А.А. НАБЕБИН

Ученым, естественникам и гуманитариям, на своих плечах поднимающих небо и раздвигающих горизонты знания, посвящаю

ДИСКРЕТНАЯ МАТЕМАТИКА

Допущено Учебно-методическим объединением вузов Российской Федерации по классическому университетскому образованию в качестве учебника для студентов высших учебных заведений, обучающихся по специальности "Прикладная математика и информатика", а также специальности "Информационные системы и технологии"

Москва
Научный мир
2010

УДК 519.1 + 519.723(075.8)
ББК 22.176.73 Н 134
Н13

Н13 **Набебин А.А.**
Дискретная математика. – М.: Научный мир, 2010. – 512 с.: ил.
ISBN 978-5-91522-190-0

Излагаются основные понятия дискретной математики: модулярная арифметика и ее использование в криптографии, элементы комбинаторики, алгебра логики и логика предикатов, теория графов, конечные автоматы. Предназначено студентам высших технических учебных заведений, специализирующимся в области прикладной математики, вычислительной техники, программирования, информатики.

РЕЦЕНЗЕНТЫ:
кафедра математической кибернетики
факультета Вычислительной математики и кибернетики
Московского государственного университета имени М.В.Ломоносова;
доктор физ.-мат. наук, профессор ***Алексеев Валерий Борисович***

НАУЧНЫЙ РЕДАКТОР:
Канд. физ.-мат. наук, доцент ***Захаров Владимир Анатольевич***

УДК 519.1 + 519.723(075.8)
ББК 22.176.73 Н 134

ISBN 978-5-91522-190-0

ВВЕДЕНИЕ

1. Множество

Понятие множества неопределимо. Это простейшее исходное понятие человечество сформировало из опыта всего своего исторического развития. То же можно сказать о смысле простейшего отношения принадлежности: элемент a принадлежит множеству A (обозначение $a \in A$) и о смысле отношения тождества (совпадения, равенства) двух элементов a и b из некоторого множества (обозначение $a = b$). Другими словами, предполагается, что читатель умеет распознавать совпадение или несовпадение двух элементов и устанавливать факт принадлежности или непринадлежности элемента множеству.

Пусть A,B,C – произвольные множества; a,b,c – элементы множеств. Итак, основными неопределяемыми отношениями в теории множеств являются следующие отношения:

$a = b$, элементы a и b равны (совпадают);

$a \in A$, элемент a принадлежит множеству A.

Пусть знак $\leftrightarrow$ означает тогда и только тогда; а знаки &, $\vee$, $\neg$, $\rightarrow$, $\exists$, $\forall$ есть логические знаки конъюнкции, дизъюнкции, отрицания, импликации, квантора общности и квантора существования. Будем использовать их в общепринятом содержательном смысле. Эти же логические символы в последующем применим при построении формул математической логики.

Введем далее следующие отношения:

$A \subseteq B \leftrightarrow \forall a\ (a \in A \rightarrow a \in B)$;

$A = B \leftrightarrow A \subseteq B \ \& \ B \subseteq A$;

$A \subset B \leftrightarrow A \subseteq B \ \& \ A \neq B$;

$A \supseteq B \leftrightarrow B \subseteq A$;

$A \supset B \leftrightarrow B \subset A$.

Обозначим через $\mathcal{P}(A)$ множество всех подмножеств множества A и пусть $\varnothing$ есть символ пустого множества. Введем операции над множествами:

$A \cup B = \{x : x \in A \vee x \in B\}$ – объединение множеств A и B;

$A \cap B = \{x : x \in A \ \& \ x \in B\}$ – пересечение множеств A и B;

$A - B = \{x : x \in A \ \& \ x \notin B\}$ – разность множеств A и B;

$A \dot{-} B = (A-B) \cup (B-A)$ – симметрическая разность множеств A и B;

$A \times B = \{(a,b) : a \in A \ \& \ b \in B\}$ – декартово произведение множеств A и B.

Примем следующие обозначения.

Декартово произведение можно распространить на несколько сомножителей, именно, $A_1 \times A_2 \times \ldots \times A_n = \{(a_1,a_2,\ldots,a_n): a_1 \in A_1,\ a_2 \in A_2,\ \ldots,\ a_n \in A_n\}$. Отсюда определим натуральную степень $A^n = A \times A \times \ldots \times A$ (n раз).

Иногда вместо $A \cap B$ пишут $A \cdot B$ или AB.

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n, \quad \bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n,$$

$$\bigtimes_{i=1}^{n} A_i = A_1 \times A_2 \times \ldots \times A_n.$$

Множества $\varnothing$ и A называются несобственными (тривиальными) подмножествами множества A. Если $A \subset B \ \&\ A \neq \varnothing$, то A есть собственное подмножество множества B.

Множество натуральных чисел $\mathbb{N} = \{0,1,2,\ldots\}$.

Множество положительных натуральных чисел $\mathbb{N}_+ = \{1,2,\ldots\}$.

Множество целых чисел $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$.

Множество $\mathbb{Z}_k = E_k = \{0,1,2,\ldots,k-1\}$.

Множество рациональных чисел $\mathbb{Q} = \{m/n: m,n \in \mathbb{Z},\ n \neq 0\}$..

Множество вещественных чисел $\mathbb{R} = (-\infty,+\infty)$.

Множество комплексных чисел $\mathbb{C} = \{x+iy: x,y \in \mathbb{R},\ i=\sqrt{-1}\}$.

Пусть A,B,C – произвольные подмножества некоторого множества U (универсума). Иногда U обозначают через 1. Пусть $\bar{A} = U-A$. Иногда $\bar{A}$ обозначают через $\neg A$. Тогда справедливы следующие (булевы) свойства операций над множествами.

1. Идемпотентность
   $A \cap A = A, \quad A \cup A = A.$
2. Коммутативность
   $A \cap B = B \cap A, \ A \cup B = B \cup A.$
3. Ассоциативность
   $A \cap (B \cap C) = (A \cap B) \cap C,$
   $A \cup (B \cup C) = (A \cup B) \cup C.$
4. Правило поглощения
   $A \cap (A \cup B) = A, \quad A \cup (A \cap B) = A.$
5. Дистрибутивность
   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$
   $A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$

6. Инволюция $\bar{\bar{A}} = A$.
7. Свойства констант
   $A \cap U = A, \quad A \cup \varnothing = A,$
   $A \cap \varnothing = \varnothing, \quad A \cup U = U.$
8. Закон исключенного третьего и закон противоречия
   $A \cup \bar{A} = U, \quad A \cap \bar{A} = \varnothing.$
9. Правила де Моргана
   $\overline{A \cap B} = \bar{A} \cup \bar{B}, \quad \overline{A \cup B} = \bar{A} \cap \bar{B}.$

Замечание. Ассоциативность позволяет записывать объединение и пересечение множеств без скобок. Они расставляются произвольным образом).

Определение. Пусть A и B – два множества. Определим функцию $f: A \to B$ как отображение, которое каждому элементу a из A ставит в соответствие некоторый элемент b из B. Это обстоятельство записывается как $b = f(a)$ или $a \vdash f(a)$. Примем, что $D(f)$ есть область определения функции f; $R(f)$ – область значений функции f; $f(A)$ – область тех значений функции f, когда аргумент функции f пробегает множество A.

Замечание. В этом определении функция f всюду определена. Частично определенная функция $f: A \to B$ есть отображение, которое каждому элементу из множества A сопоставляет не более одного элемента из множества B.

Определение. Характеристическая и представляющая функции множества A определяются соответственно как

$$\varphi(x) = \begin{cases} 1, & \text{если } x \in A, \\ 0, & \text{если } x \notin A. \end{cases} \quad \psi(x) = \begin{cases} 0, & \text{если } x \in A, \\ 1, & \text{если } x \notin A. \end{cases}$$

Определение. График функции $f: A \to B$ есть множество $G_f = \{(a,b) \in A \times B: f(a)=b\}$.

2. Мощность множества. Счетные и несчетные множества

Определение. Функция $\varphi: A \to B$ есть взаимно однозначное соответствие между множествами A и B, если

1) $\forall b \in B\ \exists a \in A\ \varphi(a)=b$,
2) $\forall a_1,a_2 \in A\ (a_1 \neq a_2 \to \varphi(a_1) \neq \varphi(a_2))$.

