I} S_i \subset \bigcup_{i \in I} D_1(t_i),$$

$$\text{то } f^{(0)} \leq \bigvee_{i \in I} t_i.$$

Так как $f^{(0)} \neq t_i$, то $|I| \geq 2$. Следовательно, сложность дизъюнктивных нормальных форм $\bigvee_{i \in I} t_i$ равна $c_G + \sum_{i \in I} c_i$. Это означает, что дизъюнктивные нормальные формы (9.18) имеют минимальную сложность.

Таблицы простых импликант. Матрицы инцидентности задачи о минимальном покрытии (9.17) называются таблицами простых импликант.

Т а б л и ц а 9.7. Таблица простых импликант из примера 9.8

Простые импликанты функции f	$D_1(f)$	Сложность
	0 2 4 6 10 11 14 29	
$t_1 = x_1 x_2 x_3 \bar{x}_4 x_5$	1	$c_G + 6c_I$
$t_2 = \bar{x}_1 x_2 \bar{x}_3 x_4$	1 1	$c_G + 5c_I$
$t_3 = \bar{x}_1 \bar{x}_2 \bar{x}_5$	1 1 1 1	$c_G + 4c_I$
$t_4 = \bar{x}_1 x_4 \bar{x}_5$	1 1 1 1	$c_G + 4c_I$

Пример 9.8. Таблица 9.7 является таблицей простых импликант функции f из задачи 9.7. При этом

$$S = D_1(f) = \{0, 2, 4, 6, 10, 11, 14, 29\};$$

$$S_1 = D_1(x_1 x_2 x_3 \bar{x}_4 x_5) = \{29\};$$

$$S_2 = D_1(\bar{x}_1 x_2 \bar{x}_3 x_4) = \{10, 11\};$$

$$S_3 = D_1(\bar{x}_1 \bar{x}_2 \bar{x}_5) = \{0, 2, 4, 6\};$$

$$S_4 = D_1(\bar{x}_1 x_4 \bar{x}_5) = \{2, 6, 10, 14\}.$$

В этой и других подобных таблицах символы 0 не ставятся и соответствующие им клетки оставляются пустыми. В крайнем правом столбце таблицы указана сложность каждой простой импликанты.

Пример 9.9. Таблица 9.8 является таблицей простых импликант функции f из примера 9.6. Все простые импликанты в данном случае имеют одинаковую сложность, поэтому в таблице она не указывается.

У п р а ж н е н и е 9.4. Определение существования или несуществования дизъюнктивной нормальной формы, которая удовлетворяет условию (9.6) и при этом может быть реализована с помощью ПЛМ из примера 9.2, сводится к решению следующей задачи о тождественности. Пусть

$$S = \{(h, j) | j \in D_1(f_h^{(0)}), 1 \leq h \leq m\}$$

Т а б л и ц а 9.8. Таблица простых импликант из примера 9.9

Простые импликанты функции f	$D_1(f)$
	0 1 5 8 10 13 14 15
$t_1 = \bar{x}_1 \bar{x}_2 \bar{x}_3$	1 1
$t_2 = \bar{x}_1 \bar{x}_3 x_4$	1 1
$t_3 = \bar{x}_2 \bar{x}_3 \bar{x}_4$	1 1 1
$t_4 = x_2 \bar{x}_3 x_4$	1 1 1
$t_5 = x_1 \bar{x}_2 \bar{x}_4$	1 1 1
$t_6 = x_1 x_3 \bar{x}_4$	1 1 1
$t_7 = x_1 x_2 x_4$	1 1 1
$t_8 = x_1 x_2 x_3$	1 1 1

и t_1, t_2, \dots, t_r — p -простые импликанты функций $f^{(1)}_1, f^{(1)}_2, \dots, f^{(1)}_m$, $f^{(1)}_1 f^{(1)}_2, \dots, f^{(1)}_{i_1} f^{(1)}_{i_2} \dots f^{(1)}_{i_l}, \dots, f^{(1)}_1 f^{(1)}_2 \dots f^{(1)}_m$, перенумерованные целыми числами от 1 до r . Если t_i — простая импликанта функции $f^{(1)}_{i_1} f^{(1)}_{i_2} \dots f^{(1)}_{i_l}$, то положим

$$S_i = \{(i_h, j) | j \in D_1(f^{(0)}_{i_h} \cdot t_i), 1 \leq h \leq l\}, \quad 1 \leq i \leq r.$$

Спрашивается, существует ли подмножество I множества $\{1, 2, \dots, r\}$ такое, что

$$S = \bigcup_{i \in I} S_i, \quad |I| \leq k. \quad (9.19)$$

Решение. Как следует из леммы 9.1, в данном случае достаточно рассмотреть лишь простые импликанты. Для произвольного подмножества I множества $\{1, 2, \dots, r\}$ определим F_h как логическую сумму всех p -простых импликант t_i функций $f^{(1)}_{i_1} f^{(1)}_{i_2} \dots f^{(1)}_{i_l}$, таких, что $i \in I$. При этом $F_h \leq f^{(1)}_h$. С другой стороны, по определению S и S_i , имеем

$$S = \bigcup_{i \in I} S_i \iff f^{(0)}_h \leq F_h \quad (1 \leq h \leq m).$$

Таким образом, сформулированная задача о тождественности эквивалентна определению существования совокупности дизъюнктивных нормальных форм, удовлетворяющих условию (9.6) и содержащих не более k элементарных конъюнкций.

Матрица смежности подмножеств S_1, S_2, \dots, S_r введенного выше множества S называется *таблицей простых импликант в случае нескольких выходов*.

9.3.2. Упрощение матриц инцидентности (таблиц простых импликант)

Существенные строки. Если в матрице инцидентности в некотором столбце j имеется ровно одна единица, расположенная в i -й строке, то эта строка i называется *существенной строкой*. Поскольку S_i является единственным подмножеством в S , содержащим элемент e_j , то условие (9.14) может выполняться только в том случае, если $i \in I$. Другими словами, S_i всегда необходимо включать в покрытие. Простые импликанты, которым в таблице простых импликант соответствуют существенные строки, называются существенными конъюнкциями. Другими словами, если имеется только одна простая импликанта t' функции $f^{(1)}$ такая, что $t \leq t'$ для некоторой минимальной конъюнкции t функции $f^{(0)}$, то t' называется *существенной конъюнкцией*. Например, все простые импликанты в табл. 9.7 являются существенными. Следовательно,

$$F = x_1 x_2 x_3 \bar{x}_1 x_5 \vee \bar{x}_1 x_2 \bar{x}_3 x_4 \vee \bar{x}_1 \bar{x}_2 \bar{x}_5 \vee \bar{x}_1 x_4 \bar{x}_5$$

является дизъюнктивной нормальной формы с минимальной сложностью, удовлетворяющей условию $f = F$.

Задача 9.16. Найдите дизъюнктивную нормальную форму минимальной сложности, удовлетворяющую условию $f^{(0)} \leq F \leq f^{(1)}$,

если $D_1(f^{(1)}) = \{0, 2, 4, 6, 10, 11, 14, 29\}$ (та же функция, что и в задаче 9.7) и $D_1(f^{(0)}) = \{0, 2, 4, 6, 10, 29\}$. Считается, что $c_I > 0$.

Указание. Таблица простых импликант получается, если из табл. 9.7 удалить столбцы, соответствующие числам 11 и 14. Конъюнкции t_1 и t_3 являются существенными. Далее из конъюнкций t_2, t_4 , имеющих символ 1 в 10-м столбце, следует выбрать ту, которая имеет меньшую сложность.

Упражнение 9.5. Пусть $f^{(0)} < f^{(1)}$ и t — простая импликанта функции $f^{(1)}$.

(1) Если существуют $t_i, t_j \in PI(f^{(1)})$ такие, что $t = (t_i \nsubseteq t_j)$, то t не является существенной конъюнкцией.

(2) Если $t \leq f^{(0)}$ и t не является существенной конъюнкцией, то t можно получить за конечное число шагов, строя согласователи отличных от t простых импликант функции $f^{(1)}$.

(3) Если $f^{(1)}$ является линейной функцией от n переменных, то дизъюнктивная нормальная форма F , имеющая минимальную сложность и удовлетворяющая условию $f^{(0)} \leq F \leq f^{(1)}$, является совершенной дизъюнктивной нормальной формой функции $f^{(0)}$.

(4) Если $f^{(1)}$ — монотонная функция и $t \leq f^{(0)}$, то t — существенная конъюнкция.

Решение. Утверждение (1) следует непосредственно из (9.11).

(2) Положим $t = t_1, PI(f^{(1)}) = \{t_1, t_2, \dots, t_l\}$. Так как t_1 не является существенной конъюнкцией то $t_1 \leq f^{(0)} \leq t_2 \vee t_3 \vee \dots \vee t_l$. Отсюда и из соотношения (3.4) табл. 7.2 получаем $t_2 \vee t_3 \vee \dots \vee t_l = t_1 \vee t_2 \vee \dots \vee t_l = f^{(1)}$. Если далее к $S_0 = \{t_2, t_3, \dots, t_l\}$ применить метод нахождения всех простых импликант, описанный в § 9.2.1, то с помощью процедуры Q находится t_1 .

(3) Из задачи 9.9 следует, что $PI(f^{(1)})$ совпадает с множеством минимальных конъюнкций функции $f^{(1)}$. Согласно утверждению задачи 9.2 у минимальных конъюнкций функций $f^{(1)}$ согласователей нет. Отсюда и из утверждения (2) настоящей задачи вытекает, что если $t < f^{(0)}$, то t является существенной, а следовательно, и минимальной конъюнкцией.

(4) Так как простые импликанты функции $f^{(1)}$ не содержат отрицаний переменных (упражнение 9.2), то они не могут иметь согласователей. Отсюда и из утверждения (2) следует, что t является существенной конъюнкцией.

Из утверждения (4) упражнения 9.5 следует, что одной из дизъюнктивных нормальных форм минимальной сложности для монотонной функции является логическая сумма всех ее простых импликант. Если коэффициент сложности c_I положителен, то других форм минимальной сложности не существует.

Задача 9.17. Как следует изменить утверждения упражнения 9.2 и утверждение (4) упражнения 9.5, если f является смешанной монотонной функцией?

Указание. Если взять отрицания части или всех переменных, то f будет монотонной. При замене переменных их отрицаниями рассматриваемая здесь сложность не изменится.

Процедура упрощения матриц инцидентности (таблиц простых импликант). R1) Если имеется существенная строка i , то i включается в I (S_i включается в покрытие), все столбцы j , элементы a_{ij} которых равны 1, и строка i из матрицы исключаются (эти столбцы j при рассмотрении условия (9.14) принимать во внимание уже нет необходимости).

R2) Если строки i, i' таковы, что $a_{ij}=1$ для всех j , для которых $a_{i'j}=1$ и $c_i \leq c_{i'}$, то говорят, что строка i покрывает строку i' . Поскольку вместо $S_{i'}$ всегда можно взять S_i , строка i' исключается.

R3) Если столбцы j, j' таковы, что $a_{ij}=1$ для всех i , для которых $a_{ij'}=1$, то говорят, что столбец j' покрывает столбец j . Если столбец j' покрывает некоторый другой столбец j , то первый можно исключить (заметьте, что, в отличие от R2, здесь исключается покрываемый столбец). Эта операция возможна по следующей причине. Из условия (9.14) следует существование числа i такого, что $e_j \in S_i, i \in I$. Поскольку столбец j' покрывает столбец j , условие $e_{j'} \in S_i$ автоматически выполняется.

R4) После выполнения операций R1—R3 матрица инцидентности упрощается, и появляется возможность выполнить операции R1—R3 вторично.

Матрица инцидентности, к которой нельзя применить ни одну из операций R1—R3, называется *неприводимой* (или циклической).

Задача 9.18. При отсутствии входов, которыми можно пренебречь, никакая строка таблицы простых импликант не может покрывать другую строку, т. е. операция R2 не может быть применена первой. При наличии входов, которыми можно пренебречь, это утверждение, вообще говоря, неверно.

Указание. Первая часть утверждения следует из определения простой импликанты. Для доказательства второй части в качестве примера можно рассмотреть таблицу простых импликант из задачи 9.16.

Задача 9.19. Если в задаче минимизации для монотонной функции f имеются входы, которыми можно пренебречь, то вместо $S = D_1(f^{(0)})$ можно начать с $S = mD_1(f^{(0)})$.

Решение. Допустим, что A лежит в $D_1(f^{(0)})$, но не лежит в $mD_1(f^{(0)})$. Тогда, согласно упражнению 6.4, существует элемент $B \in mD_1(f^{(0)})$ такой, что $B < A$. Простая импликанта функции $f^{(1)}$, как следует из упражнения 9.2, является монотонной функцией, и, следовательно, $a \in D_1(t)$ для простой импликанты t , если $B \in D_1(t)$. Другими словами, поскольку столбец, соответствующий A , покрывает столбец, соответствующий B , то первый столбец можно исключить (операция R3).

Задача 9.20. Проверьте, что табл. 9.8 простых импликант является неприводимой.

Задача 9.21. Применив операции R1—R3 к таблице простых импликант из задачи 9.16, найдите решение задачи 9.16.

Пример 9.10. Пусть $D_1(f^{(1)}) = \{0, 1, 5, 8, 10, 13, 14, 15\}$ (функция f та же, что и в примере 9.9) и $D_1(f^{(0)}) = \{0, 1, 5, 8, 10, 13, 14\}$. Таблица простых импликант получается в результате исключения из табл. 9.8 последнего столбца и представляет собой табл. 9.9, а. Перенумеруем строки и столбцы этой таблицы целыми числами 1, 2, ... соответственно сверху вниз и слева направо. Так как строки 7 и 8 покрываются соответственно строками 4 и 6, то можно применить операцию R2 и исключить их из таблицы. После этого строки 4 и 6 становятся существенными строками и можно при-

менить операцию $R1$. При этом исключаются строки 4, 6 и столбцы 3, 5, 6, 7, а строки 4, 6 помечаются знаком $*$.

В результате выполнения этих операций получается упрощенная табл. 9.9,б. Применяв вторично операцию $R2$, можно исключить строки 2, 5. После этого строки 1, 3 становятся существенными. Таким образом, дизъюнктивная нормальная форма минимальной сложности имеет вид $t_1 \vee t_3 \vee t_4 \vee t_6$.

Применение задачи о минимальном покрытии с неприводимой матрицей инцидентности. Числа применений операций $R1$, $R2$, $R3$ к матрице инцидентности из m строк и r столбцов даже при простейшем методе пропорциональны соответственно mr , m^2r , mr^2 . Так как в результате выполнения одной операции исключается, по крайней мере, одна строка или один столбец, то число операций, необходимое для получения неприводимой матрицы инцидентности, не превосходит $m^3r + mr^3$. Для нахождения решения задачи о минимальном покрытии с неприводимой матрицей инцидентности размера $m \times r$ известными в настоящее время методами, вообще говоря, требуется число операций, пропорциональное $r \cdot 2^{cm}$, где c — константа. Следовательно, более предпочтительными представляются те методы упрощения, которые, в первую очередь, минимизируют параметр m . Основные методы решения задачи о минимальном покрытии с неприводимой матрицей инцидентности — описанный метод целочисленного программирования, а также метод ветвей и границ.

Метод ветвей и границ. Обозначим через $c(M)$ сложность минимального покрытия для матрицы инцидентности M . Пусть $M(J)$, $J = \{r_1, r_2, \dots, r_h\}$ — подматрица матрицы M , которая получается при исключении из M строк r_1, r_2, \dots, r_h и всех столбцов, которые имеют хотя бы одну единицу в указанных строках. Положим

$$k(J) = \sum_{i \in J} c_i.$$

Если J совпадает с множеством M , то будем считать, что $M(J) = \emptyset$ и $c(\emptyset) = 0$.

(BR) В матрице $M(J)$ выбирается некоторый столбец (например, первый или имеющий наибольшее число единиц). Пусть i_1, i_2, \dots, i_p — строки, имеющие символ 1 в столбце l . Положим

$$c(M(J)) = \min_{1 \leq j \leq p} \{c(M(J \cup \{i_j\})) + c_{i_j}\}.$$

Далее последовательно перебираются все возможные наборы J . При этом на каждом шаге с помощью последнего равенства оп-

Таблица 9.9. Таблица простых импликант из примера 9.10 и ее упрощение

1	2	3	4	5	6	7
1	1	1				
2		1	1			
3	1			1		
4*			1		1	
5				1	1	
6*					1	1
7						1
8						1

a)

1	2	4
1*	1	1
2		1
3*	1	1
5		1

б)

ределяются наборы J , которые дают покрытия, имеющие минимальную сложность среди всех найденных ранее покрытий. Эти покрытия и соответствующие им сложности запоминаются (далее они обозначаются соответственно через I и C); ненужные наборы J исключаются из рассмотрения. Этот метод дает оптимальное решение и называется *методом ветвей и границ*.

Начальные построения. Пусть M_0 — заданная неприводимая матрица инцидентности из n строк*. Положим

$M \leftarrow M_0$, $I \leftarrow \{1, 2, \dots, n\}$, $C \leftarrow k(I)$, $J \leftarrow \emptyset$.

и рассмотрим следующую процедуру:

$P(M, J, I, C)$: (1) для M с помощью операции (BR) выбираются столбец l и строки i_1, i_2, \dots, i_r ; (2) последовательно для j от 1 до r выполняется следующая процедура $Q(j)$:

$Q(j)$: упрощается матрица $M(\{i_j\})$. Пусть J' — множество существенных строк, полученных при выполнении рассматриваемой процедуры. Предположим, что матрица $M(\{i_j\} \cup J')$ является либо пустой (т. е. в ней нет строк или столбцов; такая матрица обозначается знаком \emptyset), либо неприводимой.

(1) Случай $k(J \cup \{i_j\} \cup J') < C$.

(1.1) Если $M(\{i_j\} \cup J') = \emptyset$, то $M_0(J \cup \{i_j\} \cup J') = \emptyset$ и, следовательно, $J \cup \{i_j\} \cup J'$ является покрытием. Так как сложность этого покрытия меньше сложности любого покрытия, полученного ранее, то в качестве I и C берутся соответственно $I \leftarrow J \cup \{i_j\} \cup J'$, $C \leftarrow k(J \cup \{i_j\} \cup J')$. После этого выполняется шаг (2) операции $P(M, J, I, C)$.

(1.2) Если $M(\{i_j\} \cup J') \neq \emptyset$ то необходимо определить $c(M(\{i_j\} \cup J'))$; для этого выполняются операции $P(M(\{i_j\} \cup J'), J \cup \{i_j\} \cup J', I, C)$.

(2) Случай $k(J \cup \{i_j\} \cup J') \geq C$. При этом $k(J \cup \{i_j\} \cup J') + c(M(\{i_j\} \cup J')) \geq C$, так что никакое покрытие, содержащее в качестве своего подмножества $J \cup \{i_j\} \cup J'$, не может иметь сложность, меньшую чем C . Следовательно, эти покрытия далее рассматривать не нужно, и процедура возвращается к шагу (2) операции $P(M, J, I, C)$.

Задача 9.22. Методом ветвей и границ найдите все дизъюнктивные нормальные формы минимальной сложности функции из примера 9.9.

Неприводимую матрицу инцидентности M размера $m \times r$, каждая строка которой содержит ровно две единицы (как, например, матрица в табл. 9.8), можно рассматривать как матрицу инцидентности ребер неориентированного графа G , имеющего m вершин и r ребер. При этом рассматривавшаяся ранее задача сводится к задаче о покрытии графа G минимальным числом ребер, которая формулируется следующим образом. Требуется выбрать минимальное число ребер графа G так, чтобы произвольная вершина G была концом, по крайней мере, одного из выбранных ребер. Для

* Запись $A \leftarrow B$ означает, что в качестве A берется B .

решения этой задачи известны методы, требующие порядка $m^{2,5}$ операций [32].

Пример 9.11. Если рассматривать задачу 9.22 как задачу теории графов, то граф G будет иметь вид, изображенный на рис. 9.5. Поскольку каждый столбец табл. 9.8 содержит ровно 2 единицы, то G представляет собой один замкнутый цикл и, очевидно, минимальное покрытие графа G получается, если выбрать ребра через одно. При этом искомыми решениями будут $t_1 \vee t_4 \vee t_8 \vee t_5$ и $t_2 \vee t_7 \vee t_6 \vee t_3$.

Задача 9.23. Найдите дизъюнктивную и конъюнктивную нормальные формы минимальной сложности для функции f с $D_1(f) = \{0, 2, 3, 6, 7, 8, 9, 12, 13, 14, 15\}$.

Таблицы простых импликант в случае с несколькими выходами. Рассмотрим на следующем примере применение метода Мак-Класки, когда сложность зависит от общего числа входов элементов.

Пример 9.12. Пусть $D_1(f^{(1)}) = \{0, 1, 3, 7\}$, $D_1(f^{(1)2}) = \{1, 3, 4, 5, 7\}$ (см. задачу 9.10), $D_1(f^{(0)1}) = \{0, 1, 3, 7\}$, $D_1(f^{(0)2}) = \{3, 4, 5, 7\}$. Таблица 9.10,а является таблицей простых импликант в случае с несколькими выходами и может быть построена с помощью следующих двух операций (1) и (2).

(1) Элементам $D_1(f^{(0)1})$ и $D_1(f^{(0)2})$ ставятся в соответствие столбцы таблицы. При этом общим элементам сопоставляется столько столбцов, какова его кратность (в данном случае, два).