Замечание. Последнее можно записать как

2′) $\forall a_1,a_2 \in A\ (\varphi(a_1)=\varphi(a_2) \to a_1=a_2)$.

Определение. *Множества A и B эквивалентны* ($A \sim B$), если между их элементами можно установить взаимно однозначное со-

ответствие.

Отношение эквивалентности множеств обладает следующими свойствами.

1. $A \sim A$, рефлексивность.
2. $A \sim B \rightarrow B \sim A$, коммутативность.
3. $A \sim B \ \& \ B \sim C \rightarrow A \sim C$, транзитивность.

Определение. *Мощность множества* A (обозначение $|A|$) есть класс эквивалентных ему множеств. Мощность конечного множества есть число его элементов.

Замечание. Эквивалентные множества A и B равномощны, то есть $A \sim B \leftrightarrow |A| = |B|$.

Определение. Множество A *счетно*, если A эквивалентно множеству $\mathbb{N}$ натуральных чисел. В противном случае множество A *несчетно*.

Замечание. Множество A счетно, если A можно "пересчитать" натуральными числами.

Утверждение. Из всякого бесконечного множества можно выделить счетное подмножество.

Доказательство. Пусть A бесконечное множество. Выделим в A произвольный элемент a_0. Множество $A-\{a_0\}$ бесконечно. Выделим в нем элемент a_1. Множество $A-\{a_0,a_1\}$ бесконечно. Выделим в нем элемент a_2. И так далее. В бесконечном множестве A выделено счетное подмножество $B = \{a_0,a_1,a_2,\ldots\}$.

Утверждение. Множество $\mathbb{Q}_+$ положительных рациональных чисел счетно.

Доказательство. Расположим элементы множества $\mathbb{Q}_+$ в следующей таблице.

1, 1/2, 1/3, 1/4, 1/5, ...
2, 2/2, 2/3, 2/4, 2/5, ...
3, 3/2, 3/3, 3/4, 3/5, ...
4, 4/2, 4/3, 4/4, 4/5, ...
...

Выписываем элементы из $\mathbb{Q}_+$ по диагонали, сверху вниз, выпуская ранее встречавшиеся числа: 1,1/2,2,1/3,3,2/3,... Следовательно, множество $\mathbb{Q}_+$ счетно.

Утверждение. Объединение конечного или счетного множества счетных множеств счетно.

Доказательство. Расположим элементы множеств $A_1,A_2,A_3,\ldots$ (их число может быть и конечным) в следующей таблице.

A_1: $a_{11},a_{12},a_{13},a_{14},\ldots$

A_2: $a_{21},a_{22},a_{23},a_{24},\ldots$
A_3: $a_{31},a_{32},a_{33},a_{34},\ldots$
A_4: $a_{41},a_{42},a_{43},a_{44},\ldots$
...

Выписываем элементы из $A_1 \cup A_2 \cup A_3 \cup \ldots$ по диагонали, сверху вниз, выпуская ранее встречавшиеся элементы: a_{11},a_{12},a_{21}, $a_{13},a_{23},a_{31},\ldots$ Следовательно, множество $A_1 \cup A_2 \cup \ldots$ счетно.

Замечание. Объединение конечного множества и счетного множества счетно. Множество рациональных чисел счетно, ибо $\mathbb{Q} = \mathbb{Q}_- \cup \mathbb{Q}_+ \cup \{0\}$, где $\mathbb{Q}_-$ есть множество отрицательных рациональных чисел.

3. Мощность континуума

Утверждение. Множество C всех бесконечных последовательностей из 0 и 1 несчетно.

Доказательство. Допустим противное: существует пересчет всех бесконечных последовательностей $A_1,A_2,A_3,\ldots$ из 0 и 1:

A_1: $a_{11},a_{12},a_{13},a_{14},\ldots$
A_2: $a_{21},a_{22},a_{23},a_{24},\ldots$
A_3: $a_{31},a_{32},a_{33},a_{34},\ldots$
A_4: $a_{41},a_{42},a_{43},a_{44},\ldots$
...

Построим последовательность B: $b_1,b_2,b_3,\ldots$, где

$$b_i = \begin{cases} 1, & \text{если } a_{ii}=0, \\ 0, & \text{если } a_{ii}=1, \end{cases} \quad i=1,2,3,\ldots$$

Последовательность B лежит вне указанного пересчета. B отличается от A_1 элементом $b_1 \neq a_{11}$, от A_2 элементом $b_2 \neq a_{22}$, от A_3 элементом $b_3 \neq a_{33}$, и так далее. Следовательно, исходное множество C несчетно.

Определение. Множество A имеет *мощность континуума* c, если A эквивалентно множеству всех бесконечных последовательностей из 0 и 1.

Следствие. Множество C всех бесконечных последовательностей из 0 и 1 имеет мощность континуума: $|C| = c$.

Утверждение. Множество $\mathcal{P}(\mathbb{N})$ всех подмножеств множества натуральных чисел имеет мощность континуума.

Доказательство. Всякую бесконечную последовательность из 0 и 1 можно рассматривать как характеристическую функцию некоторого подмножества множества натуральных чисел. Следова-

тельно, множество $\mathcal{P}(\mathbb{N})$ имеет мощность континуума: $|\mathcal{P}(\mathbb{N}| = c$.

Следствие. Множество всех подмножеств множества натуральных чисел несчетно.

Утверждение. Если к бесконечному множеству добавить конечное или счетное множество элементов, то его мощность не изменится.

Доказательство. Пусть A – бесконечное множество, а B – конечное или счетное множество, причем $A \cap B = \varnothing$. Покажем, что $A \sim A \cup B$. Выделим из множества A счетное подмножество A_1. Тогда $A = C \cup A_1$, где $C = A - A_1$, и $A \cup B = (C \cup A_1) \cup B = C \cup (A_1 \cup B)$. Так как $A_1 \cup B \sim A_1$, то $A \cup B = C \cup (A_1 \cup B) \sim C \cup A_1 = A$, то есть $A \cup B \sim A$.

Утверждение. Если A – несчетное множество, а B – конечное или счетное его подмножество, то $A - B \sim A$.

Доказательство. Пусть $C = A - B$. Тогда $A = C \cup B$. Множество C несчетно, ибо в противном случае C конечно или счетно и тогда $A = C \cup B$ конечно или счетно. Множество $C \cup B \sim C$, или $A \sim C$, то есть $A \sim A - B$.

Теорема. Множество $U = [0,1)$ имеет мощность континуума c.

Доказательство. Множество U эквивалентно множеству всех последовательностей из 0 и 1.

Замечание. 1. $|[0,1]| = |(0,1)| = |(0,1]| = c$.

2. Если $a < b$, то $|[a,b]| = c$, ибо функция $y = a + x(b-a)$ отображает $[0,1]$ на $[a,b]$ взаимно однозначно.

3. $|[a,b]| = |(a,b)| = |(a,b]| = c$.

4. $|(-\infty,\infty)| = |\mathbb{R}| = c$, ибо функция $y = \operatorname{tg} x$ отображает интервал $(a,b) = (-\pi/2, \pi/2)$ на всю числовую ось $\mathbb{R}$ взаимно однозначно.

4. Кардинальные числа. Сравнение мощностей

Определение. *Мощность* множества есть класс эквивалентных между собой множеств. *Кардинальное число* или *кардинал* есть знак (символ), приписываемый мощности как классу эквивалентных между собой множеств. Мощности конечных множеств называются *финитными кардиналами*. Мощности бесконечных множеств называются *трансфинитными кардиналами*.

Пример. Счетной мощности (мощность множества натуральных чисел) присваивается кардинальное число $\aleph_0$ (алеф-нуль). Мощности множества вещественных чисел присваивается кардинальное число c (готическая буква c).

Замечание. Мощность множества A обозначается через $|A|$ или $\operatorname{card}(A)$.

Мощность конечного множества есть число его элементов.

Пусть A, B – произвольные множества и $|A|, |B|$ есть их мощности.

A'priori возможны четыре случая.

1. Множество A эквивалентно некоторому подмножеству множества B, а множество B эквивалентно некоторому подмножеству множества A.

2. Множество A эквивалентно некоторому подмножеству множества B, а множество B не эквивалентно никакому подмножеству множества A.

3. Множество B эквивалентно некоторому подмножеству множества A, а множество A не эквивалентно никакому подмножеству множества B.

4. Множество A не эквивалентно никакому подмножеству множества B, а множество B не эквивалентно никакому подмножеству множества A.

Определение.

$|A| = |B| \longleftrightarrow A \sim B$.

$|A| \leqslant |B| \longleftrightarrow A$ эквивалентно некоторому подмножеству в B.

$|A| < |B| \longleftrightarrow A$ эквивалентно некоторому подмножеству в B, а множество B не эквивалентно никакому подмножеству множества A.

$|A| \geqslant |B| \longleftrightarrow |B| \leqslant |A|$.

$|A| > |B| \longleftrightarrow |B| < |A|$.

Замечание. Случай, когда множество A не эквивалентно никакому подмножеству множества B, а множество B не эквивалентно никакому подмножеству множества A, невозможен.

Теорема (Кантор–Бернштейн). Если множество A эквивалентно некоторому подмножеству B_1 множества B, а множество B эквивалентно некоторому подмножеству A_1 множества A, то множества A и B эквивалентны (то есть имеют равную мощность).

Коротко: $|A| \leqslant |B| \ \&\ |B| \leqslant |A| \ \rightarrow\ |A| = |B|$.

Доказательство. Случаи $B_1 = B$ и $A_1 = A$ можно исключить, ибо если $B_1 = B$, то условие теоремы утверждает, что $A \sim B$, из чего, естественно, следует $A \sim B$. Случай $A_1 = A$ аналогичен.

Итак, пусть $A_1 \subset A$, $B_1 \subset B$. Пусть функции

$\varphi : A \rightarrow B_1$, $\psi : B \rightarrow A_1$

устанавливают взаимно однозначное соответствие между A и B_1 и между B и A_1, то есть $A \underset{\varphi}{\longleftrightarrow} B_1$, $B \underset{\psi}{\longleftrightarrow} A_1$.

С помощью функций φ и ψ расслоим множества A и B на "кольца" следующим образом. Имеем:

$$A \underset{\varphi}{\longleftrightarrow} B_1 \subset B, \quad B \underset{\psi}{\longleftrightarrow} A_1 \subset A,$$
$$A_1 \underset{\varphi}{\longleftrightarrow} B_2 \subset B_1, \quad B_1 \underset{\psi}{\longleftrightarrow} A_2 \subset A_1,$$
$$A_2 \underset{\varphi}{\longleftrightarrow} B_3 \subset B_2, \quad B_2 \underset{\psi}{\longleftrightarrow} A_3 \subset A_2,$$
...

Сформируем множества ("кольца"):

$$K_0^A = A - A_1, \quad K_0^B = B - B_1,$$
$$K_1^A = A_1 - A_2, \quad K_1^B = B_1 - B_2,$$
$$K_2^A = A_2 - A_3, \quad K_2^B = B_2 - B_3,$$
...

Функции φ и ψ устанавливают следующие взаимно однозначные соответствия:

$$K_0^A \underset{\varphi}{\longleftrightarrow} K_1^B, \quad K_0^B \underset{\psi}{\longleftrightarrow} K_1^A,$$
$$K_2^A \underset{\varphi}{\longleftrightarrow} K_3^B, \quad K_2^B \underset{\psi}{\longleftrightarrow} K_3^A,$$
...

Пусть множества $C = \bigcap_{i=1}^{\infty} A_i$, $D = \bigcap_{i=1}^{\infty} B_i$. Функции φ и ψ устанавливают взаимно однозначные соответствия между C и D. Если бы это было не так, то возникли бы аналогичные кольца в C и D, что по построению C и D невозможно.

Пусть множества

$$A_{\text{чет}} = \bigcup_{i=1}^{\infty} K_{2i}^A = K_0^A \cup K_2^A \cup K_4^A \cup \ldots$$

$$A_{\text{неч}} = \bigcup_{i=1}^{\infty} K_{2i+1}^A = K_1^A \cup K_3^A \cup K_5^A \cup \ldots$$

Аналогично построим множества $B_{\text{чет}}$, $B_{\text{неч}}$. Тогда

$$A = A_{\text{чет}} \cup A_{\text{неч}} \cup C, \quad B = B_{\text{чет}} \cup B_{\text{неч}} \cup D.$$

Функция φ устанавливает взаимно однозначное соответствия:

$$A_{\text{чет}} \underset{\varphi}{\longleftrightarrow} B_{\text{неч}}, \quad A_{\text{неч}} \underset{\psi}{\longleftrightarrow} B_{\text{чет}}, \quad C \underset{\varphi \text{ или } \psi}{\longleftrightarrow} D.$$

Тогда функция $\theta: A \to B$, определенная как

$$\theta(x) = \begin{cases} \varphi(x), & \text{если } x \in A_{\text{чет}} \cup C, \\ \psi(x), & \text{если } x \in A_{\text{неч}} \end{cases}$$

устанавливает взаимно однозначное соответствие между A и B. Следовательно $|A| = |B|$.

Теорема доказана.

Следствие. Если $A \subseteq B$, то $|A| \leqslant |B|$.

Кардинальные числа можно сравнивать по величине.

5. *Шкала мощностей*

Пусть A есть некоторое множество и $\mathcal{P}(A)$ есть множество всех подмножеств множества A. Очевидно, что $|A| \leqslant |\mathcal{P}(A)|$, ибо взаимно однозначное соответствие между A и частью $\mathcal{P}(A)$ устанавливается, если каждому элементу a из A сопоставить одноэлементное множество $\{a\}$ из $\mathcal{P}(A)$.

Теорема (Кантор). $|A| < |\mathcal{P}(A)|$.

Доказательство. Покажем, что $|A| \neq |\mathcal{P}(A)|$. Допустим противное: $|A| = |\mathcal{P}(A)|$ для некоторого множества A. Тогда существует взаимно однозначное соответствие $\varphi: A \to \mathcal{P}(A)$ между множествами A и $\mathcal{P}(A)$. Пусть

$A_1 = \{a \in A: a \in \varphi(a)\}$, $A_2 = \{a \in A: a \notin \varphi(a)\}$.

Тогда $A_2 = A - A_1$. Множество $A_2 \in \mathcal{P}(A)$. Пусть в нашем соответствии $\varphi(b) = A_2$ для некоторого b из A. Каждый элемент из A попадает либо в A_1, либо в A_2. Если $b \in A_1$, то по построению A_1 будет $b \in \varphi(b) = A_2$. Противоречие, ибо $b \in A_1$ и $b \in A_2$ одновременно невозможно. Если $b \in A_2$, то по построению A_2 будет $b \notin \varphi(b) = A_2$. Противоречие, ибо $b \in A_2$ и $b \notin A_2$ одновременно невозможно. Следовательно, наше предположение о равенстве $|A|$ и $|\mathcal{P}(A)|$ не верно. Остается взять $|A| \neq |\mathcal{P}(A)|$, а так как $|A| \leqslant |\mathcal{P}(A)|$, то $|A| < |\mathcal{P}(A)|$. Теорема доказана.

Иногда множество $\mathcal{P}(A)$ всех подмножеств множества A обозначается через 2^A, а мощность $\mathcal{P}(A)$ – через $2^{|A|}$. Тогда по теореме $|A| < 2^{|A|}$.

Отправляясь от произвольного множества A, по теореме Кантора можно построить возрастающую последовательность кардинальных чисел: $|A| < 2^{|A|} < 2^{2^{|A|}} < \ldots$

Отправляясь от счетного множества $\mathbb{N}$ натуральных чисел, можно построить возрастающую последовательность кардиналов:

$$\aleph_0 < 2^{\aleph_0} = c = \aleph_1 < 2^{\aleph_1} = \aleph_2 < 2^{\aleph_2} = \aleph_3 < \ldots$$

Мощности $\aleph_0$, $\aleph_1 = c$, $\aleph_2 = 2^c$, $\aleph_3 = 2^{2^c}$, ... это счетная мощнос-

ть, континуум, гиперконтинуум, гипер-гиперконтинуум и так далее.