Таблица 9.10. Таблица простых импликант из примера 9.12

Таблица а		$D_1(f_1^{(0)})$				$D_1(f_2^{(0)})$			
		0	1	3	7	3	4	5	7
p -простые импликанты функции $f_1^{(1)}$	$t_1 = \bar{x}_1 \bar{x}_2$	1	1						
p -простые импликанты функции $f_2^{(1)}$	$t_2 = x_1 \bar{x}_2$ $t_3 = x_3$					1	1	1	1
p -простые импликанты функции $f_1^{(1)} f_2^{(1)}$	$t_4 = \bar{x}_1 x_3$ $t_5 = x_2 x_3$	1	1	1	1	1			1

Таблица б	$D_1(f_2^{(0)})$	ΔN_G		ΔN_f	
	7				
t_3	1	0		1	
t_5	1	0		1	

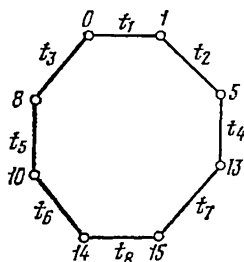


Рис. 9.5

(2) Строкам таблицы сопоставляются p -простые импликанты. При этом вначале строятся строки, соответствующие p -простым импликантам функции $f^{(1)}_1$, далее — функции $f^{(1)}_2$, и, наконец, функции $f^{(1)}_1 f^{(1)}_2$. Строятся матрицы инцидентности для строк, соответствующих p -простым импликантам функции $f^{(1)}_i$, и столбцов, соответствующих элементам $D_1(f^{(0)}_i)$. Остальные клетки таблицы остаются пустыми. Строится матрица инцидентности для строк, соответствующих p -простым импликантам функции $f^{(1)}_1 f^{(1)}_2$, и столбцов, соответствующих элементам $D_1(f^{(0)}_1)$, $D_1(f^{(0)}_2)$. Далее рассмотрим метод упрощения.

(3) Так как строка t_5 является существенной относительно столбца 7 группы $D_1(f^{(0)}_1)$, то она обязательно должна быть взята в качестве конъюнкции f_1 . Однако она не является существенной строкой для $D_1(f^{(0)}_2)$. Пока неясно, нужно ли брать строку t_5 в качестве конъюнкции f_2 . Правило упрощения $R1$ применимо только к столбцам группы $D_1(f^{(0)}_1)$. Поскольку строки t_1 и t_2 являются существенными относительно соответственно столбца 0 группы $D_1(f^{(0)}_1)$ и столбца 4 группы $D_1(f^{(0)}_2)$, то применяется правило $R1$.

(4) Правило минимизации $R2$ применяется только к строкам, соответствующим p -простым импликантам функций f_1 , f_2 , $f_1 f_2$. Здесь оно неприменимо.

(5) Правило минимизации $R3$ применяется только к столбцам, соответствующим элементам каждого множества $D_1(f^{(0)}_i)$; в других случаях неприменимо. Можно исключить столбец 3 группы $D_1(f^{(0)}_1)$ и столбцы 3, 5 группы $D_1(f^{(0)}_2)$.

(6) В результате выполнения операций $R1$, $R3$ получается табл. 9.10,б. В этой таблице в столбцах ΔN_G , ΔN_I указаны приращения соответственно числа вентиля и числа входов при выборе той или иной строки. Поскольку $\deg(t_3)=1$, то $\Delta N_G=0$ для t_3 . Поскольку строке t_5 уже соответствует элемент И для функции f_1 , то для нее также $\Delta N_G=0$. Таким образом, суммарное приращение числа входов равно 1. Так как в обоих случаях сложность оказывается одной и той же, задача имеет два решения:

$$f_1 = t_1 \vee t_5, \quad f_2 = t_2 \vee t_3 \quad \text{или} \quad t_2 \vee t_5.$$

Если имеется несколько выходов и $c_1=0$, а также в задаче о возможности реализации на ПЛМ из упражнения 9.4, правила упрощения $R1-R3$ можно применять к таблицам простых импликант в случае нескольких выходов как обычно, без какой-либо модификации. В случае (1) можно также обычным образом применять метод ветвей и границ. Если во втором случае ($c_1=c_2=\dots=c_n=1$) в процессе упрощения уже выбрано k' существенных строк, то достаточно рассмотреть только случай, когда число строк неприводимой матрицы инцидентности больше или равно $k-k'$. При этом в качестве начального значения переменной C берется $k-k'+1$. Если значение C вначале меньше $k-k'$, то шаг (1,1) операции $Q(j)$ заканчивает решение задачи и множество $J \cup \{i_j\} \cup J'$ дает ответ. В противном случае решения не существует.

9.4. О НЕ — И и НЕ—ИЛИ-схемах [17—19, 33]

Схемы из функциональных элементов, компонентами которых являются только элементы НЕ—И (НЕ—ИЛИ), называются НЕ—И (НЕ—ИЛИ)-схемами. Так как элемент НЕ—И можно рассматривать как последовательное соединение элемента И и инвертора, а элемент НЕ—ИЛИ — как последовательное соединение элемента ИЛИ и инвертора, то из упражнения 8.6 вытекает следующее утверждение.

У п р а ж н е н и е 9.6. Если НЕ—И (НЕ—ИЛИ)-схема реализует функции f_1, f_2, \dots, f_m , то двойственная ей схема, которая получается при замене каждого элемента исходной схемы элементом НЕ—ИЛИ с сохранением внешних входов и выходов и графа соединения, реализует функции $f_1^d, f_2^d, \dots, f_m^d$.

Далее в основном рассматриваются НЕ—И-схемы.

Преобразования НЕ—И-схем в схемы из элементов И и ИЛИ и обратно. Предположим, что заданная НЕ—И-схема имеет древовидную структуру (к выходу каждого элемента схемы присоединяется только один вход другого элемента). Применяя последовательно правило Де Моргана (сначала к выходному элементу), указанное на рис. 9.6 преобразование можно выполнить следующим образом.

T1) Если элемент G типа НЕ—И является выходным элементом либо если уже выполнено преобразование элемента, одним из входов которого является выход рассматриваемого эле-

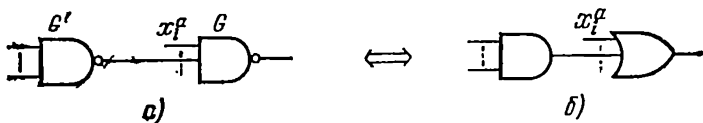


Рис. 9.6. Преобразование НЕ—И в ИЛИ и И

мента G , то G заменяется элементом И. Входами нового элемента И становятся входы первоначального элемента НЕ—И. Далее с каждым входом элемента G выполняется одно из следующих двух преобразований:

1) если рассматриваемый вход является внешним входом x_i^a , то он заменяется внешним входом x_i^a ;

2) если рассматриваемый вход — выход другого элемента G' , то выход последнего подключается к соответствующему входу элемента И без изменений.

T2) Преобразование **T1** выполняется до тех пор, пока в схеме остаются элементы НЕ—И.

Наоборот, предположим, что задана схема с древовидной структурой, состоящая только из элементов И, ИЛИ и удовлетворяющая следующим условиям: выходным элементом является элемент ИЛИ; выход любого элемента ИЛИ, не являющегося выходным, соединен со входом элемента И, а выход любого эле-

мента И соединен со входом некоторого элемента ИЛИ. Тогда допустимо следующее преобразование:

$T'1$) каждый элемент схемы заменяется элементом НЕ—И с тем же числом входов и выходов;

$T'2$) если входом элемента G типа ИЛИ является внешний вход x^a_i , то соответствующий вход элемента НЕ—И, который заменяет элемент G , заменяется внешним входом \bar{x}^a_i .

При описанных преобразованиях не изменяются функции, реализуемые схемами, конфигурация соединений, число вентилях и суммарное число входов. Следовательно, для определения НЕ—И-схемы с древовидной структурой, имеющей минимальную сложность, достаточно построить схему минимальной сложности из элементов И и ИЛИ, удовлетворяющую указанным условиям. Однако обычно использовать в качестве входов x_i и \bar{x}_i не представляется возможным, так что операция $T'2$ не всегда выполнима. Например, если в качестве внешних входов разрешается использовать только переменные x_1, x_2, \dots, x_n , а их отрицания использовать нельзя, то при необходимости во входе \bar{x}_i можно использовать в качестве инвертора элемент НЕ—И и реализовать таким образом \bar{x}_i . Однако этот прием приводит к увеличению числа элементов в схеме и, вообще говоря, не даст самую экономную схему (см. упражнение 9.8).

Пример 9.13. Пример преобразования схем приведен на рис. 9.7.

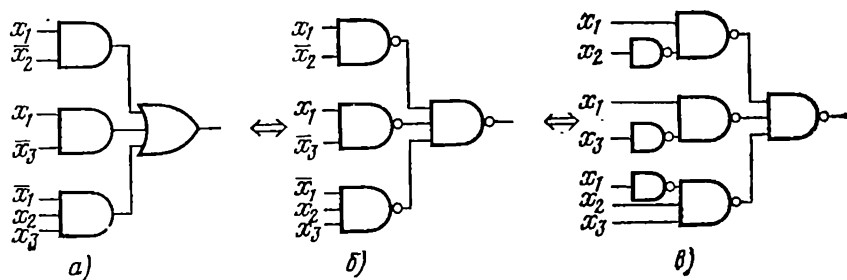


Рис. 9.7. Примеры преобразования схем

Задача 9.24. Рассмотрите описанное преобразование применительно к НЕ—ИЛИ-схеме.

Задача 9.25. Произвольная логическая функция может быть реализована НЕ—И (НЕ—ИЛИ)-схемой, глубина которой не превосходит 3.

Указание. К И—ИЛИ-схеме глубины 2 следует применить преобразования $T'1, T'2$; если при этом возникнет необходимость во входе \bar{x}_i , то в качестве инвертора следует использовать элемент НЕ—И (НЕ—ИЛИ).

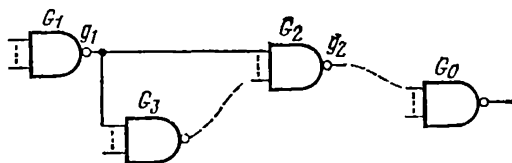
Задача 9.26. Преобразования $T1, T2, T'1, T'2$ применялись к схемам с древовидной структурой. Покажите, что если граф соединений $s(N)$ схемы N является двудольным (см. § 4.6) и все

вершины, соответствующие внешним выходам, можно окрасить в один цвет, то указанные преобразования здесь также применимы.

Указание. Решение следует из предположений и определения двудольного графа.

Упражнение 9.7. Если в НЕ—И-схеме N , показанной на рис. 9.8: 1) выход элемента G_1 (или внешний вход g_1) подсоединен ко входам элементов G_2, G_3, \dots , 2) путь элемента G_3 к выходному элементу G_0 проходит через все элементы G_2 , — то при исключении соединения от G_1 к G_3 выходная функция элемента G_0 не изменяется.

Рис. 9.8. Иллюстрации к упражнению 9.7



Решение. Пусть g_1, g_2, g_3 — выходные значения соответственно элементов G_1 (или внешнего входа g_1), G_2 и G_3 . Пусть g'_2, g'_3 — выходные значения элементов G_2, G_3 в схеме N' , получающейся из схемы N устранением соединения от G_1 к G_3 .

1) Если $g_1=0$, то $g_2=g'_2=1$.

2) Если $g_1=1$, то $g_3=g'_3$ и, следовательно, $g_2=g'_2$, т. е. всегда $g_2=g'_2$.

Значения g_3 и g'_3 , вообще говоря, могут различаться, но, как следует из условия (2), до тех пор, пока $g_2=g'_2$, это не влияет на выход элемента G_0 .

В дальнейшем предполагается, что отрицания переменных не могут быть внешними входами.

НЕ—И-схемы глубины 3. Для простоты рассмотрим только схемы с одним выходом. Как следует из упражнения 9.7, без ограничения общности достаточно рассмотреть схемы, подобные той, что указана на рис. 9.9. Входами элементов глубины 1 являются только переменные $x_1, x_2, \dots, x_l (l \leq n)$. Через I на этом рисунке обозначена часть схемы, представляющая собой сово-

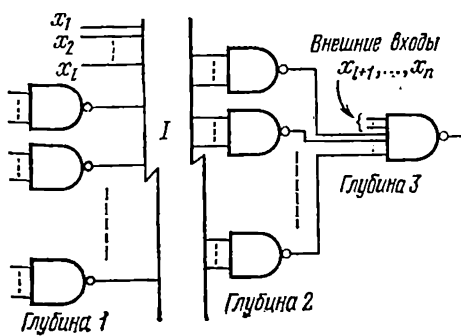


Рис. 9.9. НЕ—И-схема глубины 3

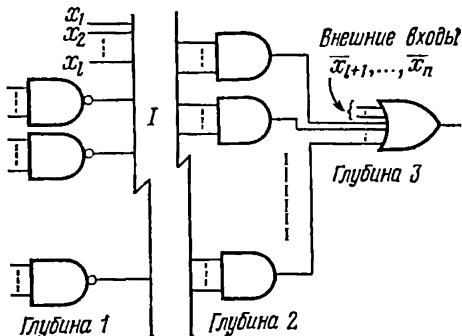


Рис. 9.10. Схема, эквивалентная схеме, изображенной на рис. 9.9

купность соединений выходов элементов глубины 1 со входами элементов глубины 2. Поскольку часть схемы, состоящая из элементов глубины 2 и 3, имеет древовидную структуру, то и с помощью преобразования $T1$ рассматриваемую схему можно преобразовать к виду, показанному на рис. 9.10. Эта схема отличается от И--ИЛИ-схем глубины 2 следующим.

1) Входами элементов И глубины 2 могут быть как переменные x_i , так и выражения типа $x_{i_1} x_{i_2} \dots x_{i_r}$.

2) Должен рассматриваться вопрос о минимизации сложности как совокупности элементов глубины 1, так и всей схемы в целом.

Упражнение 9.8. Рассмотрим функцию f , которая определяется таблицей Карно (табл. 9.11). Поскольку простыми импликантами здесь являются конъюнкции x_1 , \bar{x}_2 , $x_1 \bar{x}_3$, $\bar{x}_1 x_2 x_3$ и все они существенны, то дизъюнктивная нормальная функция минимальной сложности имеет вид

$$f = x_1 \bar{x}_2 \vee x_1 \bar{x}_3 \vee \bar{x}_1 x_2 x_3. \quad (9.20)$$

На рис. 9.7 показана НЕ--И-схема глубины 3, полученная из И--ИЛИ-схемы глубины 2 с применением преобразований $T'1$, $T'2$, при этом \bar{x}_i реализуется с помощью элемента НЕ--И, который используется в качестве ин-

вертора. Эта схема содержит семь элементов, а общее число входов равно 13.

В то же время, используя (9.20), ее можно преобразовать следующим образом: $f = x_1 (x_2 x_3) \vee \bar{x}_1 x_2 x_3$.

Поскольку $x_1 (x_2 x_3) = x_1 (x_1 x_2 x_3)$ и $\bar{x}_1 x_2 x_3 = (\bar{x}_1 x_2 x_3) x_2 x_3$, то $f = (x_1 x_2 x_3) (x_1 \vee x_2 x_3)$. В такой записи функцию f можно реализовать схемой, изображенной на рис. 9.11,а. Если далее к элементам глубины 2 и 3 этой схемы применить преобразования $T'1$ и $T'2$, то получится НЕ--И-схема, приведенная на рис. 9.11,б. Эта схема содержит четыре элемента, и суммарное число входов у нее равно 10. По сравнению со схемой на рис. 9.7,в, ее сложность уменьшилась на три элемента и на три входа.

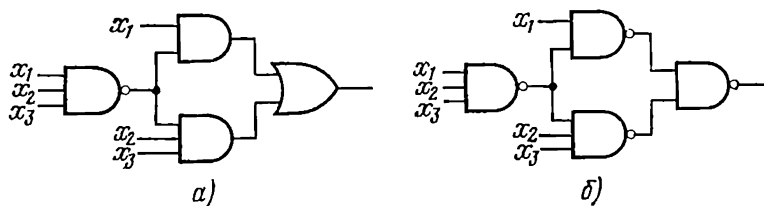


Рис. 9.11. Схемы из упражнения 9.8

Упражнение 9.9. Если в (9.20) положить $x_3=1$, то получим $x_1\bar{x}_2 \vee \bar{x}_1x_2 = x_1 \oplus x_2$.

На рис. 9.12 приведена схема, реализующая $x_1 \oplus x_2$ и получающаяся из схемы, изображенной на рис. 9.11,б. Эта схема содержит четыре элемента, общее число входов в ней равно 8.

Как видно из рис. 9.10 и упражнения 9.7, реализация функции f с помощью НЕ—И-схемы глубины 3 связана с представлением f в виде логической суммы выражений следующего типа:

$$x_{i_1} x_{i_2} \dots x_{i_\alpha} \overline{(x_{j_1} x_{j_2} \dots x_{j_\beta})} \overline{(x_{k_1} x_{k_2} \dots x_{k_\gamma})} \dots \overline{(x_{h_1} x_{h_2} \dots x_{h_\delta})}. \quad (9.21)$$

Если при этом среди $x_{j_1}, x_{j_2}, \dots, x_{j_\beta}$ и $x_{i_1}, x_{i_2}, \dots, x_{i_\alpha}$ имеется одна и та же буква, например x_{j_1} , то $\overline{(x_{j_1} x_{j_2} \dots x_{j_\beta})}$ можно заменить на $\overline{(x_{j_2} x_{j_3} \dots x_{j_\beta})}$ и приведенное выражение не изменится. Если, наоборот, x_{i_1} отличается от $x_{j_1}, x_{j_2}, \dots, x_{j_\beta}$ и $x_{j_1} x_{j_2} \dots x_{j_\beta}$ заменить на $x_{i_1} x_{j_1} x_{j_2} \dots x_{j_\beta}$, то функция (9.21) не изменится. Как показывают приведенные упражнения, такие преобразования являются одним из способов минимизации, позволяющим строить общие сомножители (...). Метод выбора выражений (9.21) с необходимой структурой (соответствующих простым импlicants при построении И—ИЛИ-схем глубины 2) и определения их логических сумм, имеющих минимальную сложность, был предложен Гимпелем [33].

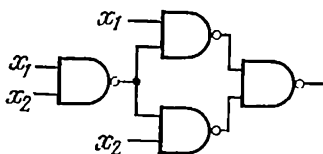


Рис. 9.12. НЕ—И-схема, реализующая $x_1 + x_2$

Упражнение 9.10. Рассмотрим представление $f = x_1 \oplus x_2 \oplus x_3$. Так как $f = (x_1 \oplus x_2) \oplus x_3$, то эту функцию можно реализовать с помощью схемы из упражнения 9.9 так, как показано на рис. 9.13. Эта схема содержит восемь элементов, имеет глубину 6, и общее число входов в ней равно 16. Большая глубина схемы приводит к большей задержке, что обычно нежелательно. На рис. 9.14 приведена схема, реализующая ту же функцию, но содержащая семь элементов, имеющая глубину 4 и 20 входов. С

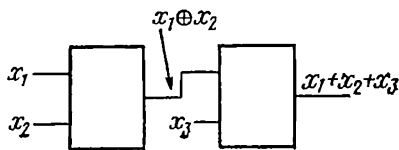


Рис. 9.13. НЕ—И-схема из упражнения 9.10

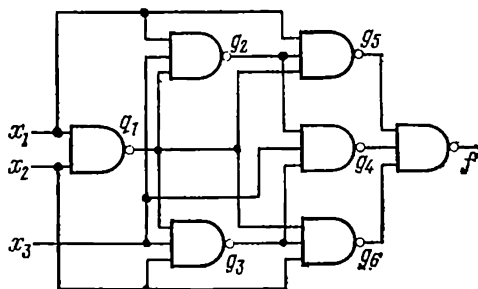


Рис. 9.14. НЕ—И-схема с минимальным числом элементов, реализующая функцию p_3

помощью ЭВМ было показано [33], что эта схема имеет наименьшее число элементов среди всех схем, реализующих функцию f и построенных из элементов, у которых число входов и коэффициентов ветвления выхода не превосходит 3.

Задача 9.27. Найдите дизъюнктивную нормальную форму минимальной сложности относительно входных переменных x_1, x_2, x_3 для функций g_1, g_2, \dots, g_6 , реализуемых в схеме, приведенной на рис. 9.14. Проверьте также, что $f = x_1 \oplus x_2 \oplus x_3$.

Задача 9.28. Добавив к схеме, приведенной на рис. 9.14, еще один элемент НЕ—И, и используя соответствующим образом выходы элементов g_1, g_2, \dots, g_6 , постройте схему, которая реализует мажоритарную функцию от трех переменных. Это будет означать, в частности, что с помощью восьми элементов можно реализовать двоичный сумматор. Проведите сравнение со схемой из задачи 8.4.

Указание. Заметьте, что $\overline{g_1 g_2 g_3} = \text{Maj}_3(x_1, x_2, x_3)$.

Задача 9.29. Постройте двоичный сумматор из элементов НЕ—ИЛИ.

Указание. Следует воспользоваться тем, что $x_1 \oplus x_2 \oplus x_3$ и Maj_3 являются двойственными функциями.

Задача 9.30. Предложите способ построения НЕ—И-схемы глубины 3 или менее, реализующий монотонную функцию f и имеющий минимальное число элементов и минимальное число входов.

Решение. Если предположить, что функция f реализуется НЕ—И-схемой глубины 3 или менее, то f может быть представлена в виде следующей логической суммы не равных тождественно нулю выражений (3.21): $f = g_1 \vee g_2 \vee \dots \vee g_r$.

Если выражение g_i содержит отрицание произведения переменных $x_{j_1} x_{j_2} \dots x_{j_p}$ то все их можно исключить и рассматривать произведения самих переменных $x_{i_1} x_{i_2} \dots x_{i_\alpha}$. В силу монотонности $g'_i \leq f$. Поскольку $g_i \leq g'_i$, то $f \leq g'_1 \vee g'_2 \vee \dots \vee g'_r \leq f$, т. е. $f = g'_1 \vee g'_2 \vee \dots \vee g'_r$.

В результате выполнения этой операции исключаются все элементы глубины 1, и сложность схемы уменьшается. В то же время функция f оказывается представленной в виде суммы элементарных конъюнкций, ни одна из которых не содержит отрицаний переменных. При этом число элементов N_G —число конъюнкций+1, а число входов $N_I = N_G + \text{сумма рангов конъюнкций}$. Из упражнений 9.2 и 9.5 следует, что логическая сумма всех простых импликант функции дает решение с минимальными числами N_G и N_I (в случае И—ИЛИ-схем глубины 2 различаются только сложности конъюнкций ранга 1). Если каждой простой импликанте l_i сопоставить взаимно однозначно элемент НЕ—И, входы которого соединить с внешними входами, соответствующими переменным, входящим в l_i , выходы всех элементов подключить ко входам еще одного элемента НЕ—И и последний рассматривать как выходной элемент, то получится НЕ—И-схема (глубины 2), имеющая минимальные параметры N_G, N_I в классе всех схем глубины 3 и менее, реализующих функцию f .

Задача 9.31. Рассмотрим НЕ—И-схему N , реализующую функцию f (как и ранее предполагается, что в качестве внешних входов нельзя использовать отрицания переменных).

1) Если f зависит от x_i и x_i является положительной (отрицательной) переменной, то в графе соединений $c(N)$ обязательно

существует путь от внешнего входа x_i к выходу схемы, проходящий через четное (нечетное) число элементов.

2) Если f зависит от x_i , но x_i не является ни положительной, ни отрицательной переменной, то в $c(N)$ существуют, по крайней мере, один путь от внешнего входа x_i к выходу схемы, проходящий через четное число элементов, и один путь, проходящий через нечетное число элементов.

Решение. (1) Пусть x_i — положительная переменная. Тогда существует набор $a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_n$ такой, что $f(a_1, a_2, \dots, a_{i-1}, 1, a_{i+1}, \dots, a_n) = 1$, $f(a_1, a_2, \dots, a_{i-1}, 0, a_{i+1}, \dots, a_n) = 0$. Пусть N' — схема, которая получается, если положить $x_j = a_j$ ($j \neq i$). Схема N' реализует x_i . Если в схеме N нет пути от входного узла x_i к выходному элементу, который проходит через нечетное число элементов, то N' не реализует функцию x_i . Утверждение (2) следует из утверждения (1).

Задачи

9.1(а). Число элементов и суммарное число входов в И—ИЛИ-схеме (ИЛИ—И-схеме) глубины 2, реализующей произвольную функцию $f \in \mathcal{B}_n$, не превосходят соответственно $2^{n-1} + 1$ и $(n+1)2^{n-1}$.

9.1(б). Если минимальные значения чисел элементов в И—ИЛИ-и ИЛИ—И-схемах глубины 2, реализующих функцию $f \in \mathcal{B}_n$, равны $2^{n-1} + 1$, то $f = c \oplus x_1 \oplus x_2 \oplus \dots \oplus x_n$, где $c = 0$ или 1 и $n \geq 2$. При этом суммарное число входов в схеме $(n+1)2^{n-1}$.

9.2. Пусть m — целое положительное число и $n = 2m + 1$. Если функция Maj_n реализуется И—ИЛИ-или ИЛИ—И-схемами глубины 2, то в обоих случаях минимальное число элементов в схеме $\binom{n}{m} + 1$; при этом суммарное число входов в схеме $(m+2)\binom{n}{m}$.

9.3(а). Предположим, что в дизъюнктивной нормальной форме $f = t_1 \vee t_2 \vee \dots \vee t_m$ элементарная конъюнкция t_i имеет вид $x_1^{a_1} x_2^{a_2} \dots x_r^{a_r}$ (при необходимости следует поменять нумерацию переменных). Если соответствующим выбором a_{r+1}, \dots, a_n можно добиться того, что при каждом $2 \leq j \leq m$ конъюнкция t_j включает x_i , a_{t_j} ($1 \leq i_j \leq n$), то $f \neq t_2 \vee t_3 \vee \dots \vee t_m$.

9.3(б). Насколько большой должна быть ширина k решетки элементов ИЛИ (см. пример 9.2) для того, чтобы кодер приоритета с $m = 3$ из примера 8.5 можно было реализовать в виде ПЛМ?

9.4. Выпишите дизъюнктивную нормальную форму минимальной сложности для двух функций, реализуемых схемой из задачи 8.4 в конце гл. 8.

В задачах 9.5—9.7 не разрешается использовать отрицания переменных в качестве внешних входов.

9.5. Пусть функция $f \in \mathcal{B}$ такова, что одна из ее переменных x_i не является ни положительной, ни отрицательной. Обозначим через L и N соответственно глубину и число элементов НЕ—И-схемы, реализующей функцию f : а) если $L \geq 3$, то $N \geq 4$; если $N = 4$,

то $L=3$; б) если $L=3$, $N=4$, то элементы в схеме соединены так, как показано на рис. 9.15, б или в. Если f не содержит положительных переменных, то в схеме на рис. 9.15, в внешний вход x_i соединяется с одним из элементов G_2 или G_3 и элементом G_4 . Если число существенных переменных n , а число отрицательных

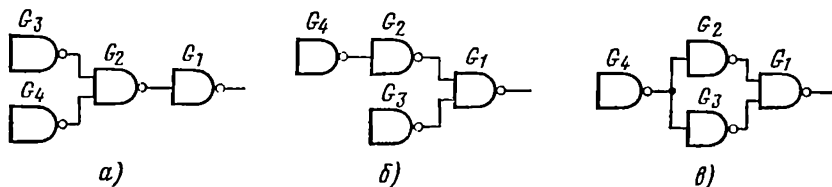


Рис. 9.15

переменных m , то суммарное число входов N_I в схеме не может быть меньше, чем $4 + 2n - m$.

9.6. НЕ—И-схемы, изображенные на рис. 9.11 и 9.12, имеют минимальные глубину, число элементов и суммарное число входов в классе всех схем, реализующих те же функции.

9.7. В классе НЕ—И-схем, реализующих функцию $x_1 x_2 \vee \bar{x}_1 \bar{x}_2 = 1 \oplus x_1 \oplus x_2$, приведенная на рис. 9.16 схема имеет минимальные глубину, число элементов и суммарное число входов. Если все элементы этой схемы заменить элементами НЕ—ИЛИ, то получится схема, имеющая минимальные глубину, число элементов и суммарное число входов в классе всех НЕ—ИЛИ-схем, реализующих $x_1 \oplus x_2$.

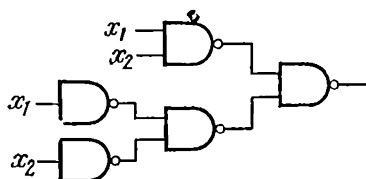


Рис. 9.16. Наилучшая НЕ—И-схема, реализующая $x_1 x_2 \vee \bar{x}_1 \bar{x}_2$

9.8. (Задача о 2-удовлетворимости). Если дана произвольная конъюнктивная нормальная форма $s_1 \cdot s_2 \cdot \dots \cdot s_m$, каждая элементарная конъюнкция s_i которой содержит не более двух букв, то за полиномиальное время можно определить, равна эта конъюнктивная нормальная форма тождественно нулю или нет.

Глава 10

Конечные автоматы

10.1. Определение конечного автомата [35—37]

Определение последовательности. Пусть A — непустое множество, состоящее из конечного числа элементов. Если написать в произвольном порядке некоторое конечное число элементов множества A , допуская возможность повторения элементов, но

помещая их друг за другом, то получится *последовательность*, состоящая из элементов A , или просто *последовательность элементов* A . Например, если $A = \{0, 1\}$, то последовательностями являются 0, 1, 00, 01, 10, 11, 000 и т. д. Слова английского текста являются последовательностями, символами которых являются 52 буквы a, A, b, B, c, C, \dots . Английский текст можно также рассматривать как последовательность, множество символов которой кроме 52 букв включает также такие знаки, как точка, запятая, пробел, вопросительный знак и др. Следует заметить, что в данном случае пробел также рассматривается как один из знаков. Последовательность, которая получается в результате n -кратной записи одной и той же буквы a , обозначается через an .

Число элементов, входящих в последовательность α (при подсчете каждый элемент учитывается столько раз, сколько он входит в последовательность), называется *длиной* последовательности и обозначается через $|\alpha|$. Для последовательности α однозначно определяется ее i -й символ ($1 \leq i \leq |\alpha|$). Наоборот, если n — целое положительное число и g — отображение множества $\{1, 2, \dots, n\}$ в A , то последовательность можно определить как пару (n, g) . При этом n является длиной последовательности, а $g(i)$ — ее i -м символом ($1 \leq i \leq n$).

Множество всех последовательностей из элементов A , имеющих конечную длину, будем обозначать через A^+ . Для удобства обращения с множествами последовательностей вводится последовательность длины 0 (*пустая последовательность*). В данной книге пустая последовательность обозначается через λ . Множество всех последовательностей конечной длины из элементов A , включающее пустую последовательность, обозначается через A^* . Так что $A^* = A^+ \cup \{\lambda\}$. Пусть $A = (a_1, a_2, \dots, a_m)$

$$\text{и } \alpha_1 = a_{i_1} a_{i_2} \dots a_{i_k} (1 \leq i_l \leq m, 1 \leq l \leq h = |\alpha_1|),$$

$$\alpha_2 = a_{j_1} a_{j_2} \dots a_{j_h} (1 \leq j_l \leq m, 1 \leq l \leq k = |\alpha_2|)$$

— две произвольные последовательности из элементов A .

1) Последовательность α_1 равна последовательности α_2 тогда и только тогда, когда $h=k$ и $i_l = j_l (1 \leq l \leq h)$.

2) Определим $\alpha_1 \cdot \alpha_2$ как последовательность $a_{i_1} a_{i_2} \dots a_{i_h} a_{j_1} a_{j_2} \dots a_{j_k}$. При этом $|\alpha_1 \cdot \alpha_2| = |\alpha_1| + |\alpha_2|$. Знак « \cdot » называется знаком *присоединения* и часто опускается. Очевидно, что для произвольных последовательностей $\alpha_1, \alpha_2, \alpha_3 \in A^+$ выполняется закон ассоциативности, и поэтому скобки можно опускать и пользоваться записью $\alpha_1 \alpha_2 \alpha_3$.

3) Будем считать, что $a^{(0)} = \lambda$ для произвольного элемента $a \in A$ и что для произвольной последовательности $\alpha \in A^*$, $\alpha = \lambda \alpha = \alpha \lambda$.

4) Последовательность $\alpha \alpha \cdot \dots \cdot \alpha$, получаемую n -кратным повторением последовательности α , будем обозначать через α^n .

Определение преобразователя. Пусть I, O — конечные множества соответственно входных и выходных символов и f — отображение I^+ на O . Тройка $M = (I, O, f)$ называется *преобразователем*, и ей обычно придается следующий смысл. Рассматриваются дискретные моменты времени $t=1, 2, \dots, i, \dots$. В каждый дискретный момент времени, начиная с $t=1$, на вход преобразователя подается входной символ. Предположим, что в момент $t=i$ таким символом является a_i . В момент времени $t=i$ (на практике с некоторой задержкой после поступления i -го входного символа) преобразователь M на основании поступившей к этому моменту времени входной последовательности $\alpha_i = a_1 a_2 \dots a_i$ формирует выходной символ $f(a_i)$. Это устройство — идеализированная модель таких реальных объектов, как автоматы для продажи различных товаров, рассматриваемые далее синхронные последовательные схемы и др.

Последовательность i выходных символов $f(a_1)f(a_1 a_2) \dots f(a_1 a_2, \dots, a_i)$, формируемую преобразователем M по входной последовательности $\alpha_i = a_1 a_2 \dots a_i$, будем обозначать через $\bar{f}(\alpha_i)$. Преобразователь M представляет собой устройство, которое последовательно преобразует входную последовательность α_i в выходную $\bar{f}(\alpha_i)$ (рис. 10.1). Это устройство обладает памятью в том смысле, что в каждый момент времени выходной символ обычно зависит от входных, поступающих во все предыдущие моменты времени. При этом по мере роста i неограниченно увеличивается и необходимый объем памяти. Попробуем ответить далее на следующий вопрос. Какими свойствами обладают преобразователи с «конечным» объемом памяти?

Предположим, что входные последовательности можно разбить определенным образом на классы s_1, s_2, \dots, s_n таким образом, что знания класса, которому принадлежит начало входной последовательности, вполне достаточно для правильного продолжения выходной последовательности последующим входным символам. Общее число возможных входных последовательностей до момента $t=i$ равно $|I|^i$. Даже если допустить на время, что $|I|=1$, все равно необходимо различать входные последовательности равной длины. Формально это утверждение можно сформулировать следующим образом.

Если $\alpha_1, \alpha_2 \in I^*$ и для произвольного $\beta \in I^*$

$$f(\alpha_1 \beta) = f(\alpha_2 \beta), \quad (10.1)$$

то α_1 и α_2 называются f -эквивалентными. Для обозначения этого отношения используется запись $\alpha_1 \overset{f}{\approx} \alpha_2$. Очевидно, отношение

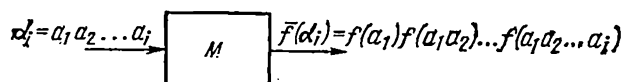


Рис. 10.1. Преобразователь

$\stackrel{f}{\approx}$ » обладает свойствами 1) рефлексивности: $a \stackrel{f}{\approx} a$; 2) симметричности: если $\alpha_1 \stackrel{f}{\approx} \alpha_2$, то $\alpha_2 \stackrel{f}{\approx} \alpha_1$; 3) транзитивности: если $\alpha_1 \stackrel{f}{\approx} \alpha_2$ и $\alpha_2 \stackrel{f}{\approx} \alpha_3$, то $\alpha_1 \stackrel{f}{\approx} \alpha_3$ и, следовательно, является отношением эквивалентности. Оно разбивает I^* на классы эквивалентности (см. § 1.4). Отношение \approx обладает, кроме того, следующим свойством.

Свойство правой инвариантности. Если $\alpha_1 \stackrel{f}{\approx} \alpha_2$, то для произвольной последовательности $\beta \in I^*$, $\alpha_1 \beta \stackrel{f}{\approx} \alpha_2 \beta$.

Доказательство. По предположению, для произвольной последовательности $\gamma \in I^*$, $f(\alpha_1 \beta \gamma) = f(\alpha_2 \beta \gamma)$, т. е. $\alpha_1 \beta \stackrel{f}{\approx} \alpha_2 \beta$.

Обозначим через S множество всех классов эквивалентности, а через $s(\alpha)$ — класс эквивалентности, содержащий последовательность α . Если $\alpha_1 \stackrel{f}{\approx} \alpha_2$, то для произвольного символа $a \in I$ имеем $f(\alpha_1 a) = f(\alpha_2 a)$, $s(\alpha_1 a) = s(\alpha_2 a)$, т. е. $f(\alpha a)$ и $s(\alpha a)$ однозначно определяются только классом эквивалентности $s(\alpha)$, которому принадлежит $\alpha \in I^*$ и символом a . Это означает, что существуют отображения $\delta: S \times I \rightarrow O$ и $\sigma: S \times I \rightarrow S$ такие, что для произвольных $\alpha \in I^*$ и $a \in I$

$$f(\alpha a) = \delta(s(\alpha), a), \quad s(\alpha a) = \sigma(s(\alpha), a).$$

Вообще говоря, если f — произвольная функция, то S может оказаться бесконечным множеством. Если S — конечное множество, то функция f называется *регулярной*. У регулярной функции области определения и области значений отображений δ и σ являются конечными. Если для входной последовательности α запомнить $s(\alpha)$ ($s(\alpha)$ называется *состоянием*), то в дальнейшем можно правильно определять выходные символы только по $s(\alpha)$ и вновь поступившим входным символам. Если же функция f нерегулярна, то при конечном объеме памяти реализовать разбиение входных последовательностей на классы эквивалентности невозможно (будут существовать входные последовательности, которые приводят к неверным выходным значениям).

Определение конечного автомата. Конечный автомат определяется как совокупность следующих пяти объектов:

- 1) S — конечное множество состояний;
- 2) I — конечное множество входных символов;
- 3) O — конечное множество выходных символов;
- 4) σ — отображение $S \times I$ в S ; эта функция определяет по текущим состоянию и входу следующее состояние и называется *функцией изменения состояний* или *функцией переходов*;
- 5) δ — отображение $S \times I$ в O ; эта функция определяет по текущим состоянию и входу текущее значение выхода и называется *выходной функцией*.

Если δ зависит как от S , так и от I , то конечный автомат называется автоматом Мили; если же δ зависит только от S , — то автоматом Мура. В автомате Мура, в отличие от автомата Мили, каждый выходной символ формируется просто по состоянию автомата, а входные символы при этом не учитываются.

Пример 10.1. Рассмотрим конечный автомат M , на вход которого поступают последовательно пары компонент двоичных представлений чисел x и y (начиная с младших разрядов), а на выходе последовательно (начиная с младших разрядов) формируется двоичное представление суммы $x+y^*$.

Пусть $x_1, x_2, \dots, x_i, \dots, y_1, y_2, \dots, y_i, \dots, s_1, s_2, \dots, s_i, \dots$ — двоичные представления чисел x, y и $x+y$ соответственно (начиная с младших разрядов). Пусть $c_0=0$ и c_i — символ переноса из $(i-1)$ -го разряда в i -й разряд. Тогда, согласно упражнению 7.5.

$$s_i = x_i \oplus y_i \oplus c_i, \quad c_{i+1} = \text{Maj}_3(x_i, y_i, c_i).$$

Таким образом, если в качестве состояния, которое запоминается после введения в конечный автомат $i-1$ разрядов двоичных представлений, взять величину c_i , то можно правильно вычислить как выходной символ s_i , так и следующее состояние. Обозначим через t_1 и t_2 состояния конечного автомата, соответствующие $c_i=0$ и $c_i=1$. Пусть $I = \{(0, 0), (0, 1), (1, 0), (1, 1)\}^{**}$, $O = \{0, 1\}$. Функции σ, δ определяются табл. 10.1, которая называется *таблицей переходов и выходных значений*. Рассмотренный конечный автомат является автоматом Мили и называется *последовательным двоичным сумматором*.

Таблица 10.1. Таблица переходов и выходных значений (пример 10.1)

Состояние конечного автомата	(0, 0)	(0, 1)	(1, 0)	(1, 1)
t_1	$t_1, 0$	$t_1, 1$	$t_1, 1$	$t_2, 0$
t_2	$t_1, 1$	$t_2, 0$	$t_2, 0$	$t_2, 1$

ответствующего ей состояния. Вершина с именем s обычно называется просто вершиной s . Из каждой вершины s проводятся ориентированные ребра во все вершины $\sigma(s, a)$, $a \in I$. Ребру, проведенному из вершины s в вершину $\sigma(s, a)$, присваивается имя $a / [\delta(s, a)]$. Других ребер диаграмма состояний $G(M)$ не имеет.

Пример 10.2. На рис. 10.2 приведена диаграмма состояний конечного автомата, описанного в примере 10.1. Вместо того

* Для простоты здесь не рассматриваются вспомогательные устройства, например используемые для указания окончания ввода при поступлении старших разрядов.

** Пары $(0, 0), \dots$ рассматриваются здесь как самостоятельные символы.

чтобы проводить i ребер с совпадающими начальными и совпадающими конечными точками, но различающимися именами, на диаграмме состояний для простоты проводится только одно ребро, на котором указываются все i имен.

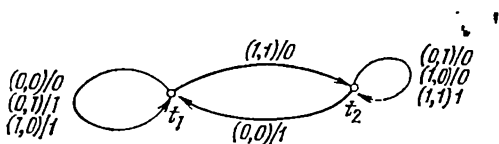


Рис. 10.2. Диаграмма состояний двоичного последовательного сумматора

Частично определенные конечные автоматы. В приведенном определении конечного автомата σ и δ считаются всюду определенными на $S \times I$ функциями. Если либо обе функции σ и δ , либо одна из них определены на $S \times I$ не всюду, т. е. являются *частичными функциями*, то конечный автомат называется *частично определенным*. Такие конечные автоматы являются моделями многих широко используемых на практике устройств. Если же обе функции, σ и δ , определены всюду, то конечный автомат называется *полностью определенным*. В дальнейшем термином «конечный автомат» будет обозначаться полностью определенный конечный автомат.

Упражнение 10.1. Вновь рассмотрим блок управления автоматом для продажи кока-колы из примера 8.4.

1) В качестве входных переменных возьмем T , F , H , указывающие соответственно, опущена или нет монета в 10 иен, 50 иен и 100 иен, а также N_{0i} и R_i , указывающие, нажаты или нет кнопки отказа от добавления воды и возврата денег соответственно.

2) Выходными сигналами в данном случае являются: N_f (не формируется ни одна из команд), $O(i, j)$ (выдачи на исполнительный механизм команды наполнения стакана кока-колой с добавлением воды, если $i=0$, без добавления воды, если $i=1$, подача стакана в окно и выдачи сдачи на сумму y иен), $R(y)$ (формирования команды выдачи y иен).