Кантор поставил проблему о существовании промежуточной мощности между $\aleph_0$ и $\aleph_1$ (континуум-гипотеза) и промежуточных мощностей между всякими $\aleph_i$ и $\aleph_{i+1}$ (обобщенная континуум-гипотеза). В работах К.Геделя и П.Коэна было установлено, что обе гипотезы не противоречат аксиоматической теории множеств (существует модель, в которой истинны аксиомы теории множеств, континуум-гипотеза, причем правила вывода сохраняют истинность выводимых формул) и не могут быть в ней доказаны (существует модель, в которой истинны аксиомы теории множеств и ложна континуум-гипотеза, причем правила вывода сохраняют истинность выводимых формул, а потому континуум-гипотеза не может быть доказана в теории множеств). Отсюда следует, что обе гипотезы независимы в аксиоматической теории множеств.

6. Унарные функции

Напомним, что функция $f: A \to B$ есть правило (отображение), которое каждому элементу из множества A сопоставляет единственный элемент из множества B. Если $f(a)=b$, то элемент b есть образ элемента a, элемент a есть прообраз элемента b. Множество A есть область определения $D(f)$ функции f. Множество B есть область значений $V(f)$ функции f.

Образ $\operatorname{Im} f = \{f(x): x \in A\}$ отображения $f: A \to B$ есть множество $f(A)$ всех значений функции f. Полный прообраз элемента $y \in B$ есть множество $f^{-1}(y)=\{x \in A: f(x)=y\}$. Полный прообраз множества $C \subseteq B$ есть множество $f^{-1}(C) = \{x \in A: f(x) \in C\}$.

На конечном множестве функцию удобно задавать таблицей. Например, пусть множества $A=\{1,2,3,4\}, B=\{1,2,3,4,5\}$, функция $f=\begin{pmatrix}1&2&3&4\\3&2&1&2\end{pmatrix}$. Здесь $f(1)=3, f(2)=2, f(3)=1, f(4)=2$. Порядок столбцов несущественнен. Область определения $D(f)=A=\{1,2,3,4\}$, область значений $V(f)=B=\{1,2,3,4,5\}$, $\operatorname{Im} f = f(A)=\{1,2,3\}$.

Функции $f: A \to B$ и $g: C \to D$ равны, если $A=C$, $B=D$, $f(x)=g(x)$ $\forall x \in A$.

Функция $I_A: A \to A$, для которой $I(x)=x$ $\forall x \in A$, называется тождественной.

Функция $f: A \to B$ есть *вложение* (*инъективная* функция), если $\forall a, a' \in A$ условие $a \neq a'$ влечет $f(a) \neq f(a')$.

Инъективная функция различные элементы из области определения переводит в различные элементы из области значений.

Функция $f: A \to B$ есть отображение *на* (*сюръективная* функция), если область значений B совпадает с образом $f(A)$, то есть если $f(A)=B$.

Функция $f: A \to B$ есть *взаимно однозначная* функция (или биекция), если f является вложением и отображением на, то есть если 1) $a \neq a' \to f(a) \neq f(a')$, 2) $\operatorname{Im} f = B$.

Если $f: A \to B$ и $C \subset A$, то функция $f: C \to B$ называется *сужением* функции f на множество C и обозначается $f|C$. Функция f называется *расширением* функции $f|C$.

Композиция $g \circ f$ функций $f: A \to B$ и $g: B \to C$ есть функция $g \circ f: A \to C$, для которой $(g \circ f)(x) = g(f(x))$ $\forall x \in A$.

Символ композиции $\circ$ иногда опускается.

Утверждение. $(h \circ g) \circ f = h \circ (g \circ f)$.

Доказательство. Пусть $f: A \to B$, $g: B \to C$, $h: C \to D$. Тогда $((h \circ g) \circ f)(x) = (hg)f)(x) = (h \circ g)(f(x)) = h(g(f(x))) = h((g \circ f)(x)) = (h \circ (g \circ f))(x)$.

Замечание. Для тождественной функции $f \circ I_A = I_B \circ f = f$.

Определение. Функция $f^{-1}: B \to A$ называется *обратной* к функции $f: A \to B$, если $f \circ f^{-1} = I_B$ и $f^{-1} \circ f = I_A$.

Замечание. g обратна к $f \longleftrightarrow f$ обратна g.

Утверждение. Если обратная функция для функции f существует, то она единственна.

Доказательство. Пусть функции f^{-1} и g обратны к функции $f: A \to B$. Тогда $f^{-1} \circ I_B = f^{-1} \circ (f \circ g) = (f^{-1} \circ f) \circ g = I_A \circ g = g$.

Следствие. Пусть для функций f и g существуют обратные функции f^{-1} и g^{-1}. Тогда справедливы утверждения:

1. $(f^{-1})^{-1} = f$.
2. $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

Доказательство. 1. Так как f^{-1} обратна к f, то f обратна к f^{-1}, то есть $f = (f^{-1})^{-1}$.

2. $(f \circ g) \circ (g^{-1} \circ f^{-1}) = f \circ (g \circ g^{-1}) \circ f^{-1} = f \circ I_B \circ f^{-1} = f \circ f^{-1} = I_B$. Аналогично $(g^{-1} \circ f^{-1}) \circ (f \circ g) = I_B$. Тогда, функция $g^{-1} \circ f^{-1}$ обратна к $f \circ g$, то есть $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

Теорема. Функция $f: A \to B$ имеет обратную функцию тогда и только тогда, когда отображение f взаимно однозначно.

Доказательство. Пусть функция f имеет обратную функцию f^{-1}. Покажем, что отображение f взаимно однозначно, то есть что $a \neq a' \to f(a) \neq f(a')$ и $B = \operatorname{Im} f$. В самом деле, пусть $f(a)=f(a')$. Тогда $a = 1_A(a) = f^{-1}(f(a)) = f^{-1}(f(a')) = 1_A(a') = a'$, то есть $f(a)=f(a') \to a=a'$, откуда $a \neq a' \to f(a) \neq f(a')$.

Пусть $b \in B$. Тогда $b = 1_B(b) = (f \circ f^{-1})(b) = f(f^{-1}(b))$, то есть всякий b есть образ некоторого $a = f^{-1}(b) \in A$. Поэтому $B = \operatorname{Im} f$.

Пусть теперь f есть взаимно однозначное отображение. Покажем, что функция f имеет обратную функцию. В самом деле, так как $B = \operatorname{Im} f$, то каждый элемент b из B есть образ в точности одного элемента a из A: $f(a) = b$. Пусть $g(b) = a$. Для соответствия $g: B \to A$ имеем:

$$(g \circ f)(a) = g(f(a)) = g(b) = a = 1_A,$$
$$(f \circ g)(b) = f(g(b)) = f(a) = b = 1_B.$$

Следовательно, g есть обратная функция для f.

Теорема доказана.

7. Отношения

Пусть $A_1, A_2, \ldots, A_n$ – произвольные множества, вообще говоря, разнородные.

Определение. *n-арное отношение* ρ^n на множествах $A_1, A_2, \ldots, A_n$ есть подмножество декартова произведения $A_1 \times A_2 \times \ldots \times A_n$.

Замечание. n-арное отношение ρ^n на множестве A есть подмножество декартова произведения $A \times A \times \ldots \times A$ (n раз). Индекс n арности (местности) отношения иногда опускается.

Иногда отношение определяют на множестве $A_1 \times A_2 \times \ldots \times A_n$.

Возможна предикатная $\rho(x_1, \ldots, x_n)$ и множественная $(x_1, \ldots, x_n) \in \rho$ формы записи отношений. Отношение ρ называют также предикатом. Для бинарного отношения используюся записи $x \rho y$ и $\rho(x, y)$. Унарное отношение $\rho \subseteq E$ есть подмножество из E.

Набор $\mathrm{a} = (a_1, a_2, \ldots, a_n) \in \rho$ (допустима запись $\rho(a_1, a_2, \ldots, a_n)$) называется элементом отношения.

Определение. Отношение *конечно*, если оно состоит из конечного числа элементов.