Предположим, что кока-кола, стаканы, сдача и вода имеются в достаточном количестве. Предположим также, что при отсутствии неисправностей автомат должен работать следующим образом (вопросы, связанные с отсутствием нужных компонент, их восполнением и контролем правильности оплаты, здесь не затрагиваются).

Пусть начальное состояние автомата s_0 . Если автомат находится в состоянии s_0 , то после поступления в него момент на сумму 50 иен и выше и непоступления к этому времени сигнала возврата денег R_i автомат формирует на выходе $O(i, x-50)$

($i=1$, если за это время поступил хотя бы один из сигналов N_{0i} , и $i=0$ в противном случае) и возвращается в состояние s_0 .

Если автомат находится в состоянии s_0 , в него опущены монеты на сумму $x \leq 50$ иен и поступил сигнал R_i , то на выход выдается команда $R(x)$ и автомат возвращается в состояние s_0 .

Для того чтобы конечный автомат удовлетворял указанным условиям, с момента выхода из состояния s_0 и до возвращения в состояние s_0 он должен помнить: а) значение $i=1$, если поступил хотя бы один из сигналов N_{0i} , и значение $i=0$ в противном случае; б) сумму $x < 50$ иен, опущенную в автомат, т. е. одну из величин 0, 10, 20, 30 или 40 иен. Следовательно, состояниями следует считать: $i=0, 1$; $x=0, 10, 20, 30, 40$.

По приведенным условиям строится таблица переходов и выходных значений (табл. 10.2). Для того чтобы описанное ранее управление имело смысл, необходимо предположить, что с момента поступления в автомат 50 или более иен и до момента,

Таблица 10.2. Таблица переходов и выходных значений из упражнения 10.1

Состояния	Входы					
	R_i	N_{0i}	H	F	T	
$S(0, x)$	$s_0, R(x)$	$S(1, x), N_f$	$s_0 O(0, 50+x)$	$s_0, O(0, x)$	$s_0, O(0, 0)$ при $x=40$	$S(0, x+10), N_f$ при $x < 40$
$S_1(1, x)$	$s_0, R(x)$	$R(1, x), N_f$	$s_0, O(1, 50+x)$	$s_0, O(1, x)$	$s_0, O(1, 0)$ при $x=40$	$S(1, x+10), N_f$ при $x < 40$

когда стакан получен покупателем, никакие команды на вход не поступают. Для того чтобы это условие выполнялось автоматически, после выдачи на выход команды $O(i, y)$ автомат должен переходить в специальное состояние s_w . В этом состоянии автомат находится до тех пор, пока с исполнительного механизма не поступит сигнал E , указывающий на то, что сдача и стакан с напитком выданы. При переходе в состояние s_w на выходе формируется команда, отличающая входы N_{0i} , R_i и исключающая возможность введения в автомат новых монет. После поступления сигнала E автомат возвращается в состояние s_0 .

Обобщение функций σ и δ . Функции σ и δ можно обобщить, определив их соответственно как отображения $S \times I^*$ в S и $S \times I^+$ в O , а также введя отображение $\bar{\sigma}: S \times I^*$ в O^* так, что для произвольного $s \in S$ выполняются условия:

$$(1) \sigma(s, \lambda) = s, \quad \bar{\delta}(s, \lambda) = \lambda; \quad (10.2)$$

$$(2) \text{ для произвольных } \alpha \in I^* \text{ и } a \in I$$

$$\sigma(s, \alpha a) = \sigma(\sigma(s, \alpha), a); \quad (10.3)$$

$$\delta(s, \alpha a) = \delta(\sigma(s, \alpha), a); \quad (10.4)$$

$$\bar{\delta}(s, \alpha a) = \bar{\delta}(s, \alpha) \bar{\delta}(\sigma(s, \alpha), a). \quad (10.5)$$

Другими словами, если вначале автомат M находился в состоянии s и в M введена последовательность входных символов $\alpha \in I^*$, то автомат в результате переходит в состояние $\sigma(s, \alpha)$, формирует выходной символ $\delta(s, \alpha)$ и полную выходную последовательность $\bar{\delta}(s, \alpha)$, соответствующую входной последовательности α .

Задача 10.1. Для произвольного $s \in S$ и произвольных $\alpha, \beta \in I^*$ справедливы следующие соотношения:

$$\sigma(s, \alpha\beta) = \sigma(\sigma(s, \alpha), \beta);$$

$$\delta(s, \alpha\beta) = \delta(\sigma(s, \alpha), \beta), \beta \in I^+;$$

$$\bar{\delta}(s, \alpha\beta) = \bar{\delta}(s, \alpha) \bar{\delta}(\sigma(s, \alpha), \beta).$$

Указание. Для решения задачи следует воспользоваться соотношениями (10.3)—(10.5) и индукцией по $|\beta|$.

Упражнение 10.2. Если на вход конечного автомата $M = (S, I, O, \sigma, \delta)$, находившегося вначале в некотором состоянии S_0 , подается один и тот же входной символ a , то самое большее через $|S| - 1$ шагов (тактов) выходная последовательность этого автомата будет периодической; ее период не превосходит $|S|$.

Решение. Рассмотрим $(\sigma s_0; a^i)$ ($i=0, 1, \dots$). Так как число состояний $|S|$ конечно, то существуют такие числа j и j' , что $\sigma(s_0; a^j) = \sigma(s_0; a^{j'})$ ($0 \leq j' \leq j$). Пусть j_0 — это минимальное из j и j' — соответствующее ему число j' . Тогда $j_0 < j \leq |S|$. Положим $j_0 - j'_0 = p$, $\sigma(s_0, a^{j'_0}) = s_1$. Пусть n — целое неотрицательное число. С помощью соотношений задачи 10.1 индукцией по n получаем

$$\sigma(s_0, a^{j'_0 + np}) = s_1;$$

$$\bar{\delta}(s_0, a^{j'_0 + np}) = \bar{\delta}(s_0, a^{j'_0}) (\bar{\delta}(s_1, a^p))^n.$$

Задача 10.2. Проанализируйте утверждение последнего упражнения на диаграмме состояний.

Задача 10.3. В упражнении 10.2 вместо a возьмите произвольную последовательность $\alpha \in I^+$ и рассмотрите выходную последовательность, соответствующую входной $\alpha, \alpha^2, \alpha^3, \dots$

Указание. Справедливы утверждения, аналогичные утверждениям из упражнения 10.2.

Последнее упражнение указывает одно из естественных ограничений, возникающих из-за конечного объема памяти автомата.

10.2. Сжатие конечных автоматов

Здесь рассматриваются две задачи. Первая связана с определением того, является ли один из двух заданных автоматов изоморфным другому или же он обладает большими возможностями. Во второй задаче считается, что задан конечный автомат и среди автоматов с теми же возможностями нужно определить тот, который является в некотором смысле простейшим.

Неразличимые состояния. Рассмотрим два конечных автомата: $M_1 = (S_1, I, O, \sigma_1, \delta_1)$ и $M_2 = (S_2, I, O, \sigma_2, \delta_2)$. (Это может быть один и тот же конечный автомат. Заметим, что у обоих автоматов входной I и выходной O алфавиты совпадают.) Если $s_1 \in S_1$, $s_2 \in S_2$ и для всех $\alpha \in I^*$ $\bar{\delta}_1(s_1, \alpha) = \bar{\delta}_2(s_2, \alpha)$ и, следовательно, $\delta_1(s_1, \alpha) = \delta_2(s_2, \alpha)$, то состояния s_1 и s_2 называются *неразличимыми*, или *эквивалентными*. Другими словами, если для произвольной входной последовательности из состояний s_1 и s_2 формируются одинаковые выходные последовательности, то с точки зрения преобразования входных последовательностей в выходные, состояния s_1 и s_2 совершенно неразличимы. Для обозначения неразличимости состояний s_1 и s_2 будем пользоваться записью $s_1 \approx s_2$.

Далее, если для всех последовательностей $\alpha \in I^*$ длины $k \geq 1$ $\bar{\delta}_1(s_1, \alpha) = \bar{\delta}_2(s_2, \alpha)$ и, следовательно, $\delta_1(s_1, \alpha) = \delta_2(s_2, \alpha)$, то состояния s_1 и s_2 называются *k -неразличимыми*, или *k -эквивалентными*. Это отношение обозначается $s_1 \approx_k s_2$. По определению, если $s_1 \approx_k s_2$, то $s_1 \approx_{k-1} s_2$. Обратное утверждение, вообще говоря, неверно. Состояния s_1 и s_2 , связанные отношением $s_1 \approx s_2$, называются *совместимыми по выходу*.

Задача 10.4. Если $s_1 \approx s_2$ для произвольной последовательности $\alpha \in I^*$, то $\sigma_1(s_1, \alpha) = \sigma_2(s_2, \alpha)$.

Решение. Из определений и задачи 10.1 следует, что для произвольной последовательности $\beta \in I^*$ имеем $\bar{\delta}_1(s_1, \alpha\beta) = \bar{\delta}_1(s_1, \alpha)\bar{\delta}_1(\sigma(s_1, \alpha), \beta) = \bar{\delta}_2(s_1, \alpha\beta) = \bar{\delta}_2(s_2, \alpha)\bar{\delta}_2(\sigma_2(s_2, \alpha), \beta)$, так что $\bar{\delta}_1(\sigma(s_1, \alpha), \beta) = \bar{\delta}_2(\sigma(s_2, \alpha), \beta)$. Отсюда и из определений получаем, что $\sigma_1(s_1, \alpha) = \sigma_2(s_2, \alpha)$.

Определение разбиения. Если подмножества $S_1, S_2, \dots, S_i, \dots$ множества S таковы, что $S = \bigcup_i S_i$, $S_i \neq \emptyset$, $S_i \cap S_j = \emptyset$, $i \neq j$, то $\{S_1, S_2, \dots, S_i, \dots\}$ называется разбиением множества S . Каждое из подмножества S_i называется при этом *блоком* разбиения. Пусть R — это отношение эквивалентности на множестве S . Разбиение множества S , каждый блок которого состоит из всех элементов, принадлежащих некоторому классу эквивалентности отношения R (см. § 1.4), называется разбиением S , соответствующим R .

Рассмотрим два разбиения: $P_1 = \{S_1^{(1)}, S_2^{(1)}, \dots, S_i^{(1)}, \dots\}$ и $P_2 = \{S_1^{(2)}, S_2^{(2)}, \dots, S_j^{(2)}, \dots\}$ множества S . Если для каждого $S_j^{(2)}$ ($j=1, 2, \dots$) существует i_j такое, что $S_j^{(2)} \subset S_{i_j}^{(1)}$, т. е. если каждый блок разбиения P_2 вложен в некоторый блок разбиения P_1 , то P_2 называется *измельчением* P_1 . Это обозначается так: $P_2 \leq P_1$. Если $P_1 \leq P_2$ и P_1 не совпадает с P_2 ($P_1 \neq P_2$), то говорят, что P_2 является *истинным измельчением* P_1 , и пишут $P_1 < P_2$. Пусть R — отношение эквивалентности на множестве S . Разбиение P множества S называют R -совместными, если в каж-

дом блоке этого разбиения произвольные два элемента, x и y , связаны отношением $R(xRy)$.

Задача 10.5. (1) Если P_R — это разбиение множества S , соответствующее отношению эквивалентности R , и P — R -совместимое разбиение S , то $P \leq P_R$.

(2) Пусть R_1, R_2 — отношения эквивалентности на S и $R_2 \subset R_1$ (если выполняется R_2 , то выполняется и R_1). Если P_1 и P_2 — разбиения S , соответствующие R_1 и R_2 , то $P_2 \leq P_1$.

(3) Если разбиения P_1 и P_2 множества S таковы, что $P_1 \leq P_2$ и $P_2 \leq P_1$, то $P_1 = P_2$.

Указание. Все утверждения задачи могут быть легко получены непосредственно из определений.

Далее на время предположим, что $M_1 = M_2 = M = (S, I, O, \sigma, \delta)$.

Задача 10.6. Отношения « \approx » и « \approx_k » удовлетворяют законам рефлексивности, симметричности и транзитивности, т. е. являются отношениями эквивалентности.

Решение. Следует из определений. Обозначим через $P^{(\omega)}$ и $P^{(k)}$ разбиения множества S , соответствующие отношениям эквивалентности « \approx » и « \approx_k ». Согласно утверждению (2) задачи 10.5, имеем

$$P^{(\omega)} \leq P^{(k+1)} \leq P^{(k)}. \quad (10.6)$$

Совместимые по выходу разбиения. Если разбиение P множества S является измельчением разбиения $P^{(1)}$, т. е. если для произвольных двух состояний s_1 и s_2 , принадлежащих одному блоку разбиения P , для всех $a \in I$ справедливо соотношение $\delta(s_1, a) = \delta(s_2, a)$, то P называется *совместимым по выходу*.

Возможность подстановки. Если для произвольных двух состояний s_1 и s_2 , принадлежащих одному блоку разбиения P множества S , и для произвольного входного символа $a \in I$ оба элемента $\sigma(s_1, a)$ и $\sigma(s_2, a)$, также принадлежат одному блоку рассматриваемого разбиения, то говорят, что P — допустимая подстановка.

Задача 10.7. $P^{(\omega)}$ — допустимая подстановка.

Решение. Следует из определений и утверждения задачи 10.4.

Лемма 10.1. Если разбиение P множества S является допустимой подстановкой, совместимой по выходу, то $P \leq P^{(\omega)}$.

Доказательство. Пусть s_1 и s_2 — два произвольных состояния, принадлежащие одному блоку разбиения P . Индукцией по $|a|$ покажем, что для произвольного элемента $a \in I^+$ справедливо соотношение $\delta(s_1, a) = \delta(s_2, a)$, т. е. $s_1 \approx s_2$.

При $|a| = 1$ справедливость доказываемого утверждения следует из совместимости P по выходу. Предположим, что доказываемое утверждение справедливо при $|a| \leq k$. Пусть $|a| = k + 1$. Положим $a = aa_1$, $|a_1| = k$. Тогда $\delta(s_1, aa_1) = \delta(\sigma(s_1, a), a_1)$, $\delta(s_2, aa_1) = \delta(\sigma(s_2, a), a_1)$.

Из допустимости подстановки следует, что $\sigma(s_1, a)$ и $\sigma(s_2, a)$ принадлежат одному блоку разбиения P . Отсюда и по предположению индукции получаем $\delta(s_1, aa_1) = \delta(s_2, aa_1)$.

Следствие 10.1. Пусть n — число состояний и $n \geq 2$. Тогда существует число $k \leq n-1$ такое, что

$$P^{(k)} = P^{(k+i)} = P^{(\omega)}, \quad i > 0.$$

Доказательство. Из (10.6) следует, что $P^{(1)} \geq P^{(2)} \geq \dots \geq P^{(n-1)} \geq P^{(n)}$. Если $P^{(1)}$ имеет только один блок, то оно является допустимой подстановкой, и из леммы 10.1 и соотношения 10.6 вытекает справедливость следствия; при этом $P^{(1)} = P^{(\omega)}$. Предположим, что $P^{(1)}$ имеет два или более блоков. Допустим, что $P^{(j-1)} \neq P^{(j)}$, $1 < j \leq k$. Тогда так как $P^{(j)}$ является истинным измельчением $P^{(j-1)}$, то $P^{(j)}$ содержит, по крайней мере, на один блок больше, чем $P^{(j-1)}$. С другой стороны, число блоков $P^{(j)}$ не превосходит n . Следовательно, $k \leq n-1$. Если k — наибольшее число, обладающее указанным свойством ($k=1$, если $P^{(1)} = P^{(2)}$), то $P^{(k)} = P^{(k+1)}$. Так как $k \geq 1$, то $P^{(k)}$ является совместимым по выходу. Далее, так как произвольные два состояния, s_1 и s_2 , принадлежащие одному блоку разбиения $P^{(k)}$, принадлежат также одному блоку разбиения $P^{(k+1)}$, то $\sigma(s_1, a) \approx \sigma(s_2, a)$ для произвольного символа $a \in I$. Поэтому $P^{(k)}$ — допустимая подстановка и, согласно лемме 10.1, $P^{(k)} \leq P^{(\omega)}$. Из соотношения (10.6) и утверждения (3) задачи 10.5 получаем $P^{(k)} = P^{(k+i)} = P^{(\omega)}$, т. е. доказываемое следствие справедливо также при $i > 0$.

Следствие 10.2. Если состояния s_1 и s_2 не являются эквивалентными, то существует входная последовательность α длины $n-1$ или менее, для которой $\delta(s_1, \alpha) \neq \delta(s_2, \alpha)$.

Доказательство. Согласно следствию 10.1, состояния s_1 и s_2 не могут быть $(n-1)$ -эквивалентными.

Пример 10.3. Рассмотрим конечный автомат с $S = \{0, 1, \dots, n-1\}$, $I = O = \{0, 1\}$ и функциями σ и δ . При $i \in S$, $j \in I$

$$\sigma(i, 0) = i;$$

$$\sigma(i, 1) = i+1 \quad (i+1 \text{ вычисляется по модулю } n);$$

$$\delta(i, j) = 0, \quad 0 \leq i < n-1; \quad \delta(n-1, j) = j.$$

Этот автомат является одной из разновидностей n -ичных счетчиков. Если он, находясь вначале в состоянии 0, получает входную последовательность, число символов 1 в которой кратно n , то на выходе формируется символ 1; в противном случае выход равен 0. Диаграмма состояний приведена на рис. 10.3. Поскольку для произвольного $\alpha \in I^+$, $|\alpha| \leq n-2$, $\delta(0, \alpha) = \delta(1, \alpha) = 0$, то в следствии 10.2 $n-1$ нельзя заменить на $n-2$.

Эквивалентность конечных автоматов. Рассмотрим два конечных автомата: $M = (S, I, O, \sigma, \delta)$ и $M' = (S', I, O, \sigma', \delta')$, имеющих одни и те же входные и выходные алфавиты. Если для каждого $s \in S$ существует, по крайней мере, одно состояние конечно-

го автомата M , которое эквивалентно s , то говорят, что M' покрывает M , и пишут $M \leq M'$. Если $M \leq M'$ и $M \geq M'$, то M и M' называют эквивалентными и пишут $M \approx M'$. Эквивалентные автоматы невозможно различить по выполняемым ими отображениям входных последовательностей в выходные. По определению, отношение « \approx » является отношением эквивалентности.

Предположим далее, что задан конечный автомат $M = (S, I, O, \sigma, \delta)$, и остановимся на задаче определения среди всех конечных автоматов, покрывающих заданный автомат, того, который имеет минимальное число состояний.

Лемма 10.2. Предположим, что $M \leq M' = (S', I, O, \sigma, \delta)$. Тогда число блоков $*P^{(\omega)}$ разбиения $P^{(\omega)}$ множества S с помощью отношения эквивалентности состояний M удовлетворяет неравенству

$$*P^{(\omega)} \leq |S'|.$$

Доказательство. По предположению, для каждого $s \in S$ существует, по крайней мере, одно состояние $s' \in S'$ такое, что $s \approx s'$. Пусть $\varphi(s)$ — одно из таких состояний s' . Предположим, что состояния s_1 и s_2 конечного автомата M неэквивалентны. Если допустить, что $\varphi(s_1) = \varphi(s_2)$, то $s_1 = \varphi(s_1) = \varphi(s_2) = s_2$. Получаем противоречие. Следовательно, $*P^{(\omega)} \leq |S'|$.

Лемма 10.3. Существует конечный автомат M' с числом состояний $*P^{(\omega)}$ такой, что $M \approx M'$.

Доказательство. Пусть B_1, B_2, \dots, B_l — блоки разбиения $P^{(\omega)}$. Построим $M' = (S', I, O, \sigma', \delta')$ следующим образом. Множество S' состоит из элементов $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_l$, взаимно однозначно соответствующих блокам B_1, B_2, \dots, B_l . Так как $P^{(\omega)}$ — допустимая подстановка (задача 10.7), то для произвольного элемента $a \in I$ множество $\{\sigma(s, a) | s \in B_i\}$ содержится в некотором блоке B_j ; j зависит только от i и a .

По определению, положим $\sigma'(\bar{B}_i, a) = \bar{B}_j$. Поскольку $P^{(\omega)}$, очевидно, является совместимым по выходу, то для каждого $a \in I$ все величины $\delta(s, a)$, $s \in B_i$, которые определяются равенством $\delta'(\bar{B}_i, a) = \delta(s, a)$, оказываются совпадающими. Покажем далее, что $M \approx M'$, т. е. что $\delta(s, a) = \delta'(\bar{B}_i, a)$ для произвольных $s \in B_i$ ($1 \leq i \leq l$), $a \in I^+$, так что $s \approx \bar{B}_i$.

Воспользуемся для этого индукцией по $|\alpha|$. При $|\alpha| = 1$ справедливость доказываемого утверждения следует из определения δ' . Предположим, что доказываемое утверждение справедливо при $|\alpha| \leq k$. Пусть $\alpha = a\alpha_1$ и $|\alpha_1| = k$. Из задачи 10.1 получаем $\delta(s, a\alpha) = \delta(\sigma(s, a), \alpha)$, $\delta'(\bar{B}_i, a\alpha) = \delta'(\sigma'(\bar{B}_i, a), \alpha)$. С другой стороны, если $\sigma(s, a) \in B_j$, то из определения σ' следует что $\sigma'(\bar{B}_i, a) = \bar{B}_j$. Поскольку по предположению индукции $\delta(\sigma(s, a), \alpha) = \delta'(\bar{B}_j, \alpha)$, то $\delta(s, a\alpha) = \delta'(\bar{B}_i, a\alpha)$.

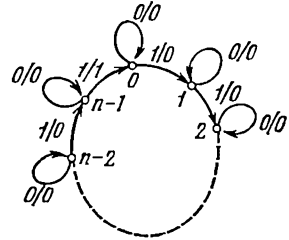


Рис. 10.3. Диаграмма состояний из примера 10.3

Конечный автомат M' , построенный при доказательстве последней леммы, далее обозначается через $M_{P(\omega)}$.

Конечный автомат M , имеющий наименьшее число состояний среди всех автоматов, которые ему изоморфны, называется *конечным автоматом с минимальным числом состояний* или *простейшим автоматом*.

Следствие 10.3. Конечный автомат M является простейшим тогда и только тогда, когда никакие два его состояния неэквивалентны.

Доказательство. Допустим, что $s_1, s_2 \in S$ и $s_1 \approx s_2$. Тогда $*P(\omega) < |S|$. В таком случае, согласно лемме 10.3, автомат не является простейшим. Получили противоречие.

Теорема 10.1. Пусть $M' = (S', I, O, \sigma', \delta')$, $M'' = (S'', I, O, \sigma'', \delta'')$ — два конечных автомата, покрывающих конечный автомат $M = (S, I, O, \sigma, \delta)$ и имеющих минимальное число состояний. Тогда

1) $M \approx M' \approx M''$;

2) существует взаимно однозначное отображение $\varphi: S' \rightarrow S$, такое, что, для произвольных $s \in S'$, $a \in I$ (рис. 10.4)

$$\varphi(\sigma'(s, a)) = \sigma(\varphi(s), a);$$

$$\delta'(s, a) = \delta''(\varphi(s), a).$$

Автоматы M и M'' , для которых выполняются эти соотношения, называются *изоморфными*. Изоморфные автоматы можно рассматривать как полностью идентичные.

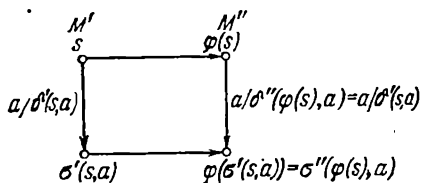


Рис. 10.4. Определение изоморфизма конечных автоматов

Доказательство. (1) Из лемм 10.2 и 10.3 следует, что как у M , так и у M' число состояний равно $*P(\omega)$. Как видно из доказательства леммы 10.2, если имеется такое состояние $s' \in S'$, для которого не существует состояния s такого, что $s \approx s'$, $s \in S$, то $|S| > *P(\omega)$ и мы пришли к противоречию.

Следовательно, $M \leq M'$. С другой стороны, по предположению, $M \leq M'$. Таким образом, $M \approx M'$. Аналогично доказывается, что $M \approx M''$; т. е. $M \approx M' \approx M''$. Это означает, что оба конечных автомата являются простейшими.

(2) Так как $M' \approx M''$, то для каждого $s' \in S'$ существует, по крайней мере, одно состояние $s'' \in S''$ такое, что $s' \approx s''$. Если имеется два или более таких состояний, то они будут эквивалентными. Однако, как показывает следствие 10.3, это противоречит тому, что конечный автомат M'' является простейшим. Аналогичное рассуждение можно провести относительно S'' . Это означает, что существует взаимно однозначное отображение φ множества S' в S'' , порождающего пары эквивалентных состояний. Поскольку $s' \approx \varphi(s')$ для каждого $s' \in S'$, то для произвольного элемента

$a\alpha \in I + \bar{\delta}'(s', a\alpha) = \bar{\delta}''(\varphi(s'), a\alpha)$. Отсюда и из задачи 10.1 получаем, что

$$\delta'(s', a) = \delta''(\varphi(s'), a);$$

$$\delta'(\sigma'(s', a), \alpha) = \delta''(\sigma''(\varphi(s'), a), \alpha),$$

так что $\sigma'(s', a) \approx \sigma''(\varphi(s'), a)$. Следовательно, $\sigma''(s'), a) = \varphi(\sigma'(s', a))$.

Таким образом получается следующий метод определения простейшего конечного автомата.

Метод определения простейшего конечного автомата. Если задан конечный автомат $M = (S, I, O, \sigma, \delta)$, то простейший эквивалентный ему конечный автомат можно найти следующим образом. Вначале находится $P^{(\omega)}$. Методом, использованным при доказательстве леммы 10.3, строится $M_{P^{(\omega)}}$. С точностью до обозначения состояний этот метод дает единственное решение.

Рассмотрим далее метод определения $P^{(\omega)}$. При этом будем пользоваться леммой 10.1.

1) Положим $P = P^{(1)}$.

2) Для каждого $a \in I$ проводится разбиение блока B_i разбиения P на подблоки; разбиение осуществляется на основании блока разбиения P , которому принадлежит состояние $\sigma(s, a)$, получаемое при входе a из состояния s , принадлежащего B_i . Если B_i разбивается на два или более подблоков, то в качестве P берется разбиение, получающееся в результате указанного разбиения B_i на подблоки. Операция (2) повторяется. Если B_i невозможно разбить на подблоки ни по одному из входов, берется следующий блок. Если не удастся разбить ни по одному из входов ни один подблок, то получившиеся в результате разбиения P является допустимой подстановкой.

3) Разбиение P совпадает с $P^{(\omega)}$.

Задача 10.8. Покажите, что с помощью описанного метода действительно находится $P^{(\omega)}$.

Решение. Так как число блоков не может превышать $|S|$, то описанный процесс должен закончиться. Поскольку при этом разбиение $P^{(1)} = P$ последовательно измельчается, то $P^{(1)} = P_1 > P_2 > P_3 > \dots > P_n$. Для двух состояний, s_1 и s_2 , принадлежащих различным блокам P_i , имеем $s_1 \sim s_2^*$. В этом можно убедиться индукцией по i . Действительно, при $i=1$ сформулированное утверждение очевидно. Предположим, что оно справедливо при $1 \leq i < j$, и рассмотрим случай $i=j$. Так как s_1, s_2 принадлежат одному блоку разбиения P_{j-1} и при некотором входном символе a состояния $\sigma(s_1, a)$ и $\sigma(s_2, a)$ принадлежат различным блокам разбиения P_{j-1} , то достаточно рассмотреть случай, когда они принадлежат различным блокам разбиения P_j . По предположению, $\sigma(s_1, a) \sim \sigma(s_2, a)$. Согласно задаче 10.4, $s_1 \sim s_2$. С другой стороны, как следует из леммы 10.1, конечное разбиение P является измельчением $P^{(\omega)}$ и, следовательно, $P = P^{(\omega)}$.

Пример 10.4. Найдем простейший вид конечного автомата с $S = \{1, 2, \dots, 8\}$, $I = O = \{0, 1\}$, функции σ и δ которого задаются табл. 10.3.

* Запись $s_1 \sim s_2$ означает, что состояния s_1 и s_2 неэквивалентны.

Вначале имеем $P_1 = P^{(1)} = \{(1, 2, 4, 5, 7, 8), (3, 6)\}$. Взяв первый блок и рассмотрев входной символ 0, получаем $P_2 = \{(1, 4, 8), (2, 5, 7), (3, 6)\}$. Вновь обратившись к первому блоку и рассмотрев тот же входной символ 0, получаем $P_3 = \{(1, 4), (3), (2, 5, 7), (3, 6)\}$. Так как P_3 уже допустимая подстановка, то $P^{(\omega)} = P_3$. Если ввести состояния s_1, s_2, s_3 и s_4 , соответствующие блокам $(1, 4), (8), (2, 5, 7)$ и $(3, 6)$, то с помощью процедуры, описанной при доказательстве леммы 10.3, можно найти $\sigma_{P^{(\omega)}}, \delta_{P^{(\omega)}}$. Таблица 10.4 представляет собой получающуюся в результате таблицу этих функций.

Т а б л и ц а 10.3. Таблица определения σ, δ из примера 10.4

Состояние	Входные символы	
	0	1
1	8,0	3,0
2	3,0	5,0
3	1,1	3,0
4	8,0	6,0
5	3,0	2,0
6	4,1	6,0
7	6,0	7,0
8	7,0	1,0

Т а б л и ц а 10.4. Таблица функций $\sigma_{P^{(\omega)}}, \delta$

Состояние	Входные символы	
	0	1
s_1	$s_2, 0$	$s_4, 0$
s_2	$s_3, 0$	$s_1, 0$
s_3	$s_4, 0$	$s_3, 0$
s_4	$s_1, 1$	$s_4, 0$

Что касается числа операций, необходимых для определения $P^{(\omega)}$ описанным способом, то: 1) число операций, выполняемых при переходе от P_{j-1} к P_j , в худшем случае пропорционально $|S| \times |I|$; 2) так как процесс измельчения заканчивается самое большое через $|S| - 1$ шагов, то общее число операций не превосходит по порядку $|S|^2 \times |I|$. Известны методы определения $P^{(\omega)}$, требующие выполнения порядка $|S| \log_2 |S|$ операций, где $|S|$ — число состояний [25].

Задача 10.9. Постройте диаграммы состояний для конечного автомата из примера 10.4 и его простейшего представления.

Определение эквивалентности двух конечных автоматов. Предположим, что заданы два конечных автомата: $M = (S, I, O, \sigma, \delta)$ и $M' = (S', I, O, \sigma', \delta')$. Изменив, в случае необходимости, обозначения состояний S' , предположим, что $S \cap S' = \emptyset$. Определим конечный автомат $M'' = (S \cup S', I, O, \sigma'', \delta'')$ следующим образом. Для каждого $a \in I$ положим

$$\sigma''(s'', a) = \sigma(s'', a), \delta''(s'', a) = \delta(s'', a),$$

если $s'' \in S$, и

$$\sigma''(s'', a) = \sigma'(s'', a), \delta''(s'', a) = \delta'(s'', a),$$

если $s'' \in S'$. Пусть P'' — разбиение множества S'' на классы эквивалентности множества S'' , соответствующее отношению эквивалентности между состояниями M'' . Из построения P'' следует, что для любых $s \in S, s' \in S'$ отношение $s \approx s'$ имеет место тогда и только тогда, когда s и s' принадлежат одному блоку разбиения P'' . Отсюда и определения отношения $M \leq M'$ получаем следующее следствие.

Следствие 10.4. $M \leq M'$ тогда и только тогда, когда все блоки разбиения P'' содержат, по крайней мере, по одному состоянию из S' .

Известны также методы определения эквивалентности и минимизации для не полностью определенных конечных автоматов. Одним из таких методов является метод Унгера [37]. Однако среди них нет методов, которые были бы столь же эффективны, как описанный.

10.3. Синхронные конечные автоматы [17—24, 37]

Цифровые устройства, не удовлетворяющие условиям в определении схемы из функциональных элементов (§ 8.1), называются *последовательными схемами*. Последовательные схемы произвольного вида называют также *асинхронными схемами*. Поскольку скорость распространения сигнала в различных элементах схемы обычно различна, а также по многим другим причинам, в асинхронных схемах происходят довольно сложные переходные процессы [37].

Пример 10.5. (Триггер-защелка SR типа). Рассмотрим НЕ—ИЛИ-схему, изображенную на рис. 10.5. Пусть d_1, d_2 — временные задержки двух элементов НЕ—ИЛИ. Предположим, что в момент времени $t_0 = \max(d_1, d_2)$ значения внешних входов S, R фиксируются и далее оставляются неизменными. Обозначим через $Q_1(t), Q_2(t)$ значения внешних выходов Q_1, Q_2 в момент времени t . Будем считать, что временные задержки, вносимые проводниками, включаются в задержки соответствующих элементов, и поэтому ими можно пренебречь.

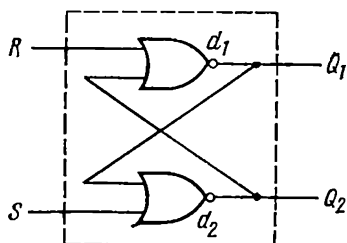


Рис. 10.5. Триггер-защелка SR -типа на элементах НЕ—ИЛИ

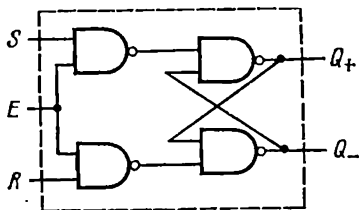


Рис. 10.6. Триггер SR -типа

1) Пусть $S=R=0$. Предположим, что в интервале времени $t_0 - \max(d_1, d_2) \leq t \leq t_0$ значения функций $Q_1(t), Q_2(t)$ не изменяются и $Q_1(t) = \overline{Q_2(t)}$. Очевидно, что $Q_1(t) = Q_1(t_0), Q_2(t) = \overline{Q_2(t_0)}$ и при $t_0 \leq t$. Таким образом, текущие значения внешних выходов зависят не только от внешних входов, но и предшествующих значений внешних выходов. Если считать, что $Q_1(t_0) = \overline{Q_2(t_0)}$, то $Q_1(t_0 + j(d_1 + d_2)) = Q_1(t_0), Q_2(t_0 + j(d_1 + d_2)) = \overline{Q_2(t_0)}$.

$=\overline{Q_i(t_0)}$ ($i=1, 2$; $j=0, 1, 2, \dots$), т. е. даже при фиксированных внешних входах внешние выходы могут меняться. Поскольку точные значения задержек d_1 и d_2 обычно не известны, а кроме того, они могут меняться при длительной работе устройства, в данном случае выходные значения предсказать правильно трудно. Такие выходные значения называются *неустойчивыми**.

2) Пусть $S=1, R=0$. С момента времени t_0 выход $Q_2=0$. Следовательно, с момента времени t_0+d_1 выход $Q_1=1$. При $S=0$ и $R=1$ индексы 1 и 2 меняются местами.

3) Пусть $S=R=1$. С момента времени t_0 оба выхода $Q_i=0$ ($i=1, 2$).

4) Положим $d=\max(d_1, d_2)$. Предположим, что разрешенными наборами значений внешних входов S, R являются $(0, 0)$, $(0, 1)$ и $(1, 0)$. Допустим, что каждый набор значений входов удерживается на входе схемы в течение интервала времени $2d$ или больше. Предположим также, что начальным набором значений входов является либо набор $(0, 1)$, либо $(1, 0)$. Как следует из изложенного, в течение интервала времени $2d$ после изменения значений входов устанавливаются входные значения, и до тех пор, пока входные значения вновь не будут изменены, не

изменяются и выходные значения (последние называются устойчивыми). В табл. 10.5 приведены выходные значения, соответствующие различным наборам значений входов. Через Q'_1 в этой таблице обозначено значение Q_1 , соответствующее предшествующему набору значений входов. Обычно $Q_1=\overline{Q_2}$. Поскольку выход-

Таблица 10.5. Таблица связи выходов со входами триггера-защелки RS-типа

S	R	Q_1	Q_2
0	0	Q'_1	$\overline{Q'_1}$
0	1	0	1
1	0	1	0
1	1	Запрещенный вход	

ные значения зависят от Q'_1 , то рассмотренная схема в некотором смысле обладает памятью.

Пример 10.6. Пусть d — максимальное значение задержки элементов схемы, изображенной на рис. 10.6. Предположим, что если значение на внешнем входе $E=1$, то в течение интервала времени длительностью $2d$ или более значения на внешних входах S, R остаются неизменными и могут быть любыми, кроме $(1, 1)$. При этом входы связаны с выходами так, как описано в примере 10.5. Отличие здесь лишь в том, что выходы принимают указанные значения с задержкой во времени, которая в данном случае больше величины задержки элементов глубины 1. При $E=0$ выходные значения фиксируются и в дальнейшем остаются

* В данном случае рассматривается идеализация, основанная на понятии чистой задержки. Однако в реальных схемах на выходе обычно устанавливается одно из состояний $(Q_1, Q_2)=(0, 1)$ или $(1, 0)$. Однако при этом невозможно установить, в какое именно из этих двух состояний перейдет схема.

неизменными. Эта схема называется *триггером SR-типа*, или просто *SR-триггером*. В реальных условиях выходные значения могут устанавливаться по входам S и R в течение очень короткого промежутка времени, когда значение E изменяется с 0 на 1 (или наоборот). Такие схемы называются триггерами с динамическим входом. Кроме триггеров с динамическим входом существуют также триггеры *MS-типа** и др.

Триггеры. Триггеры различных типов — важные конструктивные компоненты последовательных схем. Они могут иметь несколько информационных входов z_1, \dots, z_l (обычно $l=1$ или 2), вход E , на который подается тактовый сигнал (далее сигнал E), вход для установки начального значения и два выхода, Q_+ , Q_- (в устойчивом состоянии значение одного выхода является дополнением к значению другого). Обозначим через $z_i(t)$, $E(t)$, $Q_+(t)$, $Q_-(t)$ значения z_i , E , Q_+ , Q_- в момент времени t (значениями этих функций являются символы 0 и 1, показывающие, что соответствующий сигнал не превышает или превышает заданный порог). Изменение значения E с a на \bar{a} будет называться далее E -переходом $a \rightarrow \bar{a}$. Функционирование триггера описывается следующей совокупностью характеристик: положительными постоянными τ_s (время установления**), τ_h (время удержания**), d (время распространения), двоичной постоянной a (полярность), логической функцией g от $l+1$ переменных (характеристическая функция). Эти характеристики связаны между собой следующим образом.

Для установления начальных значений все входы удерживаются в течение определенного времени в фиксированных состояниях, после чего осуществляется E -переход $a \rightarrow \bar{a}$. В течение этого промежутка времени $Q_+ = Q_-$ и значения Q_+ , Q_- остаются постоянными (при этом говорят о том, что происходит установление не выхода). После этого происходит следующее.

Пусть t_1 — момент некоторого E -перехода $a \rightarrow \bar{a}$, и t_2 — момент следующего за ним E -перехода $\bar{a} \rightarrow a$. Предположим, что $t_1 + \tau_h \leq t_2$; в момент t_1 устанавливается выход. По способу реакции на сигнал E различают следующие два типа триггеров.

(ЗМ) Триггеры *MS-типа*.

а) В интервале времени от $t_1 - \tau_s$ до t_2 значения входов z_1, \dots, z_l остаются постоянными.

б) Если в интервале времени $t_1 - \tau_s \leq t < t_1$ функция $E(t) = a$, то в интервале времени от t_1 до t_2

$$Q_+(t) = \overline{Q_-(t)} = Q_+(t_1). \quad (10.7)$$

в) Если в промежутке от t_2 до $t_2 + d$ происходит следующий E -переход $a \rightarrow \bar{a}$, то, начиная с момента $t_2 + d$ до возникновения

* В оригинале в качестве терминов, переводимых в данной книге как триггер с динамическим входом и триггер *MS-типа*, были использованы английские термины «edge triggered flip-flop», «master-slave flip-flop».

** Терминология здесь не является установившейся.

следующего E -перехода $a \rightarrow \bar{a}$, в любой момент времени t выход определяется равенством

$$Q_+(t) = \overline{Q_-(t)} = g(Q_+(t_1), z_1(t_1), \dots, z_l(t_1)). \quad (10.8)$$

(3Е) Триггеры с динамическим входом.

а) С момента времени $t_1 - \tau_s$ до $t_1 + \tau_h$ значения входов остаются постоянными.

б) Если в интервале времени от t_2 до $\max(t_2, t_1 + d)$ не происходит E -перехода $a \rightarrow \bar{a}$, то в произвольный момент времени t до следующего E -перехода $a \rightarrow \bar{a}$ выход определяется соотношением (10.8).

Если хотя бы одно из указанных условий (которые называются условиями функционирования) не выполняется, то выходные значения не определены.

1) D -триггер: $l=1$, вместо z_1 используется символ D и

$$g(Q_+, D) = D, \quad (10.9)$$

т. е. значение $Q_+(t)$ равно входу D в момент времени, непосредственно предшествующий последнему E -переходу $a \rightarrow \bar{a}$. Этот триггер выполняет функции элемента задержки.

2) SR -триггер: $l=2$, вместо z_1, z_2 используются символы S, R и функция g определяется табл. 10.6 (такие таблицы называются характеристическими). При $S(t_1)R(t_1)=1$ SR -триггер не функционирует. Следовательно, в качестве входных сигналов должны использоваться только такие сигналы, для которых $S(t_1)R(t_1)=0$. Входы S и R используются соответственно для установки триггера в состояния 0 и 1. Из характеристической таблицы SR -триггера и условия $SR=0$ следует, что

$$\begin{aligned} g(Q_+, S, R) &= \bar{S}\bar{R}Q_+ \vee S\bar{R} = \bar{S}\bar{R}Q_+ \vee S\bar{R} \vee SR = \\ &= \bar{S}RQ_+ \vee S = \\ &= \bar{R}Q_+ \vee S, \quad SR=0. \end{aligned} \quad (10.10)$$

3) JK -триггер: $l=2$, вместо z_1 и z_2 используются соответственно символы J и K . Характеристической таблицей JK -триггера является табл. 10.7. Эти триггеры отличаются от SR -триггеров тем, что при $J=K=1$ функция g определена и равна \bar{Q}_+ . При этом

$$g(Q_+, J, K) = \bar{K}Q_+ \vee J\bar{Q}_+. \quad (10.11)$$

Таблица 10.6. Характеристическая таблица SR -триггера

$S(t_1)$	$R(t_1)$	g
0	0	$Q_+(t_1)$
0	1	0
1	0	1
1	1	Не определена

Таблица 10.7. Характеристическая таблица JK -триггера

$J(t_1)$	$K(t_1)$	g
0	0	$Q_+(t_1)$
0	1	0
1	0	1
1	1	$\bar{Q}_+(t_1)$

4) *T*-триггер: $l=1$, вместо z_1 используется символ *T*, а функция g получается из характеристической функции *JK*-триггера подстановкой $J=K$. Из (10.11) следует, что

$$g(Q_+, T)g(Q_+T) = Q_+ \oplus T. \quad (10.12)$$

Задача 10.10. Покажите, что если в *SR*-триггере положить $S=\bar{R}=D$, то получится *D*-триггер.