8. Отношение эквивалентности

Пусть A – произвольное множество.

Определение. Бинарное отношение $\sigma \subseteq A \times A$ есть *отношение эквивалентности* (обозначение $a \sim b$), если оно удовлетворяет следующим аксиомам: $\forall a, b, c \in A$

1) $a \sim a$, рефлексивность,

2) $a \sim b \to b \sim a$, коммутативность,

3) $a \sim b \;\&\; b \sim c \to a \sim c$, транзитивность.

Обозначение. $a \sim b$, $\sigma(a,b)$, $(a,b) \in \sigma$, $a \sigma b$.

Определение. *Разбиение множества* A есть семейство непустых попарно непересекающихся подмножеств из A, в объединении (в сумме) дающих все A: $A = \bigcup_{i \in I} A_i$, $A_i \cap A_j = \varnothing \;\forall i \neq j$. Подмножества A_i называются смежными классами разбиения.

Пример. $A = \{0,1,2,3,4,5\} = \{0,1,5\} \cup \{2\} \cup \{3,4\}$.

Теорема. 1. Каждому отношению эквивалентности, определенному на множестве A, соответствует некоторое разбиение множества A. 2. Каждому разбиению множества A соответствует некоторое отношение эквивалентности.

Коротко: между классом всех определенных на множестве A эквивалентностей и классом всех разбиений множества A существует взаимно однозначное соответствие.

Доказательство. 1. Пусть σ есть отношение эквивалентности, определенное на множестве A. Пусть $a \in A$. Построим множество $K_a = \{x \in A : x \sim a\}$ всех элементов x, эквивалентных a. Оно обозначается также через $[a]_\sigma$. Множества K_a называются *смежными классами A по σ*, или классами эквивалентности.

Заметим, что если $b \in K_a$, то $b \sim a$. Покажем, что $a \sim b \leftrightarrow K_a = K_b$. В самом деле, пусть $a \sim b$. Пусть произвольный элемент $c \in K_a$. Тогда $c \sim a$, $a \sim b$, $c \sim b$, $c \in K_b$ и потому $K_a \subseteq K_b$. Аналогично показываем, что $K_b \subseteq K_a$. Тогда $K_a = K_b$. Пусть теперь $K_a = K_b$. Тогда $a \in K_a = K_b$, $a \in K_b$, $a \sim b$. Утверждение доказано.

Если два класса K_a и K_b имеют общий элемент c, то они совпадают. В самом деле, если $c \in K_a$, $c \in K_b$, то $b \sim c$, $c \sim a$ и $b \sim a$, откуда $K_a = K_b$. Поэтому всякие два класса эквивалентности либо не пересекаются, либо (в случае непустого пересечения) совпадают. Всякий элемент c попадает в класс эквивалентности K_c. Поэтому система смежных классов есть разбиение множества A.

2. Пусть имеем некоторое разбиение множества A. Определим на A отношение $\sim$, положив $a \sim b \leftrightarrow$ элементы a, b принадлежат одному и тому же классу разбиения. Отношение $\sim$ удовлетворяет аксиомам 1) $a \sim a$, 2) $a \sim b \to b \sim a$, 3) $a \sim b \,\&\, b \sim c \to a \sim c$ и потому оно есть отношение эквивалентности.

Замечание. 1. Разбиение множества A на одноэлементные подмножества $A = \bigcup_{a \in A} \{a\}$ и разбиение A, состоящее из одного только множества A, называются тривиальными (несобственными) разбиениями.

2. Разбиение A на одноэлементные подмножества соответствует отношению эквивалентности, которое есть равенство.

3. Разбиение множества A, состоящее из одного только множества A, соответствует отношению эквивалентности, содержащему все множество $A\times A$.

4. $a\sigma b \longleftrightarrow [a]_\sigma = [b]_\sigma$.

Определение. Совокупность классов эквивалентности множества A называется *фактор-множеством* A/σ множества A по эквивалентности σ.

Определение. Отображение $p: A\to A/\sigma$, при котором $p(a) = [a]_\sigma$, называется *каноническим* (естественным).

9. Каноническое разложение функции

Пусть $f:A\to B$ есть некоторая функция. Определим на A отношение $\sigma \subseteq A\times A$, положив $\forall a,a'\in A$ $(a\sim a' \longleftrightarrow f(a)=f(a'))$. Отношение σ есть отношение эквивалентности, ибо $\forall a,b,c\in A$

1) $a\sim a$, ибо $f(a)=f(a)$,

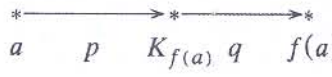
2) $a\sim b \to b\sim a$, ибо $f(a)=f(b) \to f(b)=f(a)$,

3) $a\sim b \,\&\, b\sim c \to a\sim c$, ибо $f(a)=f(b)\ \&\ f(b)=f(c) \to f(a)=f(c)$.

Введенное отношение σ называется *ядерной эквивалентностью* для отображения f. Классы эквивалентности A/σ есть полные прообразы элементов множества B при отображении f, то есть

$$A_b = f^{-1}(b) = \{a\in A: f(a)=b\}.$$

Отображение f можно разложить в композицию двух отображений согласно следующему рисунку:



Имеет место равенство $f = q*p$, то есть

$$f(a) = q(p(a)).$$

Представление $f = q*p$ называется каноническим разложением (представлением) функции f.

Пример. Получить каноническое разложение функции

$$f: E_{10}\to E_{10},\ f=0112105533=\begin{pmatrix}0&1&2&3&4&5&6&7&8&9\\0&1&1&2&1&0&5&5&3&3\end{pmatrix}.$$

Область определения $D(f)=E_{10}$. Область значений $\mathrm{Im}(f)=\{0,1,2,3,5\}$. Классы эквивалентности:

$K_0 = [0]_\sigma = f^{-1}(0) = \{0,5\},\ q(K_0) = 0,$

$K_1 = [1]_\sigma = f^{-1}(1) = \{1,2,4\},\ q(K_1) = 1,$

$K_2 = [2]_\sigma = f^{-1}(2) = \{3\},\ q(K_2) = 2,$

$K_3 = [3]_\sigma = f^{-1}(3) = \{8,9\},\ q(K_3) = 3,$

$K_5 = [5]_\sigma = f^{-1}(5) = \{6,7\},\ q(K_5) = 5.$

Функции p,q задаются следующим образом.

$$p(a) = K_{f(a)} = \begin{pmatrix}0&1&2&3&4&5&6&7&8&9\\K_0&K_1&K_1&K_2&K_1&K_0&K_5&K_5&K_3&K_3\end{pmatrix},$$

$D(p) = E_{10},\ Im(p) = \{K_0,K_1,K_2,K_3,K_5\};$

$$q(K_a) = a = \begin{pmatrix}K_0&K_1&K_2&K_3&K_5\\0&1&2&3&5\end{pmatrix},$$

$D(q) = \{K_0,K_1,K_2,K_3,K_5\},\ Im(q) = \{0,1,2,3,5\};$

$f(a) = q(p(a)).$

Замечание. Каноническое разложение функции можно использовать в кодировании. Например, в рассмотренном примере вместо сообщения $b_1b_2b_3b_4b_5b_6b_7 = 2355131$ в алфавите $\mathrm{Im}(f)$ отправляется сообщение $a_1a_2a_3a_4a_5a_6a_7$, где всякое $a_i\in K_i$, например, $a_1a_2a_3a_4a_5a_6a_7$=3967284. Декодируется оно как

$f(a_1)f(a_2)f(a_3)f(a_4)f(a_5)f(a_6)f(a_7)$=2355131.

10. Определение группы, кольца, поля

Приведем некоторые определения из абстрактной алгебры, на которые мы будем опираться в последующем изложении. Позже мы скажем о них подробнее.

Определение. *Группа* есть множество G на котором определены бинарная операция $x*y$ (умножение), унарная операция x^{-1} (обратный элемент), отмеченный элемент e (единица), удовлетворяющие следующим аксиомам: $\forall x,y,z\in G$

1. $(x*y)*z = x*(y*z)$.
2. $x^{-1}*x = x*x^{-1} = e$.
3. $x*e = e*x = x$.

Группа G коммутативна (или абелева), если дополнительно

4. $x*y = y*x$.