Решение следует из соотношений (10.9) и (10.10).

Задача 10.11. На рис. 10.7 приведена схема *SR*-триггера *MS*-типа. Поясните принцип ее работы.

Решение. Изображенный на рис. 10.7 *SR*-триггер *MS*-типа состоит из двух триггеров: *M* и *S*. При E_2 -переходе $1 \rightarrow 0$ вначале устанавливается выход триггера *S*. Далее при E_1 -переходе $0 \rightarrow 1$ входы считываются в триггер *M*. При следующем E_2 -переходе $0 \rightarrow 1$ выходные значения триггера *M* переписываются в триггер *S*, при следующем за ним E -переходе $1 \rightarrow 0$ внешние входы триггера *S* отключаются от триггера *M*.

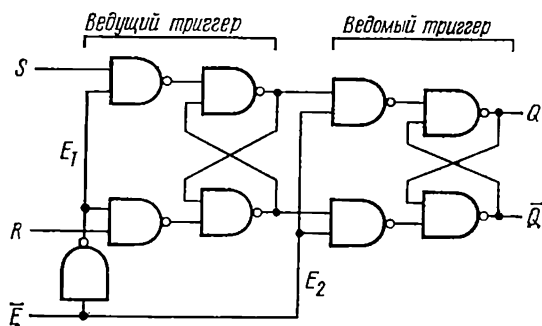


Рис. 10.7. Схема *SR*-триггера *MS*-типа

Синхронные последовательные схемы. Такие схемы имеют несколько меньшее быстродействие, но поскольку они выгодно отличаются простотой проектирования, то их широко используют на практике. В синхронных последовательных схемах синхронизация осуществляется посредством подачи последовательности тактирующих импульсов, идеальный вид которых показан на рис. 10.8. Обозначим через ω , Δ соответственно ширину импульса и интервал между тактирующими импульсами. Выбрав соответствующим образом моменты времени t_j , будем считать, что j -й переход $a \rightarrow \bar{a}$ происходит в момент времени t_j . Очевидно, $t_{j+1} - t_j = \omega - \Delta$.

Логические схемы со структурой, показанной на рис. 10.9, называются *синхронными последовательными схемами*. На этом рисунке C — это схема из функциональных элементов с $n+2k$ входами и $m+p$ выходами. Пусть $f_i (1 \leq i \leq m+p)$ — функции, реализуемые этой схемой. Для простоты предположим, что все триггеры F_1, F_2, \dots, F_k имеют по два входа ($l=2$) и полярность a . Обозначим через τ_s, τ_h и d соответственно максимальные зна-

чения времени установления, времени удержания и задержки распространения. Общие тактирующие импульсы подаются на входы E всех триггеров. Информационные входы триггеров F_1, \dots, F_k обозначаются соответственно через z_1, \dots, z_{2k} ($p=2k$). Максимальное и минимальное значения задержки схемы S обо-

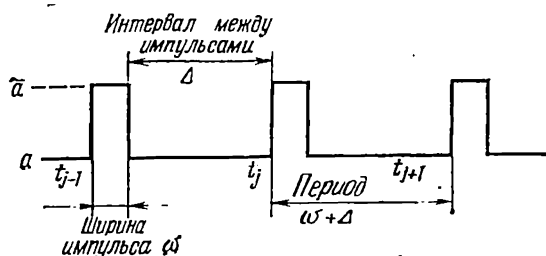


Рис. 10.8. Тактирующие импульсы

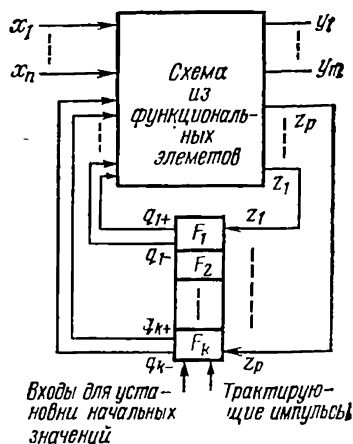


Рис. 10.9. Синхронная последовательная схема

значаются через d_{\max} и d_{\min} . Будем предполагать что все триггеры являются триггерами MS -типа. Для того чтобы выполнялись описанные условия функционирования триггеров, должны также соблюдаться следующие условия.

1) Начальное значение каждого триггера F_i устанавливается в момент времени $t_0 + \omega + d$.

2) $\tau_h < \omega$ — условие (1) функционирования триггера.

3) Для выполнения условия ЗМ (а) необходимо, чтобы значения z_1, \dots, z_{2k} определялись в момент $t_j - \tau_s$ ($j \geq 1$) и удерживались до момента $t_j + \omega$ по значениям $q_{1+}, q_{1-}, \dots, q_{k+}, q_{k-}$ выходов триггеров, которые устанавливаются в момент $t_{j-1} + \omega + d$ и удерживаются неизменными до $t_j + \omega$ (далее эти значения q_{i+} обозначаются через $q_i(j)$) по значениям внешних входов x_1, \dots, x_n . Для этого, в свою очередь, необходимо, чтобы выполнялось неравенство $d + d_{\max} + \tau_s \leq \Delta$, а кроме того, при каждом j внешние входы x_1, \dots, x_n оставались неизменными с момента времени $t_j - d_{\max} - \tau_s$ до $t_j + \omega - d_{\min}$ (эти значения x_i обозначаются далее через $x_i(j)$). При выполнении этих условий требования условия ЗМ (б) и (в) выполняются автоматически.

Условия (1)—(3) называются условиями функционирования рассматриваемых здесь синхронных последовательных схем. Как будет показано, при выполнении этих условий синхронные последовательные схемы с точки зрения реализуемых ими преобразований входных последовательностей в выходные можно рассматривать как конечные автоматы (в условиях функционирования, вообще говоря, не требуется, чтобы входной сигнал E был периодическим).

Определим функции δ_i , h_i следующим образом:

$$\delta_i(x_1, \dots, x_n, q_1, \dots, q_k) = f_i(x_1, \dots, x_n, q_1, \bar{q}_1, q_2, \bar{q}_2, \dots, q_k, \bar{q}_k), \\ 1 \leq i \leq m;$$

$$h_i(x_1, \dots, x_n, q_1, \dots, q_k) = f_i(x_1, \dots, x_n, q_1, \bar{q}_1, q_2, \bar{q}_2, \dots, q_k, \bar{q}_k), \\ m+1 \leq i \leq m+p.$$

Ввиду условия (3) значения внешних выходов y_1, \dots, y_m остаются постоянными в интервале времени с $t_j - \tau_s$ до $t_j + \omega$. Обозначим эти значения через $y_1(j), \dots, y_m(j)$. При этом

$$y_i(j) = \delta_i(x_1(j), \dots, x_n(j), q_1(j), \dots, q_k(j)); \quad (10.13)$$

$$z_i(j) = h_i(x_1(j), \dots, x_n(j), q_1(j), \dots, q_k(j)). \quad (10.14)$$

Величины z_i , h_i называют соответственно *переменными возбуждения* и *функциями возбуждения*. Если g_i — характеристическая функция триггера F_i , то

$$q_i(j+1) = q_i(q_i(j), z_{2i-1}(j), z_{2i}(j)). \quad (10.15)$$

По значениям входов $X(j) = (x_1(j), \dots, x_n(j))$ и внутренним состояниям $Q(j) = (q_1(j), \dots, q_k(j))$ значения выходов $Y(j) = (y_1(j), \dots, y_m(j))$ определяются соотношением (10.13), а следующие внутренние состояния $Q(j+1) = (q_1(j+1), \dots, q_k(j+1))$ — соотношениями (10.14) и (10.15). Величины $q_i (1 \leq i \leq k)$ называются *переменными состояниями*.

Пусть $S = V_h$, $I = V_n$, $O = V_m$, $\delta = (\delta_1, \delta_2, \dots, \delta_m)$ и σ — это отображение $S \times I$ в S , определяемое соотношениями (10.14) и (10.15). Тогда синхронную последовательную схему, изображенную на рис. 10.9, можно рассматривать как конечный автомат $(S, I, O, \sigma, \delta)$, определив $x_i(j)$, $y_i(j)$ и $q_i(j)$ так, как было указано. Если все триггеры являются D -триггерами, то $\sigma = (h_1, h_2, \dots, h_k)$.

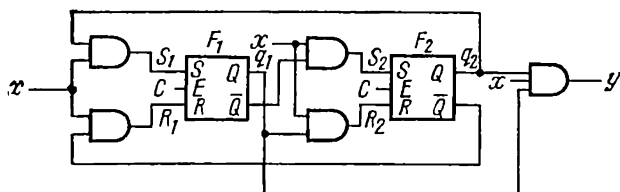
Пример 10.7. Рассмотрим синхронную последовательную схему с одним входом и одним выходом, изображенную на рис. 10.10. Здесь F_1 и F_2 — SR -триггеры, C — тактирующие импульсы. Из приведенной на рис. 10.10 схемы видно, что

$$y(j) = x(j) q_1(j) q_2(j);$$

$$S_1(j) = x(j) q_2(j), \quad R_1(j) = x(j) \bar{q}_2(j);$$

$$S_2(j) = x(j) \bar{q}_1(j), \quad R_2(j) = x(j) q_1(j).$$

Условие $S_1 R_1 = S_2 R_2 = 0$ выполняется. Символы q_1 , q_2 представляют собой переменные состояния, а пары их значений — внутренние состояния. По характеристической таблице SR -триггера (табл. 10.6) строится таблица переходов и выходных значений (табл. 10.8). Как следует из приведенных равенств, если $x(j) = 0$, то $S_1(j) = R_1(j) = S_2(j) = R_2(j) = 0$ и, следовательно, $q_1(j+1) = q_1(j)$, $q_2(j+1) = q_2(j)$, т. е. состояние не изменяется. Если внутренним состояниям $(1, 0)$, $(0, 0)$, $(0, 1)$, $(1, 1)$ сопоставить состояния 1, 2, 3, 4, то диаграмма состояний рассматриваемой схе-



Нижний элемент представляет собой SR-триггер

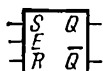


Рис. 10.10. Схема из примера 10.10.
Нижний элемент представляет собой SR-триггер

мы будет в точности совпадать с диаграммой состояний, приведенной на рис. 10.3 при $n=4$. А именно, находясь первоначально в состоянии 0, конечный автомат формирует на выходе символ 1, если на его вход поступает входная последовательность с равным 1 последним символом, число символов 1 в которой кратно 4 и 0 остальных случаях. Эта схема представляет собой один из видов четверичных счетчиков.

Задачи, возникающие при построении последовательных схем. Задан конечный автомат $M(S, I, O, \sigma, \delta)$. Требуется построить последовательную схему, которая его реализует. Обычно вначале конечный автомат M упрощают (вообще говоря, это делать не обязательно, но часто желательно).

Таблица 10.8. Таблица значений $(q_1(j+1), q_2(j+1)), y(j)$

$q_1(j), q_2(j)$	$x(j)$	
	0	1
(0, 0)	(0, 0), 0	(0, 1), 0
(0, 1)	(0, 1), 0	(1, 1), 0
(1, 0)	(1, 0), 0	(0, 0), 0
(1, 1)	(1, 1), 0	(1, 0), 1

1) Выбираются целые числа $k \geq \log_2 |S|$, $n \geq \log_2 |I|$, $m \geq \log_2 |O|$ и инъективные отображения $\varphi_S: S \rightarrow V_k$; $\varphi_I: I \rightarrow V_n$; $\varphi_O: O \rightarrow V_m$. Задача возникает уже при выборе отображений φ_I и φ_O . Задача выбора φ_S называется задачей распределения состо-

яний. Дело в том, что часто в зависимости от выбора отображения φ_S могут получаться конечные автоматы различной сложности. Этому вопросу посвящен ряд исследований [37], но до сих пор общего эффективного метода решения этой задачи нет.

2) Определим отображение $\sigma': V_k \times V_n \rightarrow V_k$ следующим образом. Для каждого $s \in S$ и каждого $a \in I$ определим $\sigma'(\varphi_S(s))$, $\varphi_I(a)$ с помощью равенства

$$\sigma'(\varphi_S(s), \varphi_I(a)) = \varphi_S(\sigma(s, a)). \quad (10.16)$$

* Отображение φ множества A в множество B называется инъективным, если $\varphi(a_1) \neq \varphi(a_2)$ для произвольных двух различных элементов, a_1 и a_2 , из A . Область значений отображения φ обозначается через $\varphi(A)$.

На множестве $V_k \times V_n = \varphi_S(S) \times \varphi_I(I)$ будем считать σ' неопределенным. Аналогично определяется отображение $\delta': V_k \times V_n \rightarrow \varphi_O(O)$. А именно, для каждого $s \in S$ и каждого $a \in I$ положим

$$\delta'(\varphi_S(s), \varphi_I(a)) = \varphi_O(\delta(s, a)). \quad (10.17)$$

На множестве $V_k \times V_n = \varphi_S(S) \times \varphi_I(I)$ это отображение также будем считать неопределенным.

3) Пусть x_1, x_2, \dots, x_n — входные переменные, y_1, y_2, \dots, y_m — выходные переменные и q_1, q_2, \dots, q_k — переменные состояния. Определим тип используемых триггеров. Рассмотрим случай $l=2^*$. Пусть z_1, z_2, \dots, z_{2k} — переменные возбуждения. Задача состоит в построении комбинационной схемы, показанной на рис. 10.9. Обозначим через $\delta'_1, \delta'_2, \dots, \delta'_m$ компоненты δ' . При этом

$$y_i = \delta'_i(x_1, \dots, x_n, q_1, \dots, q_k), \quad 1 \leq i \leq m. \quad (10.18)$$

4) Пусть $\sigma'_1, \sigma'_2, \dots, \sigma'_k$ — компоненты σ' и g_i — характеристическая функция i -го триггера. Как следует из (10.14) и (10.15), нужные функции возбуждения достаточно определить так, чтобы они удовлетворяли условию

$$\sigma'_i(q_1, \dots, q_k, x_1, \dots, x_n) = g_i(q_i, h_{2i-1}, h_{2i}). \quad (10.19)$$

После того как заданы (частичные) логические функции σ'_i и g_i , задача заключается в определении по (10.19) (частичных) логических функций h_{2i-1} и h_{2i} . Для этого следует для каждого набора $(b_1, b_2, \dots, b_k) \in V_k$ и каждого набора $(a_1, a_2, \dots, a_n) \in V_n$ определить пару $(c_1, c_2) \in V_2$ так, что

$$\sigma'_i(b_1, \dots, b_k, a_1, \dots, a_n) = g_i(b_i, c_1, c_2), \quad (10.20)$$

и функции h_{2i-1}, h_{2i} , для которых $h_{2i-1}(a_1, \dots, a_n, b_1, \dots, b_k) = c_1$, $h_{2i}(a_1, \dots, a_n, b_1, \dots, b_k) = c_2$. Обычно существует несколько пар (c_1, c_2) , удовлетворяющих условию (10.20).

Пример 10.8. Рассмотрим SR-триггер. Если положить $\sigma'_i(b_1, \dots, b_k, a_1, \dots, a_n) = b'_i$, то из (10.10) будет следовать, что

$$b'_i = \bar{c}_2 b_i \vee c_1, \quad c_1 \cdot c_2 = 0,$$

т. е. при заданных Q'_+ и Q_+ задача заключается в определении совокупности S, R , удовлетворяющей условию $Q'_+ = \bar{R}Q_+ \vee S$, $S \cdot R = 0$. Например, если $Q'_+ = Q_+ = 0$, то $S = 0$, а R можно выбрать равным как 0, так и 1 (обозначим выбранное значение R через d). Если $Q'_+ = 1, Q_+ = 0$, то $S = 1, R = 0$. Таким образом находится полное решение задачи (табл. 10.9). Эта таблица называется *таблицей возбуждения* SR-триггера. Запись $Q'_+ = d$ означает, что левая часть равенства (10.10) не определяется (обычно σ' является частичной функцией). Как видно из последних двух строк таблицы, пару (S, R) , удовлетворяющую условию $SR = 0$, можно выбрать тремя способами. Аналогично строятся

* При $l=1$ или если у разных триггеров значения различны, все происходит аналогично.

Т а б л и ц а 10.9. Таблица возбуждения SR -триггера

Q_+	Q'_+	S	R
0	0	0	d
0	1	1	0
1	0	0	1
1	1	d	0
0	d	d	d
1	d	d	d

$\}SR=0$

Т а б л и ц а 10.10. Таблица возбуждения JK -триггера

Q_+	Q'_+	J	K
0	0	0	d
0	1	1	d
1	0	d	1
1	1	d	0
0	d	d	d
1	d	d	d

таблицы возбуждения JK -триггера и T -триггера (табл. 10.10 и 10.11).

Задача 10.12. Проверьте, что табл. 10.10 и 10.11 построены правильно.

Т а б л и ц а 10.11. Таблица возбуждения T -триггера

Q_+	Q'_+	T
0	0	0
0	1	1
1	0	1
1	1	0
0	d	d
1	d	d

Т а б л и ц а 10.12. Таблица значений функций J_1, K_1, J_2, K_2 из примера 10.9

$q_1(j), q_2(j)$	$x(j)$	
	0	1
(0, 0)	0, d 0, d	0, d 1, d
(0, 1)	0, d d , 0	1, d d , 1
(1, 0)	d , 0 0, d	d , 0 1, d
(1, 1)	d , 0 d , 0	d , 1 d , 1

Пример 10.9. Рассмотрим конечный автомат из примера 10.3 при $n=4$, когда в качестве внутреннего состояния, соответствующего состоянию i , используется двоичное представление числа i , а в качестве F_1 и F_2 используются SR -триггеры. Обозначим через J_1, K_1, J_2, K_2 соответствующие входы триггеров F_1 и F_2 и найдем функции возбуждения, которые определяют их значения. Например, при $q_1(j), q_2(j)=(0, 0)$, $x(j)=1$ имеем $q_1(j+1)=0$, $q_2(j+1)=1$. Отсюда и из табл. 10.10, в свою очередь, следует, что $J_1(j)=0$, $K_1(j)=d$, $J_2(j)=1$, $K_2(j)=d$. Таким образом получается табл. 10.12, определяющая значения функций J_1, K_1, J_2, K_2 . Если подобрать соответствующим образом значение d , то $J_1=K_1=xq_2$, $J_2=K_2=x$ (здесь опущен аргумент (j)). Так как $J_1=K_1$, $J_2=K_2$, то автомат аналогичен T -триггеру.

Задача 10.13. Постройте JK -триггер из SR -триггеров, рассматривая его как синхронную последовательную схему.

Указание. Постройте таблицу значений функций возбуждения. Далее подберите значение дополнительного входа d так, что $S=JQ_+$, $R=KQ_+$.

Если использовать последовательные ПЛМ, представляющие собой обычные ПЛМ (см. пример 9.2), и некоторую совокупность

триггеров (схема из функциональных элементов на рис. 10.9 заменяется на ПЛМ), то можно реализовать многие достаточно сложные последовательные схемы соединением элементов И в решетке [21].

Задачи

10.1. Пусть $I=O=\{0, 1\}$, $\alpha \in I^+$ и функция $f(\alpha)$ равна 1, если в последовательности α содержится одинаковое число символов 0 и 1; в противном случае $f(\alpha)=0$. Докажите, что не существует конечного автомата, который давал бы то же преобразование входных последовательностей в выходные, что и преобразователь (I, O, f) .

10.2. Пусть $I=\{r$ (повторная установка), s (тактовый сигнал), $O=\{0, 1\}$ и $f(\alpha)=1$ для $\alpha=a_1rc^n$, где $a_1 \in I^*$, n — целое число, являющееся квадратом другого целого числа, и $f(\alpha)=0$ для остальных α . Докажите, что не существует конечного автомата, который давал бы то же преобразование входных последовательностей в выходные, что и преобразователь (I, O, f) .

10.3. Для заданного конечного автомата Мили $M=(S, I, O, \sigma, \delta)$ существует конечный автомат Мура $M'=(S', I, O, \sigma', \delta')$, обладающий следующим свойством. Существует отображение $\varphi: S \rightarrow S'$, такое, что для произвольных $s \in S$, $\alpha \in I^+$, $\alpha \in I$, $\delta(s, \alpha) = \delta'(\varphi(s), \alpha\alpha)$.

10.4. Пусть $M=(S, I, O, \sigma, \delta)$ — конечный автомат. Для $s \in S$ и произвольного $\alpha \in I^+$ положим $f_s(\alpha) = \delta(s, \alpha)$. Покажите, что f_s является регулярной.

10.5. Найдите дизъюнктивную нормальную форму минимальной сложности выходной функции и функции возбуждения при реализации на SR-триггерах параллельного двоичного сумматора из примера 10.1.

10.6. В предположении, что нельзя использовать отрицания x и y , реализуйте s, S, R из предыдущей задачи на элементах НЕ—И.

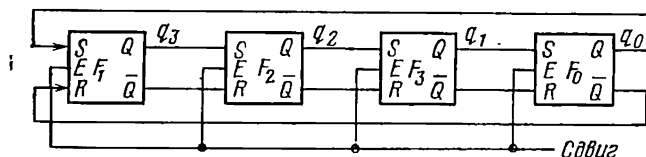


Рис. 10.11. Односторонний циклический регистр сдвига

10.7. На рис. 10.11 показан четырехкаскадный односторонний циклический регистр сдвига на SR-триггерах MS-типа. Объясните принцип его работы (входы, используемые для установки начальных значений регистров, на рисунке не изображены).

10.8. Найдите таблицы значений функции возбуждения триггеров T_1 и T_2 , если в примере 10.9 в качестве F_1 и F_2 используются T -триггеры. Начертите схему, используя элементы И.

10.9. Пусть $I\{[,]\}$, $O=\{0, 1\}$, $a \in I^*$ и $f(a)=1$, если левые и правые скобки расставлены в a правильно, и $f(a)=0$ в противном случае. Покажите, что функция f не является регулярной и не существует конечного автомата, который давал бы то же отображение входных последовательностей в выходные, что и преобразователь (I, O, f) .

Решения задач

К главе 1:

1.1. 1) 17, 2) 23.

1.2. Положим $d=(ac, b)$. Очевидно, $(b, c)|d$. Так как $d|ac$ и $d|bc$, то $d|(ac, bc)=(c)$. Отсюда и из делимости $d|b$, $d|c$ получаем $d|(b, c)$, т. е. $d=(b, c)$.

1.3. Решение первого уравнения b_1+m_1y . Для того чтобы оно удовлетворяло второму уравнению, должно выполняться сравнение $m_1y \equiv b_2-b_1 \pmod{m_2}$. Из теоремы 1.16 следует, что решение существует. Заметим, что если x_1 и x_2 — решения системы уравнения, то x_1-x_2 кратно как m_1 , так и m_2 . Следовательно, оно кратно $[m_1, m_2]$. Согласно (1.11), $[m_1, m_2]=m_1m_2$, так что $x_1-x_2 \equiv 0 \pmod{m_1m_2}$.

1.4. Если $b \in S$, то $0=b-b \in S$, так что $0-b=-b \in S$. Следовательно, $a+b=a-(-b) \in S$.

1.5.

Таблица сложения A

	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Таблица умножения B

	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

К главе 2:

2.1. Пусть k — целое число такое, что $p^k \leq n < p^{k+1}$. Среди чисел $1, 2, \dots, n$ имеется $\left\lfloor \frac{n}{p} \right\rfloor$ чисел, кратных p . Среди них $\left\lfloor \frac{n}{p^2} \right\rfloor$ чисел кратно p^2 . Среди последних $\left\lfloor \frac{n}{p^3} \right\rfloor$ чисел кратно p^3 . Таким образом, показатель степени числа p равен

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots + \left\lfloor \frac{n}{p^k} \right\rfloor.$$

2.2. Из $\binom{n}{k}$ последовательно вычитаем каждое слагаемое правой части. Используя формулу (2.7), получаем

$$\binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}, \quad \binom{n-1}{k-1} - \binom{n-2}{k-1} = \binom{n-2}{k-2}, \dots, \\ n \binom{n-k+1}{1} - \binom{n-k}{1} = 1.$$

2.3. Очевидно, что $\mu(1)=1$. Если $(n_1, n_2)=1$, то n_1, n_2 не имеют общих простых делителей и, следовательно, выполняется также равенство $\mu(n_1 n_2) = \mu(n_1) \mu(n_2)$.

2.4. Свойство размещения, заключающееся в том, что оно не содержит соответственно букв a, b, c , обозначим через $p(1), p(2), p(3)$. При этом $N=4^r$, $N_1=3^r$, $N_{1,1_2}=2^r$, $N_{123}=1$. Согласно принципу включения и исключения, искомое число равно $4^r - 3 \cdot 3^r + 3 \cdot 2^r - 1$.

2.5. В n столбцах не может быть более чем n шашек. Число способов выбора столбцов, содержащих k шашек, равно $\binom{n}{k}$. Так как в каждом столбце шашку можно поместить как в верхней, так и нижней половинах, то $r_k = 2^k \binom{n}{k}$. Следовательно,

$$R_n = \sum_{k=0}^n 2^k \binom{n}{k} x^k = (1+2x)^n.$$

К главе 3:

$$3.1. u_n = \frac{5}{8} \cdot 3^n + \frac{5}{8} (-1)^n - \frac{1}{4}.$$

$$3.2. \text{ Решая уравнение } u_{n+1} - 2u_n = 4^n, u_1=3, \text{ находим } u_n = \frac{1}{2} \cdot 2^n + \frac{1}{2} 4^n.$$

3.3. Положим $f_{n+1} - f_n = n+1$, $f_1=2$. Так как 1 является простым характеристическим корнем, то $f_n = \frac{1}{2} n^2 + \frac{1}{2} n + 1$.

3.4. Из (3.54) видно, что частное решение в данном случае имеет вид

$$n^{(k)} / (k+r)^{(r)}.$$

Из (3.46) следует, что характеристическим уравнением здесь является $(x-1)^r$. Общим решением при этом оказывается

$$c_0 + c_1 n + \dots + c_{r-1} n^{r-1},$$

где c_0, c_1, \dots, c_{r-1} — неопределенные постоянные.

3.5. Решая уравнение $u_{n+2} - 4u_{n+1} + 4u_n = 0$, $u_1=4$, $u_2=12$, получаем $u_n = (1+n)2^n$.

К главе 4:

4.1. Если $(a-b)$ -путь не существует, то граф не является связанным. Пусть S_a, S_b — связанные компоненты, содержащие соответственно a и b . Вершина a в S_a является единственной имеющей нечетную степень. Это противоречит утверждению упражнения 4.1.

4.2. Предположим, что $P_1 = a_0, x_1, a_1, \dots, x_l, a_l$; $P_2 = a'_0, x'_1, a'_1, \dots, x'_l, a'_l$ общих вершин не имеют. Пусть $V_1 = \{a_0, a_1, \dots, a_l\}$, $V_2 = \{a'_0, a'_1, \dots, a'_l\}$; $P' : a_l, y_1, \dots, y_k, a'_l$ — кратчайший путь между V_1 и V_2 ($k \geq 1$). Если к участку пути P_1 от a_0 до a_l (или от a_l до a_0) присоединить P' и к полученному в результате пути присоединить участок пути P_2 от a'_l до a'_0 (или от a'_l до a'_0), то получится путь, длина которого превышает l . Пришли к противоречию.

4.3.

$$M = \begin{bmatrix} 1 & \dots & \dots & 1 & 1 & 1 & \dots & \dots \\ \vdots & 1 & \dots & \dots & 1 & 1 & \dots & 1 \\ \vdots & \vdots & 1 & \dots & \dots & \dots & \dots & 1 \\ \vdots & \vdots & \vdots & 1 & \dots & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & 1 & \dots & \dots & 1 \end{bmatrix} \quad Q = \begin{bmatrix} 1 & \dots & \dots & 1 & \dots & \dots & \dots \\ \vdots & 1 & \dots & \dots & 1 & \dots & \dots \\ \vdots & \vdots & 1 & \dots & \dots & 1 & \dots \\ \vdots & \vdots & \vdots & 1 & \dots & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 1 & 1 \end{bmatrix}.$$

4.4. Существуют ребра $x \in F - F'$, $x' \in F' - F$. Главный цикл в дереве T , определяемый хордой x' , обязательно включает ребро x . В противном случае существовал бы цикл, состоящий только из ребра F' . Следовательно, с помощью элементарного преобразования дерева путем добавления к T ребра x' и исключения x можно получить T' .

4.5. Предположим, что указанные в условии задачи вершины b и c существуют. Тогда так как вершины b и c , принадлежащие в G одной связной компоненте, в G' принадлежат разным связным компонентам, то степень связности G больше степени связности G' .

Пусть a — разделяющая вершина. Тогда обязательно найдутся две вершины, b и c , которые принадлежат одной связной компоненте G и разным связным компонентам G' . Это означает, что в G все $(b-c)$ -пути проходят через a .

К главе 5:

5.1. Рассмотрим ориентированную цепь максимальной длины $\vec{a}, \vec{x}_1, a_1, \dots, \vec{x}_l, a_l$. Если положительная степень вершины a_l не равна нулю, то $\vec{x}_{l+1} = [a_l, a_{l+1}]$. Если бы вершина a_{l+1} совпадала с одной из вершин a_0, a_1, \dots, a_{l-1} , то это означало бы, что в графе существует замкнутый цикл. Следовательно, a_{l+1} отлична от перечисленных вершин. Но тогда, добавляя \vec{x}_{l+1} к рассматривавшейся цепи, можно получить цепь большей длины. Пришли к противоречию.

5.2. Пусть a_1, \dots, a_n — вершины графа \vec{G} . Для вершин a_1, a_2 найдется вершина b_2 такая, что $b_2 \Rightarrow a_1$, $b_2 \Rightarrow a_2$. Для b_2 и a_2 найдется вершина b_3 такая, что $b_3 \Rightarrow b_2$, $b_3 \Rightarrow a_3$. Следовательно, из b_3 можно достичь a_1, a_2, a_3 . Если эту процедуру продолжить, то через некоторое число шагов мы придем к тому, что b_n является корнем.

5.3.

$$M = \begin{bmatrix} 1 & \vdots & \vdots & \vdots & \vdots & -1 & -1 & -1 & \vdots & \vdots & \vdots \\ \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & -1 & 1 & \vdots \\ \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 1 \\ \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 1 & \vdots \end{bmatrix};$$

$$Q = \begin{bmatrix} -1 & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & -1 & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & -1 & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & -1 & \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & -1 & \vdots & \vdots & \vdots & \vdots & 1 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & -1 & \vdots & \vdots & \vdots & \vdots & 1 \end{bmatrix}.$$

5.4. Допустим, что циклом является путь $b_0, x_1, b_1, \dots, x_l, b_l = b_0$.

Ориентируем этот цикл так, чтобы его вершины проходились в последовательности $b_0, b_1, \dots, b_{l-1}, b_l = b_0$. Пусть e_1, \dots, e_l — величины, принимающие значение $+1$, если введенная ориентация соответствующего ребра в последовательности x_1, \dots, x_l совпадает с исходной ориентацией этого ребра, и -1 в противном случае. Для каждого целого числа $b (1 \leq b \leq \varphi + 1)$ определим вектор $e^{(b)}$, b -я компонента которого равна 1 , если $b \neq \varphi + 1$; остальные компоненты этого вектора всегда равны 0 (если $b = \varphi + 1$, то $e^{(b)}$ — нулевой вектор).

Пусть $a^{(1)}, \dots, a^{(l)}$ — вектор-столбцы A , соответствующие x_1, \dots, x_l и

$$e^{(k)} = e_k (a^{(b_k)} - a^{(b_{k+1})}), k = 1, \dots, l.$$

Следовательно, $e_1 a^{(1)} + \dots + e_l a^{(l)} = 0$.

5.5. Для произвольного целого числа $b (1 \leq b \leq \varphi)$ найдется цепь $b = b_0, x_1, \dots, x_l, b_l = p$, соединяющая вершину b в базисной точке $\varphi + 1 = p$. Ориентируем эту цепь в направлении от b к p . Точно так же, как и при решении предыдущей задачи, введем величины e_1, \dots, e_l , принимающие значения $+1$ и -1 в соответствии с совпадением или несовпадением введенной и исходной ориентации ребер x_1, \dots, x_l соответственно. Имеем

$$e_1 a^{(1)} + \dots + e_l a^{(l)} = e^{(b)} - e^{(p)} = e^{(b)}.$$

Следовательно, в виде линейной комбинации φ вектор-столбцов, соответствующих ребрам дерева T , можно представить все базисные векторы $e^{(1)}, \dots, e^{(\varphi)}$.

К главе 6:

6.1. Из (6.6) имеем $b \cap c \leq b \leq a \cup b$, $b \cap c \leq c \leq a \cup c$, $a \leq a \cup b$, $a \leq a \cup c$. Из (6.7) получаем $a \cup (b \cap c) \leq a \cup b$, $a \cup (b \cap c) \leq a \cup c$. Из (6.8) имеем $a \cup (b \cap c) \leq (a \cup b) \cap (a \cup c)$. Из закона двойственности получается вторая формула.

6.2. В силу дистрибутивности $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$. Отсюда, из неравенства $a \leq c$ и (6.5) получаем $a \cup c = c$.

6.3. Из дистрибутивности следует, что

$$\begin{aligned} (\text{Правая часть}) &= (a \cup b) \cap ((a \cap b) \cup c) = ((a \cup b) \cap (a \cap b)) \cup ((a \cup b) \cap c) = \\ &= (a \cap b) \cup (b \cap c) \cup (c \cap a). \end{aligned}$$

При выводе последнего равенства использовалось то, что $(a \cup b) \cap (a \cap b) = a \cap b$, что, в свою очередь, является следствием неравенства $a \cap b < a < a \cup b$.

6.4. Пусть U — множество из m элементов. Согласно теореме 6.1 и следствию из теоремы 6.1 рассматриваемая булева алгебра изоморфна $p(U)$, а в силу утверждения задачи 7.3 также и V_m . Если представить V_m в виде m -мерного единичного куба и каждое его ребро ориентировать от вершины, которой соответствует m -мерный вектор большего веса в направлении вершины, которой соответствует m -мерный вектор меньшего веса, то на плоскости получится диаграмма Хассе.

К главе 7:

7.1. В данном случае достаточно показать, что высказывание $(P \rightarrow Q) \cdot (Q \rightarrow R)$ истинно, т. е. что из истинности $P \rightarrow Q$ и $Q \rightarrow R$ следует истинность $P \rightarrow R$. Если P ложно, то $P \rightarrow R$ истинно. Если P истинно, то истинно Q . Следовательно, также истинны R и $P \rightarrow R$.

7.2. Из задачи 7.32 следует, что имеется восемь симметричных функций от двух переменных; если исключить константы 0 и 1 , то останется шесть функций.

Согласно упражнению 7.4, этими функциями являются $x_1 x_2$, $\bar{x}_1 \bar{x}_2$, $x_1 \oplus x_2$, $1 \oplus x_1 \oplus x_2$, $x_1 \vee x_2$, $\bar{x}_1 \vee \bar{x}_2$. Так как $\overline{0 \cdot 1} = 1 \neq 0 = \overline{0 \cdot 0} \cdot \bar{1}$, то $\bar{x}_1 \bar{x}_2$ не удовлетворяет закону ассоциативности. Такова же функция $\bar{x}_1 \vee \bar{x}_2$. Оставшиеся четыре функции, как следует из таблиц формул I и III, удовлетворяют закону коммутативности.

7.3. (1) Коэффициенты c_1, c_2, \dots, c_n в определении пороговой функции, как легко проверить, таковы, что

$$f_{X,A} \geq f_{X,\bar{A}}, \text{ если } \sum_{j=1}^r c_j a_j \geq \sum_{j=1}^r c_j (1 - a_j);$$

$$f_{X,A} \leq f_{X,\bar{A}}, \text{ если } \sum_{j=1}^r c_j a_j \leq \sum_{j=1}^r c_j (1 - a_j).$$

(2) Если положить $X = \{1, 3\}$, $A = \{1, 0\}$, то $f_{X,A} = x_2$, $f_{X,\bar{A}} = x_1$. Следовательно, функция f смешанной монотонной не является. Из (1) также следует, что она не является и пороговой.

7.4. Если в определении $D_1(f)$ пороговой функции f положить

$$D_1(g_{2j-1}) = \left\{ (a_1, a_2, \dots, a_n) \left| \sum_{i=1}^n c_i a_i \geq c^{(2j-1)} \right. \right\};$$

$$D_1(g_{2j}) = \left\{ (a_1, a_2, \dots, a_n) \left| \sum_{i=1}^n (-c_i) a_i \geq 1 - c^{(2j)} \right. \right\},$$

то получится искомое представление.

$$\begin{aligned} 7.5. \quad \bar{f}^{(sd)}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 1) &= \bar{f}^d(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = f(x_1, x_2, \dots, x_n) = \\ &= f^{sd}(x_1, x_2, \dots, x_n, 1); \\ \bar{f}^{sd}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, \bar{0}) &= \bar{f}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = f^d(x_1, x_2, \dots, x_n) = \\ &= f^{sd}(x_1, x_2, \dots, x_n, 0). \end{aligned}$$

7.6. Имеем $f(0, x_2, \dots, x_n) = \bar{f}(1, \bar{x}_2, \dots, \bar{x}_n) = \bar{g}(1, \bar{x}_2, \dots, \bar{x}_n) = g(0, x_2, \dots, x_n)$, т. е. $f_{x_1=0} = g_{x_1=1} (a=0, 1)$. Следовательно, $f = g$.

7.7. Обозначим $|D_1(f)|$ через $|f|_n$ и запишем f в виде $f = f_0 \oplus x_1 \cdot f_1$, где f_0, f_1 — многочлены от переменных x_2, \dots, x_n , $\deg f_0 < r$, $\deg f_1 \leq r-1$ (f_0 — сумма всех конъюнкций f , не содержащих x_1 , а f_1 — сумма конъюнкций, включающих x_1). Имеем

$$|f|_n = |f_{x_1=0}|_{n-1} + |f_{x_1=1}|_{n-1} = |f_0|_{n-1} + |f_0 \oplus f_1|_{n-1}.$$

Воспользуемся индукцией по n . При $n=1$ доказываемое неравенство очевидно. Предположим, что оно верно при $n-1$.

(1) Если $f_0 \equiv 0$, имеем $f_1 \neq 0$, $\deg f_1 \leq r-1$. По предположению индукции, $|f|_n = |f_1|_{n-1} \geq 2^{n-r}$.

(2) Если $f_0 \neq 0$ и $f_0 \oplus f_1 \equiv 0$, то $\deg f_0 \leq r-1$. Далее доказательство проводится, как в (1).

(3) Если $f_0 \neq 0$ и $f_0 + f_1 \neq 0$, то, по предположению индукции,

$$|f|_n \geq 2^{n-r-1} + 2^{n-r-1} = 2^{n-r}.$$

И, наконец, если $f \equiv x_1 x_2 \dots x_r$, то $|f|_n = 2^{n-r}$.

7.8. Для элементарной конъюнкции t ранга r из предыдущей задачи имеем $|D_1(t)| = 2^{n-r}$. Если $n > r$, то

$$\sum_{(a_1, a_2, \dots, a_n) \in V_n} t_{x_1=a_1, x_2=a_2, \dots, x_n=a_n} = 0.$$

К главе 8:

8.1. Выход s у FA_1 равен $x_1 \oplus x_2 \oplus x_3$; выход s у FA_2 — $(x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5) = S^{(5)}_{1,3,5}$; выход c у FA_1 — $\text{Maj}_3(x_1, x_2, x_3)$, выход c у FA_2 — $\text{Maj}_3(x_1 \oplus x_2 \oplus x_3, x_4, x_5)$. Выходом s у FA_3 является $\text{Maj}_3(x_1, x_2, x_3) \oplus \text{Maj}_3(x_1 \oplus x_2 \oplus x_3, x_4, x_5)$. Так как последние две функции симметричны соответственно по $\{x_1, x_2, x_3\}$ и $\{x_4, x_5\}$, то путем классификации функций по весам w_1 и w_2 значений их переменных легко показать, что выход s у FA_3 равен $S^{(5)}_{2,5}$. Так как $S^{(5)}_{2,5} = S^{(5)}_{2,5} \cdot S^{(5)}_{1,3,5}$, то выход c у FA_4 равен $S^{(5)}_3$.

8.2. По определению,

$$\sum_{i=0}^6 c_i 2^i = 10 \sum_{i=0}^3 a_i 2^i + \sum_{i=0}^3 b_{0i} 2^i,$$

где $b_{0i} = b_i$. Так как $c_0 = b_{00} = b_0$, то

$$\sum_{i=1}^6 c_i 2^{i-1} = 5 \sum_{i=1}^3 a_i 2^i + \sum_{i=1}^3 b_{0i} 2^{i-1}.$$

Если положить $5a_0 + 4b_{03} + 2b_{02} + b_{01} = 8b_3 + 4b_{12} + 2b_{11} + b_{10}$, то $c_1 = b_{10}$. Следовательно,

$$\sum_{i=2}^6 c_i 2^{i-2} = 5(4a_3 + 2a_2) + 5a_1 + 4b_{13} + 2b_{12} + b_{11}.$$

Если положить $5a_1 + 4b_{13} + 2b_{12} + b_{11} = 8b_{23} + 4b_{22} + 2b_{21} + b_{20}$, то $c_2 = b_{20}$. Повторяя эту процедуру достаточное число раз, можно получить схему, приведенную на рис. 8.15.

8.3. Представим f в виде многочлена F . Положим c_0 равным постоянному члену F . Для каждого члена $x_{j_1} \cdot x_{j_2} \dots x_{j_m}$ степени 1 и выше, имеющего в F коэффициент 1, используем m следующих друг за другом компонент: c_{i+1}, \dots, c_{i+m} ; при этом вход s у $c_{i+1}, \dots, c_{i+m-1}$ полагаем равным 0; а входы x полагаем по порядку равными $x_{j_1}, \dots, x_{j_m-1}$. Пусть вход s у c_{i+m} равен 1, а вход x равен x_{j_m} . Если вход y_1 у c_{i+1} положить равным F_i , а y_2 равным 1, то выходом z_1 у c_{i+m} будет $F_i \oplus x_{j_1} x_{j_2} \dots x_{j_m}$. При этом выход z_2 равен 1. При использовании дизъюнктивной нормальной формы следует взять компоненты z выходными функциями $z_1 = y_1 \vee c x y_2, z_2 = c \vee x y_2$.

8.4. Если пара (x_i, y_i) принимает значение $(0, 1)$ или $(1, 0)$, то $x_{i+1} = x_i, y_{i+1} = y_i$. Если значением (x_i, y_i) является $(1, 1)$, то следует положить $x_{i+1} = y_{i+1} = 1$ при $a_i = b_i$; $x_{i+1} = 1, y_{i+1} = 0$ при $a_i > b_i$ и $x_{i+1} = 0, y_{i+1} = 1$ при $a_i < b_i$. Набор $(x_i, y_i) = (0, 0)$ является избыточным.