Замечание. 1. Группа обозначается через $(G,\{*,^{-1},e\})$ или просто буквой G. Группа конечна, если она имеет конечное число элементов.

2. Иногда знак произведения $*$ заменяется точкой, иногда опускается совсем и тогда говорят о мультипликативной группе. Если вместо умножения используется знак сложения, то го-

ворят об аддитивной группе. Иногда в мультипликативной группе единица e заменяется на 1, а в аддитивной группе на 0.

3. Степень $a^n = a \cdot a \cdot \ldots \cdot a$ (n раз). По определению полагаем $a^0=e$ и $a^{-n} = (a^{-1})^n$.

Пример. 1. Множество целых чисел $\mathbb{Z}$ со сложением есть есть абелева группа $(\mathbb{Z},\{+,-,e=0\})$.

2. Множество рациональных чисел $\mathbb{Q}$ без нуля с умножением есть абелева группа $(\mathbb{Q}-\{0\},\{\cdot,^{-1},1\})$.

Определение. Множество $H \subseteq G$ есть *подгруппа* группы G, если H есть группа по отношению к операциям, определенным в G.

Замечание. Множество $\langle a \rangle = \{a^n : n \in \mathbb{Z}\}$ всех целых степеней элемента a группы G образует абелеву группу, порожденную элементом a. Множество $\langle a \rangle$ есть подгруппа группы G.

Определение. Группа G называется *циклической*, если существует элемент $a \in G$, для которого $G = \langle a \rangle$. Элемент a называется *генератором* группы G. *Порядок* $\mathrm{ord}(a)$ элемента a группы есть порядок циклической подгруппы $\langle a \rangle$, порожденной элементом a. Если подгруппа $\langle a \rangle$ бесконечна, то $\mathrm{ord}(a) = \infty$.

Определение. *Кольцо* $(R,\{+,\cdot\})$ есть множество R, на котором определены две операции (функции):

сумма $x+y$: $R \times R \to R$ и произведение $x \cdot y$: $R \times R \to R$, удовлетворяющие следующим аксиомам.

1. $(x+y)+z = x+(y+z)$. 2. $x+y = y+x$.
3. $\exists 0 \in R\ x+0 = x$. 4. $\forall x \in R\ \exists (-x) \in R\ x+(-x)=0$.
5. $(xy)z = x(yz)$. 6. $(x+y)z = xz+yz$, $x(y+z) = xy+xz$.

Кольцо *коммутативно*, если произведение коммутативно:

7. $xy = yx$.

Замечание. Кольцо R конечно, если множество R содержит конечное число элементов. Кольцо есть коммутативная группа по сложению.

Определение. *Поле* $(F,\{+,\cdot\})$ есть множество F, на котором определены две операции (функции):

сумма $x+y$: $R \times R \to R$ и произведение $x \cdot y$: $R \times R \to R$, удовлетворяющие следующим аксиомам.

1. $(x+y)+z = x+(y+z)$. 2. $x+y = y+x$.
3. $\exists 0 \in F\ x+0 = x$. 4. $\forall x \in F\ \exists (-x) \in F\ x+(-x)=0$.
5. $(xy)z = x(yz)$. 6. $xy = yx$.
7. $\exists e \in F\ \forall x \in F\ x \cdot e = x$. 8. $\forall x \in F-\{0\}\ \exists x^{-1} \in F\ x \cdot x^{-1}=e$.
9. $(x+y)z = xz+yz$.

Замечание. Поле F конечно, если множество F содержит конечное число элементов. Поле есть коммутативная группа по сложению. Поле без нуля есть коммутативная циклическая группа по умножению.

Работа состоит из пяти частей.

В первой части излагается теория целых чисел, сравнения целых чисел, модулярная арифметика и ее применение в криптографии. Приводятся минимально необходимые сведения из абстрактной алгебры. Ее подробное изложение можно найти в специальных руководствах по абстрактной алгебре. Есть глава о рекуррентных уравнениях и рекуррентных последовательностях.

Во второй части излагаются основные понятия комбинаторики.

В третьей части рассматриваются функции алгебры логики (булевы функции), логика предикатов и их применение.

Четвертая часть посвящена теории графов.

В пятой части излагается теория конечных автоматов и ее связь с логикой одноместных предикатов второго порядка.

Работа написана по материалам лекций автора по курсам "Математическая логика и теория алгоритмов" и "Дискретная математика" в Московском энергетическом институте и в Российском государственном социальном университете. Эти курсы (или им аналогичные) начинали в МЭИ Д.А.Поспелов, В.Н.Вагин, В.П.Кутепов, А.Б.Фролов, А.А.Болотов, повлиявшие на выбор и характер излагаемого автором материала.

Автор выражает искреннюю признательность заведующему кафедрой математической кибернетики Московского государственного университета В.Б.Алексееву и сотруднику этой кафедры В.А.Захарову за рецензии, научную редакцию книги и критические замечания. Автор благодарен также профессорам Московского энергетического института Ю.А.Дубинскому, А.А.Амосову, В.М.Лавыгину за поддержку и помощь в издании книги.

Книга предназначена для студентов высших технических учебных заведений, специализирующимся в области прикладной математики, вычислительной техники, программирования, информатики.

Часть 1. МОДУЛЯРНАЯ АРИФМЕТИКА

1. ДЕЛИМОСТЬ

1.1. Позиционная система счисления

Пусть $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$ есть множество целых чисел, $\mathbb{N} = \{0,1,2,3,\ldots\}$ есть множество натуральных чисел, $\mathbb{N}_+$ есть множество положительных натуральных чисел.

Сложение, вычитание, умножение, деление целых чисел определяются обычным образом.

Пусть a,b,c есть целые числа. Примем следующие обозначения.

$b \mid a$, b делит a (без остатка).

$b \nmid a$, b не делит a (без остатка).

$a:b$, a делится на b (без остатка).

$\lfloor a/b \rfloor$ есть частное от деления a на b.

$\mathrm{mod}(a,b)$ есть остаток от деления a на b.

Определение. $b \mid a$, если $a=b \cdot q$ при некотором q. *a кратно b*, если $a=b \cdot q$ при некотором q. Число b есть *собственный делитель* a, если $b \mid a$, $b \neq \pm 1$, $b \neq \pm a$.

Замечание. 1. Если a кратно b и b кратно c, то a кратно c. В самом деле, $a=bq_1$, $b=cq_2$, откуда $a=bq_1=(cq_2)q_1=c(q_2q_1)=cq$, где $q=q_2q_1$.

2. $1 \mid a$, $a \mid a$. Если $a \mid b$ и $b \mid a$, то $a=+b$ или $a=-b$.

3. Если $a \mid b$, $b \mid c$, то $a \mid c$.

4. Если $a \mid b$, то $a \mid bc$ $\forall c \in \mathbb{Z}$.

5. Если $k \in \mathbb{Z}$, $k \neq 0$, то $a \mid b \longleftrightarrow ka \mid kb$.

6. Если $a \mid b$, $a \mid c$, то $a \mid (bk+cl)$ $\forall k,l \in \mathbb{Z}$.

7. Если $a_1 \mid b_1, \ldots, a_n \mid b_n$, то $(a_1 \cdot \ldots \cdot a_n) \mid (b_1 \cdot \ldots \cdot b_n)$.

8. Если $a \mid b$, то $a^n \mid b^n$ $\forall n \in \mathbb{N}$.

9. Если целые $l,\ldots,n,p,q,\ldots,s$ делятся на b и удовлетворяют равенству $k+l+\ldots+n = p+q+\ldots+s$ are , то k делится на b. В самом деле, $l=l_1b,\ldots,n=n_1b$, $p=p_1b$, $q=q_1b,\ldots,s=s_1b$ и $k=p+q+\ldots+s-l-\ldots-n=(p_1+q_1+\ldots+s_1-l_1-\ldots-n_1)b$.

Теорема (деление с остатком). Пусть $b \in \mathbb{N}_+$. Всякое $a \in \mathbb{Z}$ можно единственным образом представить в виде $a = bq+r$, где $q \in \mathbb{Z}$, $0 \leqslant r < b$.