8.5. (а) Предположим, что t_i, u_j общих букв не имеют. Если для каждого $2 \leq k \leq n$ конъюнкция t_k включает x_k , то положим $a_k = b_k$. Если ни x_k , ни \bar{x}_k и u_j не входят, то положим $a_k = 1$. По определению,

$$t_i > x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}, u_j \geq \bar{x}_2^{a_2} \bar{x}_3^{a_3} \dots \bar{x}_n^{a_n},$$

т. е. $g(a_2, a_3, \dots, a_n) = g^d(\bar{a}_2, \bar{a}_3, \dots, \bar{a}_n) = \bar{g}(a_2, a_3, \dots, a_n) = 1$. Получили противоречие.

(б) Так как отрицание, Майн и суперпозиция самодвойственных функций — самодвойственные функции, то выходная функция h также самодвойственна. Пусть $f(a_1, a_2, \dots, a_n) = 1$. Если $a_1 = 1$, то существует такая конъюнкция t_i , что при $x_k = a_k$ ($2 \leq k \leq n$) из равенства t_i единице вытекает равенство единице всех букв $z_{i1}, z_{i2}, \dots, z_{in}$. При этом равны 1 все выходы компонент i -й строки схемы, приведенной на рис. 8.18. Так как $\bar{x}_1 = \bar{a}_1$, то выходы $(i+1)$ -й,, m -й строк равны 1 и $h(a_1, a_2, \dots, a_n) = 1$. Если $a_1 = 1$, то вместо строк следует рассматривать столбцы. Следовательно, $f \leq h$. Из задачи 7.42 получаем $f = h$.

8.6. Если $f(a, a, \dots, a) = \bar{a}$, то $f^d(\bar{a}, \bar{a}, \dots, \bar{a}) = a$. Если предположить, что f ни линейна, ни монотонна, ни самодвойственна, то таковой не будет и ее двойственная функция f^d . Отсюда и из теоремы 8.1 следует, что система функций $\{g^d, g^{d_2}, \dots, g^{d_m}\}$ также является полной.

К главе 9:

9.1. Для $B = (b_1, b_2, \dots, b_{n-1}) \in V_{n-1}$ положим $B_0 = (b_1, b_2, \dots, b_{n-1}, 0)$, $B_1 = (b_1, b_2, \dots, b_{n-1}, 1)$ и разобьем V_n на 2^{n-1} пар (B_0, B_1) . Если для каждого $B \in V_{n-1}$ положить:

- 1) $t(B) = 0$ при $f(B_0) = f(B_1) = 0$;
- 2) $t(B) = x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}}$ при $f(B_0) = f(B_1) = 1$;

$$6) t(B) = x_1^{b_1} x_2^{b_2} \dots x_{n-1}^{b_{n-1}} x^n \text{ при } f(B_0) = \bar{a}, f(B_1) = a,$$

то $f = \bigvee_{B \in V_{n-1}} t(B)$. Число членов в этой сумме не превосходит 2^{n-1} . То же самое можно проделать с f^d . Утверждение (а) доказано.

Пусть условия утверждения (б) выполняются. Если при некотором B выполняется условие (1), то минимальное число элементов не превосходит 2^{n-1} . Пришли к противоречию. Если выполняется условие (2), то для f^d при некотором B будет выполняться условие (1), но это противоречит утверждению упражнения 9.1. Следовательно, для всех B выполняется только условие (3).

До настоящего момента рассматривались пары, отличающиеся последними компонентами. Точно так же можно рассмотреть пары, отличающиеся только i -ми компонентами. Таким образом, $f(A) = \bar{f}(A')$ для произвольных наборов $A, A' \in V_n$ таких, что $d(A, A') = 1$. Отсюда и из решения задачи 9.9 следует утверждение (б).

9.2. Согласно утверждению задачи 7.46 функция Майн самодвойственна. Из упражнения 9.1 следует, что достаточно рассматривать только дизъюнктивную нормальную форму минимальной сложности. Так как f монотонна, то, согласно упражнению 9.2, все простые импликанты имеют ранг $m+1$, а их число равно $\binom{n}{m+1} = \binom{n}{m}$. Из утверждения 4 упражнения 9.6 следует, что все простые импликанты являются существенными строками. Отсюда следует решение задачи.

9.3. (а) С одной стороны, $t_1 \geq x_1^{a_1} x_2^{a_2} \dots x_{r-1}^{a_{r-1}} x_{r+1}^{a_{r+1}} \dots x_n^{a_n}$. С другой стороны, $t_j \geq x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ для $2 \leq j \leq m$. Следовательно, утверждение (а) справедливо.

(6) Согласно лемме 9.1 и утверждению задачи 9.11, достаточно рассмотреть частичные суммы следующих дизъюнктивных нормальных норм для $EO = E\bar{x}_1\bar{x}_2 \dots \bar{x}_8, y_1, y_2, y_3$ (для простоты E опускается):

$$\begin{aligned}y_1 &= E(x_6 \vee x_7 \vee x_8 \vee \bar{x}_7 \vee x_6 \vee x_6); \\y_2 &= E(x_6 \vee x_7 \vee x_3 \bar{x}_5 \bar{x}_6 \bar{x}_7 \vee x_3 \bar{x}_5 \bar{x}_6 \vee x_3 \bar{x}_5 \bar{x}_6); \\y_3 &= E(x_6 \vee x_6 \bar{x}_7 \vee x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \vee x_4 \bar{x}_5 \bar{x}_7 \vee x_2 \bar{x}_3 \bar{x}_5 \bar{x}_7).\end{aligned}$$

Согласно утверждению (а), из y_1 нельзя исключить x_8, x_7, x_5 , из y_2 — $x_8, x_7, x_3 \bar{x}_5 \bar{x}_6$, из y_3 — $x_8, x_6 \bar{x}_7 x_2 \bar{x}_3 \bar{x}_5 \bar{x}_7$. Так как $x_6 = x_6 \bar{x}_7 \not\subset x_7$, то, согласно задаче 9.4, в y_1 не нужно x_6 . В y_2 должно быть оставлено одно из произведений $x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7$ или $x_4 \bar{x}_5 \bar{x}_6$, в y_3 — $x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7$ или $x_4 \bar{x}_5 \bar{x}_7$ (если одно из этих произведений исключается, то другое, как следует из утверждения (а), исключено быть не может). Если оставить $x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7$, то $x_4 \bar{x}_5 \bar{x}_6 = x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \not\subset x_7$, $x_4 \bar{x}_5 \bar{x}_7 = x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7 \not\subset x_6 \bar{x}_7$. При этом из y_2 исключается $x_4 \bar{x}_5 \bar{x}_6$, а из y_3 — $x_4 \bar{x}_5 \bar{x}_7$. С другой стороны,

$$GS = E(x_1 \vee x_2 \vee x_3 \vee x_4) \vee y_1 \vee y_2 \vee y_3.$$

При $GS=1$ равенство $y_1 y_2 y_3 = 0$ имеет место только в том случае, если $x_1 \bar{x}_2 \bar{x}_3 \dots \bar{x}_8 = 1$. Единственной p -простой импликантной функции GS , включающей $x_1 \bar{x}_2 \bar{x}_3, \dots, \bar{x}_8$, является x_1 . Как следует из изложенного, достаточно иметь $E x_1, E x_5, E x_7, E x_8, E x_6 \bar{x}_7, E x_3 \bar{x}_5 \bar{x}_6, E x_2 \bar{x}_3 \bar{x}_5 \bar{x}_7, E x_4 \bar{x}_5 \bar{x}_6 \bar{x}_7, E \bar{x}_1 \bar{x}_2 \dots \bar{x}_8$. Они дают единственное решение с минимальным числом, так что $k \geq 9$.

9.4. Опуская индексы, обозначим входные переменные через x, y, a, b , а выходные переменные через x', y' . Пусть f, g — выходные функции, дающие соответственно $x' = y' = 0$ на избыточном входном наборе $x = y = 0$. Функции $f, g, f^{(1)}, g^{(1)}$ (на избыточных входных наборах значения функций равны 1; на остальных входных наборах эти функции совпадают с f и g), как следует из решения задачи 9.4, имеют вид

$$\begin{aligned}f &= x \bar{y} \vee xy (ab \vee ab \vee \bar{a} \bar{b}), f^{(1)} = f \vee \bar{x} \bar{y}; \\g &= \bar{x} y \vee xy (ab \vee \bar{a} b \vee \bar{a} \bar{b}), g^{(1)} = g \vee \bar{x} \bar{y}; \\f^{(1)} g^{(1)} &= xy (ab \vee \bar{a} \bar{b}) \vee \bar{x} \bar{y}.\end{aligned}$$

Пользуясь группой формул I из табл. 7.1 и утверждением (1) задачи 7.17, получаем $f^{(1)} = \bar{y} \vee xa \vee xb$, $g^{(1)} = \bar{x} \vee yb \vee ya$. Согласно теореме 9.1, p -простыми импликантами функции $f^{(1)}$ являются \bar{y}, xa и xb , p -простыми импликантами функции $g^{(1)}$ — \bar{x}, yb и ya , p -простыми импликантами функции $f^{(1)} g^{(1)}$ — $xyab, xy\bar{a}\bar{b}$ (конъюнкция $\bar{x}\bar{y}$ исключается, так как она равна 1 только на избыточном наборе). Точно так же, как и в пункте (а) предыдущей задачи, можно убедиться, что \bar{y} и \bar{x} исключить нельзя. Одновременно нельзя исключить $xa, xb(yb, ya)$. Если исключить одну из этих конъюнкций, то необходимо восполнить либо $xyab$, либо $xy\bar{a}\bar{b}$. Следовательно, число конъюнкций не меньше чем 5. Так как $xb = (xa \not\subset xy\bar{a}\bar{b}) \not\subset \bar{y}$, $y\bar{a} = (yb \not\subset xy\bar{a}\bar{b}) \not\subset \bar{x}$, то, пользуясь утверждением задачи 9.4, легко получить, что $f \leq \bar{y} \vee xa \vee xy\bar{a}\bar{b}$, $g \leq \bar{x} \vee yb \vee xy\bar{a}\bar{b}$ дают минимальное и по числу конъюнкций и по числу входов решение.

К главе 10:

10.1. Предположим, что существует конечный автомат $M = (S, I, O, \sigma, \delta)$, удовлетворяющий условиям теоремы, и что в момент времени $t=1$ он находится в состоянии s_0 . Для произвольного $\alpha \in I^+$ положим $\delta(s_0, \alpha) = f(\alpha)$.

В доказательстве утверждения упражнения 10.2 используются j', j , соответствующие $a=0$. А именно, существуют j' и j такие, что $\sigma(s_0, 0^{j'}) = \sigma(s_1, 0^j) = s_1$, $0 \leq j' < j$. Из задачи 10.1 следует, что

$$\delta(s_0, 0^{j'} 1^j) = \delta(\sigma(s_0, 0^{j'}), 1^j) = \delta(s_1, 1^j);$$

$$\delta(s_0, 0^j 1^j) = \delta(\sigma(s_0, 0^j), 1^j) = \delta(s_1, 1^j).$$

Из предположений и определения f получаем

$$\delta(s_0, 0^{j'} 1^j) = f(0^{j'} 1^j) = 0; \quad \delta(s_0, 0^j 1^j) = f(0^j 1^j) = 1.$$

Получили противоречие.

10.2. Предположим, что существует конечный автомат $M = (S, I, O, \sigma, \delta)$, удовлетворяющий условиям задачи, и что при некотором $s_0 \in S$

$$\delta(s_0, rc^n) = f(rc^n), \quad n \geq 1. \text{ Если положить } \sigma(s_0, r) = s_1, \text{ то } \delta(s_0, rc^n) = \delta(s_1, c^n) = f(rc^n).$$

Имеем $f(rc^m) = 1$, $m \geq 1$. В других случаях $f(rc^n) = 0$. Так как $(m+1)^2 - m^2 = 2m+1$, то при увеличении m интервал увеличивается. Однако это противоречит утверждению упражнения 10.2.

10.3. Если определить M' так, что $S' = S \times O$, $\sigma'((s, b), a) = (\sigma(s, a), \delta(s, a))$; $\delta'((s, b), a) = b$, то получится автомат Мура. Выбирая $b_0 \in O$ и полагая $\Phi(s) = (s, b_0)$, рассматриваемое соотношение можно доказать индукцией по $|a|$. Если хранить в памяти выходные значения в качестве компонент состояния, то число состояний увеличится. Выход автомата M в момент времени $t=i$ представляет собой выход автомата M' в момент времени $t=i+1$.

10.4. Без какой-либо потери общности M можно считать простейшим. Далее следует показать, что $\alpha_1 \approx_{f_s} \alpha_2 \Leftrightarrow \sigma(s, \alpha_1) = \sigma(s, \alpha_2)$. Заметьте, что число классов эквивалентности отношения \approx_{f_s} не превосходит $|S|$.

10.5. Так как число состояний в данном случае равно двум, то достаточно использовать один триггер. Пусть q — переменная состояния; состоянию t_1 сопоставим $q=0$, а состоянию t_2 — $q=1$. С помощью табл. 10.1 и 10.9 находим таблицу значений S и R . Выходной функцией является $s = x \oplus y \oplus q$. При таком распределении состояний дизъюнктивной нормальной формой минимальной сложности будет $s = xyq \vee \bar{x} \bar{y} q \vee \bar{x} y \bar{q} \vee x \bar{y} \bar{q}$, $S = xy$, $R = \bar{x} \bar{y}$. Если состояниям t_1 и t_2 сопоставить соответственно $q=1$ и $q=0$, то q и \bar{q} поменяются местами. Из табл. 10.9 следует, что надо поменять местами S и R .

10.6. Одним из решений задачи является использование НЕ—ИЛИ-схемы, двойственной для схемы из задачи 9.28. Если при этом положить $S = \text{Maj}_3(x, y, g)$ (получается, если все d положить равными 1) и $R = \bar{x} \bar{y}$ (выход g^1), то рассматриваемые функции можно реализовать восемью элементами.

10.7. Пусть q_i — внутреннее состояние триггера F_i (при определении входе Q). При подаче на тактовый вход импульса со входа происходит переход $q_i \rightarrow q_{i+1}$ (индексы вычисляются по модулю 4).

10.9. Так как для различных целых положительных чисел i, j выполняются равенства $f([i]^1) = 0$, $f([i]^j) = 0$, $f([j]^1) = 0$, $f([j]^j) = 1$, то последовательности $[i]^1, [j]^j$ не являются f -эквивалентными. Следовательно, f не является регулярной и, согласно упражнению 10.4, не существует конечного автомата, который давал бы то же отображение входных последовательностей в выходные, что и упомянутый в задаче преобразователь.

Список литературы

1. Birkhoff G., Bartee T. C. Modern Applied Algebra. — New York: McGraw-Hill, 1970.
2. Birkhoff G., MacZahe S. A survey of Modern Algebra — New York, 3d ed., MacMillan Company, 1965.
3. Ли. Введение в комбинаторику. I. — Томотати, 1972.
4. Холл М. Комбинаторика. Пер. с англ./Под ред. А. О. Гельфонда и В. Е. Тарканова. — М.: Мир, 1970.
5. Такахаси Т. Разностные уравнения. Байфукан, 1961.
6. Ли. Введение в комбинаторику. II. — Томотати, 1972.
7. Харари. Теория графов.
8. Исни Д. Теория схем. — Корона, 1977.
9. Одзак Х., Сиракава И., Онага К. Теория графов. — Корона, 1975.
10. Янага, Кодaira. Введение в современную математику. — Иванами, 1961.
11. Одзак, Киносита, Сиракава. Теория информационных систем, I. — Корона, 1970.
12. Одзак, Киносита. Дискретная алгебра. — Томотати, 1966.
13. Нодзак С. Теория коммутационных схем. — Томотати, 1972.
14. Muroga S. Logical Design and Switching Theory. CS-391, Univ. of Illinois, 1976.
15. Курихара, Накамура. Математическая логика, I. — Томотати, 1975.
16. Мурога, Ибараки, Китабаси. Пороговая логика. — Сангэ, 1976.
17. McCluskey E. J. Introduction to the Theory of Switching Circuits. — McGraw-Hill, 1965.
18. Хагивара Х. Проектирование ЭВМ. I. Логические схемы. — Асакура, 1969.
19. Фиста М. Логическое проектирование цифровых вычислительных машин. — Асакура, 1960.
20. Gschwind H. W. and McCluskey E. J.: Design of Digital Computers — Springer Verlag, 1975.
21. Lee S. C. Digital Circuits and Logic Design. — Prentice — Hall, 1976.
22. Brzozowski J. A. and Yoeli M. Digital Networks. — Prentice Hall, 1976.
23. Deem W., Muchow K. and Zeppa A.: Digital Computers, Circuits and Concepts. — Reston Publ. Co., 1977.
24. Kostopoulos G. K. Digital Engineering. — John Wiley, 1975.
25. Aho A. V., Hopcroft J. E. and Ullman J. D. The Design and Analysis of Computer Algorithms. — Addison — Wesley, 1974.
26. The Fairchild Semiconductor TTL Data Book, Fairchild — Semiconductor, 1972.
27. The TTL Data Book for Design Engineers, Second Edition. — Texas Instruments, 1976.
28. Ибуки, Намура, Нодзак. Общая теория логических отношений. — Дэнсисун гаккай ромбунси, 1963, с. 46, № 7, с. 55.
29. Касами, Токура, Ивадари, Инагаки. Теория кодирования: Пер. с япон./Под ред. Б. С. Цыбакова и С. И. Гельфанда. — М.: Мир, 1978.
30. Миякава, Ивадари, Имаи. Теория кодирования. — Сёкодо, 1973.
31. Фу Т. С. Целочисленные вычисления и сети. — Байфукан, 1975.
32. Even S. and Kariv O. An $O(n^{2.5})$ Algorithm for Maximum Matching in General Graphs. — Proc. of 16th Ann. Sympos. on Foundations of Computer Science, p. 100—112, 1975.
33. Maley G. A. and Earle J. The Logic Design of Transistor Digital Computers. — Prentice-Hall Inc., 1967.
34. Gimpel J. E. The Minimization of TANT Networks. — IEEE Trans. Electronic Computers, Vol. EC-16, № 1, p. 18—38 (Feb. 1967).
35. Хонда Х. Теория автоматов и языков. — Корона, 1972.
36. Киносита К. Введение в теорию автоматов. — Асакура, 1973.
37. Тома Н. Теория последовательных схем. — Сёкодо, 1976.

Оглавление

	Стр.
Предисловие к русскому изданию	3
Предисловие к японскому изданию	4
Глава 1. Свойства целых чисел	6
1.1. Упорядоченность	6
1.2. Наибольший общий делитель	9
1.3. Разложение на простые сомножители	12
1.4. Сравнимость целых чисел	14
1.5. Области целостности и поля	17
Задачи	21
Глава 2. Размещения, сочетания, принцип включения — исключения	21
2.1. Размещения без повторов	21
2.2. Размещения и сочетания с повторениями	25
2.3. Формула обращения Мёбиуса	29
2.4. Принцип включения — исключения	32
2.5. Размещения с запрещенными позициями	35
Задачи	39
Глава 3. Рекуррентные уравнения	40
3.1. Производящие функции	40
3.2. Решение однородного линейного рекуррентного уравнения	42
3.3. Метод решения неоднородного линейного рекуррентного уравнения	45
3.4. Разностные формулы	47
3.5. Нахождение частного решения	50
3.6. Приложения к задачам, связанным с периодическими структурами	53
Задачи	55
Глава 4. Графы	55
4.1. Понятие графа	55
4.2. Связность	60
4.3. Деревья	63
4.4. Коцкклы	67
4.5. Векторные пространства, связанные с графами	70
4.6. Двудольные графы	74
Задачи	76
Глава 5. Ориентированные графы	77
5.1. Ориентированные пути	77
5.2. Сильная связность	80
5.3. Ориентированные деревья	82
5.4. Матрицы инцидентности	85
5.5. Закон Кирхгофа	88
Задачи	92
Глава 6. Булевы алгебры	92
6.1. Определение булевой алгебры	92
6.2. Отношения порядка в булевой алгебре	96
6.3. Булевы кольца	103
6.4. Представления булевых алгебр	106
Задачи	111
Глава 7. Булевы функции	112
7.1. Определение	112
7.2. Операции	117
7.3. Методы задания булевых функций. Базисные функции	122
7.4. Разложение булевых функций	129
Задачи	135
Глава 8. Применение теории булевых функций	136
8.1. Схемы из функциональных элементов	136
8.2. Суперпозиции и полные системы функций	147
8.3. Булевы формулы	155
8.4. Коды с обнаружением и исправлением ошибок	161
Задачи	167
Глава 9. Методы минимизации	169
9.1. Задачи минимизации булевых формул	169
9.2. Метод Квайна-Мак-Класски	174
9.2.1. Метод, основанный на согласователях	177
9.2.2. Таблица Карно и коды с единичным расстоянием	184
9.3. Задача о покрытии множеств	187
9.3.1. Постановка задачи	187
9.3.2. Упрощение матриц инцидентности (таблиц простых импликант)	190
9.4. О НЕ-И и НЕ-ИЛИ-схемах	197
Задачи	203
Глава 10. Конечные автоматы	204
10.1. Определение конечного автомата	204
10.2. Сжатие конечных автоматов	211
10.3. Синхронные конечные автоматы	219
Задачи	229
Решения задач	230
Список литературы	239