Доказательство. Одно такое представление $a=bq+r$, $0 \leqslant r < b$, можно получить, если bq есть наибольшее кратное для b, не большее a. Если $a=bq_1+r_1$, $0 \leqslant r_1 < b$, есть другое такое представление, то вычитание дает $0=b(q-q_1)+r-r_1$, $r_1-r=b(q-q_1)$, r_1-r кратно b. Так как $|r_1-r|<b$, то $r_1-r=0$, $r_1=r$, $q_1=q$.

Замечание. Если $a,b \in \mathbb{Z}$, $b \neq 0$, то $\mathrm{mod}(a,b) = a - \lfloor a/b \rfloor \cdot b$.

Пример. Пусть $b=12$.

$129 = 12 \cdot 10 + 9$, $0 \leqslant 9 < 12$;

$-65 = 12 \cdot (-5) - 5 + 12 - 12 = 12 \cdot (-6) + 7$, $0 \leqslant 7 < 12$;

$5 = 12 \cdot 0 + 5$, $0 \leqslant 5 < 12$;

$-5 = 12 \cdot (-1) + 7$, $0 \leqslant 7 < 12$;

$204 = 12 \cdot 17 + 0$, $0 \leqslant 0 < 12$.

Теорема. Для всяких целых $a \geqslant 1$, $h \geqslant 2$ при некотором $s \geqslant 0$ существует единственное представление a в виде

$$a = a_s h^s + a_{s-1} h^{s-1} + \ldots + a_1 h + a_0, \qquad (1.1)$$

где $0 \leqslant a_i \leqslant h-1$ $(i=0,1,\ldots,s-1)$, $1 \leqslant a_s \leqslant h-1$.

Доказательство. *Существование*. Индукция по a.

Базис. $a=1$. Тогда $1=1 \cdot h^0$, $a_0=1$, $s=0$.

Предположение индукции. Допустим, что теорема верна для всякого натурального $a < n$.

Шаг индукции. Покажем, что теорема верна для $a=n$. По предыдущей теореме $n=hb+r$, $0 \leqslant r \leqslant h-1$, $b<n$. Возможны два случая.

1. $b=0$. Тогда $n=r$. Представление (1.1) выполняется при $s = 0$, $c_0=r$.

2. $b \geqslant 1$. Так как $1 \leqslant b < n$, то по предположению индукции $b = b_u h^u + b_{u-1} h^{u-1} + \ldots + b_0$ при некотором $u \geqslant 0$ и $0 \leqslant b_i \leqslant h-1$ $(i=0,1,\ldots,u)$, $b_u \geqslant 1$. Тогда $n = hb+r = h(b_u h^u + b_{u-1} h^{u-1} + \ldots + b_0) + r = b_u h^{u+1} + b_{u-1} h^u + \ldots + b_0 h + r$, и мы опять имеем представление в виде (1.1). Существование доказано.

2. *Единственность*. Если

$$a = a_s h^s + \ldots + a_1 h + a_0 = a'_u h^u + \ldots + a'_1 h + a'_0, \qquad (1.2)$$

то $a = h(a_s h^{s-1} + \ldots + a_1) + a_0 = h(a'_u h^{u-1} + \ldots + a'_1) + a'_0$. Так как представление $a=hq+r$, $0 \leqslant r \leqslant h-1$, единственно, то $a_0=a'_0$ и $d = a_s h^{s-1} + \ldots + a_1 = a'_u h^{u-1} + \ldots + a'_1$. Аналогично, $a_1=a'_1$, $a_2=a'_2$ и так далее. Пусть $s<u$. Тогда $a_0=a'_0$, $a_1=a'_1,\ldots,$ $a_s=a'_s$. Удалим одинаковые слагаемые $a_0, a_1 h, \ldots, a_s h^s$ в (1.2) и получим $a'_u h^u + \ldots + a'_{s+1} h^{s+1} = 0$. Противоречие, ибо $a'_u \geqslant 1$. Поэтому $s<u$ невозможно. Неравенство $s>u$ тоже невозможно. Остается $s=u$.

Теорема доказана.

Определение. Представление (1.1) называется *представлением числа a (в системе счисления) по основанию h*. Числа a_s, $a_{s-1},\ldots,a_0$ называются *цифрами* числа a по основанию h и тог-

да пишут, что по основанию h число $a=(a_s a_{s-1}\ldots a_0)_h$.

Замечание. Представление (1.1) числа a можно рассматривать как многочлен степени s относительно h, который можно использовать для представления в компьютере сверх больших чисел (порядка нескольких сот цифр в десятиричном представлении) и для производства целочисленных арифметических операций над ними – сложения, умножения, вычитания, нахождения частного и остатка при их делении, переход от одной системы счисления к другой и т.д.

1.1.1. Алгоритм вычисления h-ричной записи 10-ричного числа a

ВХОД. Натуральные числа $a>0$ и $h\geq 2$.
ВЫХОД. h-ричная запись числа $a=(a_t a_{t-1}\ldots a_1 a_0)_h$.
1. $i := 0$.
2. Пока $q \neq 0$, выполнять следующее.
   2.1. $r := \mathrm{mod}(a,h)$, $q := (a-r)/h$.
   2.2. $a := q$, $a_i := r$.
   2.3. $i := i+1$.
3. Вернуть a.

Пример. Записать 10-ричное число 160 в 7-ричной системе.
Решение. По основанию h число $a_{10} = (a_t a_{t-1}\ldots a_1 a_0)_h$.
$i := 0$, $a := 160$,
$r := \mathrm{mod}(a,h) = \mathrm{mod}(160,7) = 6$, $q := (a-r)/h = (160-6)/7 = 22$,
$a_0 := r = 6$, $i := i+1 = 0+1 = 1$; $a := q = 22$,
$r := \mathrm{mod}(a,h) = \mathrm{mod}(22,7) = 1$, $q := (a-r)/h = (22-1)/7 = 3$,
$a := q = 3$, $a_1 := r = 1$, $i := i+1 = 1+1 = 2$.
$r := \mathrm{mod}(a,h) = \mathrm{mod}(3,7) = 3$; $q := (a-r)/h = (3-3)/7 = 0$.
$a := q = 0$, $a_2 := r = 3$, $i := i+1 = 2+1 = 3$.
4. Вернуть $(a_t a_{t-1}\ldots a_1 a_0)_h$.
Ответ. $a = 160_{10} = (a_2 a_1 a_0)_7 = 316_7$.

1.2. Простые числа

Определение. Натуральное число $p\geq 2$ есть *простое число*, если p делится только на 1 и на p, то есть p не имеет собственных делителей. Целое $a>2$ есть *составное число*, если a имеет собственные делители.

Замечание. 1. Наименьший положительный делитель q целого $a>1$ есть простое число. В самом деле, пусть $q \mid a$. Если q есть есть составное число, то q имеет делитель q_1, для которого

$1<q_1<q$. Так как $q_1 \mid q$, $q \mid a$, то $q_1 \mid a$. Противоречие с минимальностью q.

2. Если $q>1$ есть наименьший делитель составного целого $a>1$, то $q \leq \sqrt{a}$. В самом деле, так как q есть наименьший делитель a, то $a=qa_1$, $a_1 \geq q$. Перемножим оба выражения и получим $aa_1 \geq qa_1 q$, $a \geq q^2$, $q \leq \sqrt{a}$.

Теорема. Существует бесконечно много простых чисел.

Доказательство. Пусть $p_1, p_2, \ldots, p_k$ есть простые числа. Если число $s = p_1 p_2 \ldots p_k + 1$ простое, то s есть новое простое число. Если s составное, то наименьший больший единицы делитель p для s есть новое простое число. Делитель p не есть один из $p_1, \ldots, p_k$, иначе $p \mid s$, $p \mid (p_1 p_2 \ldots p_k + 1)$, и тогда $p \mid 1$. Противоречие.

Теорема (число простых чисел). Пусть $\pi(x)$ есть число простых чисел, не превосходящих x. Тогда

1. $\lim\limits_{x\to\infty} \dfrac{\pi(x)}{x/\ln x} = 1$.

2. Для $x\geq 17$ $\pi(x) > \dfrac{x}{\ln x}$; для $x>1$ $\pi(x) < 1.25506\dfrac{x}{\ln x}$.

3. Для $x\geq 17$ $\dfrac{x}{\ln x} < \pi(x) < 1.25506\dfrac{x}{\ln x}$.

1.2.1. Тест Миллера–Рабина для простоты числа

ВХОД. Нечетное целое $n\geq 3$ и параметр безопасности $t\geq 1$.
ВЫХОД. Ответ "простое" или "составное" на вопрос: "Является ли n простым числом?"
1. Найти s и нечетное r, для которых $n-1 = 2^s r$.
2. Для i от 1 до t выполнить следующее.
   2.1. Выбрать случайное целое a, $2 \leq a \leq n-1$.
   2.2. Вычислить $y = a^r \pmod n$.
   2.3. Если $y\neq 1$ и $y\neq n-1$, то выполнить следующее.
   $j := 1$.
   Пока $j \leq s-1$ и $y \neq n-1$, выполнить следующее.
   Вычислить $y := y^2 \pmod n$.
   Если $y = 1$, то вернуть "составное".
   $j := j+1$.
   Если $y \neq n-1$, то вернуть "составное".

3. Вернуть "простое".

Замечание. Вероятность получить неверный ответ для целого положительного n меньше $(1/4)^t$.

1.3. Факторизация целых чисел

Всякое целое a и простое p могут иметь общими делителями только 1 или p. В последнем случае a делится на p.

Если произведение нескольких множителей делятся на простое p, то хотя бы один множитель делится на p. Допустим противное: все множители не делятся на p. Тогда произведение этих множителей не делится на p. Противоречие. Тогда хотя бы один множитель делится на p.

Теорема (основная теорема арифметики). Всякое целое большее единицы число можно факторизовать (разложить в произведение (положительных) простых сомножителей) единственным образом с точностью до порядка сомножителей.

Доказательство. Пусть целое $a>1$. Пусть p_1 есть наименьший (положительный) простой делитель a. Тогда $a=p_1a_1$. Если $a_1=1$, нужная факторизация получена. Если $a_1>1$, то аналогично получаем $a_1=p_2a_2$. Если $a_2=1$, нужная факторизация получена. Если $a_2>1$, то получаем $a_2=p_3a_3$. И так далее. Последовательность $a,a_1,a_2,\ldots$ убывает. Поэтому процесс закончится при некотором $a_{n-1}=p_na_n$, $a_n=1$. В результате получаем факторизацию $a=p_1p_2\cdots p_n$.

Покажем единственность этой факторизации. Допустим существование другой: $a=q_1q_2\cdots q_s$ и пусть для определенности $s\geqslant n$. Тогда $p_1p_2\cdots p_n = q_1q_2\cdots q_s$. Правая часть равенства делится на простое q_1. Тогда левая часть равенства делится на q_1 и хотя бы один ее множитель делится на q_1. Пусть $q_1 \mid p_1$. Тогда $p_1=q_1$. Сократим равенство на q_1 и пусть $p_2\cdots p_n = q_2\cdots q_s$. Аналогично получим

$$q_2 = p_2,\ p_3\cdots p_n = q_3\cdots q_s,$$
$$q_3 = p_3,\ p_4\cdots p_n = q_4\cdots q_s,$$
$$\ldots$$
$$q_n = p_n,\ 1 = q_{n+1}\cdots q_s.$$

Поэтому $q_{n+1}=\ldots=q_s=1$ и факторизация единственна.

Замечание. 1. Простые сомножители в факторизации могут повторяться. Пусть $p_1<\ldots<p_k$ есть все различные сомножители в факторизации числа a и a_i есть число вхождений простого p_i

в факторизацию. Тогда представление $a = p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}$ есть *каноническая факторизация* числа a, которая единственна.

2. Если $a = p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}$ есть каноническая факторизация целого a, то $d \mid a \leftrightarrow d = p_1^{b_1}p_2^{b_2}\ldots p_k^{b_k}$, где $0\leqslant b_1\leqslant a_1$, $\ldots, 0\leqslant b_k\leqslant a_k$. Поэтому число a имеет $(a_1+1)(a_2+1)\ldots(a_k+1)$ различных делителей.

3. Иногда каноническая факторизация $a = p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}$ включает все отсутствующие простые числа p между 2 и p_k в виде p^0.

4. Распознавание простоты целого числа с 125 цифрами в его десятиричном представлении существующими методами может быть выполнено в несколько минут. Факторизация такого числа потребует миллионы лет компьютерных вычислений, то есть практически неосуществима.

Замечание. Если положительное целое n удовлетворяет неравенствам $b^{k-1} \leqslant n < b^k$, то n имеет k цифр по основанию b. Логарифмируем неравенства по основанию b и получаем $k-1 \leqslant \log_b n < k$, откуда $k = \lfloor \log_b n \rfloor + 1 = \left\lfloor \frac{\ln n}{\ln b} \right\rfloor + 1$.

1.4. Наибольший общий делитель

Определение. *Общий делитель* од$(a,b,\ldots,l)$ целых $a,b,\ldots,l$ есть всякое целое, которое делит каждое из $a,b,\ldots,l$.

Пример. Целое 3 есть общий делитель для 18,24,36. Целое 6 есть тоже общий делитель для 18,24,36.

Определение. *Наибольший общий делитель* нод$(a,b,\ldots,l)$ или $(a,b,\ldots,l)$ чисел $a,b,\ldots,l$ есть наибольший положитель-ный делитель среди всех общих делителей для $a,b,\ldots,l$. По определению полагают, что нод$(0,\ldots,0)=0$.

Замечание. 1. нод$(a,b,\ldots,l)$ есть наибольшее положительное целое, которое делит каждое целое из $a,b,\ldots,l$.

2. d=нод(a,b), если 1) d=cd(a,b), 2) if $c \mid a$, $c \mid b$, то $c \mid d$.

Определение. Целые $a,b,\ldots,l$ *взаимно просты*, если $(a,b,\ldots,l) = 1$.

Замечание. Если целые числа попарно взаимно просты, то они взаимно просты. Обратное неверно.

Пример. нод$(18,24,36)=6$, нод$(12,24,36)=12$, $(14,28)=7$, $(8,13,21)=1$, $(0,a)=a\ \forall a\neq 0$, $(1,a)=1\ \forall a\neq 0$.

1. Если $a=bq+c$, то множество общих делителей для a и b

совпадает с множеством общих делителей для b и c. В частности, $(a,b)=(b,c)$.

2. Если $c=0$ и не все целые $a,\ldots,b$ равны нулю, то $(a,\ldots,b,c)=(a,\ldots,b)$.

Теорема. Если целые $a>1$, $b>1$ и их канонические факторизации $a = p_1^{a_1}p_2^{a_2}\ldots p_s^{a_s}$, $b = p_1^{b_1}p_2^{b_2}\ldots p_s^{b_s}$, где $p_1,\ldots,p_s$ есть различные простые делители для a или b, то

$$d = (a,b) = p_1^{\min(a_1,b_1)}\cdot p_2^{\min(a_2,b_2)}\cdot\ldots\cdot p)_s^{\min(a_s,b_s)}.$$

Доказательство. Пусть $a>1$, $b>1$ и $p_1,\ldots,p_s$ есть множество всех простых чисел, которые делят хотя бы один из a,b. Если простое p из $p_1,\ldots,p_s$ отсутствует в канонической факторизации a, то добавим к ней множитель p^0. Далее имеем следующее.

1. $d>0$.

2. Так как $a_1\geqslant\min(a_1,b_1),\ldots,a_s\geqslant\min(a_s,b_s)$, то $d\,|\,a$, $d\,|\,b$.

3. Если $h\,|\,a$, $h\,|\,b$, то $h = p_1^{d_1}p_2^{d_2}\ldots p_s^{d_s}$, где $d_1\leqslant a_1$, $d_1\leqslant b_1$, откуда $d_1 \leqslant \min(a_1,b_1)$,

...

$d_s\leqslant a_s$, $d_s\leqslant b_s$, откуда $d_s \leqslant \min(a_s,b_s)$.

Тогда $h\,|\,d$. Следовательно $d = (a,b)$.

Замечание. 1. Если $a_1>1,\ldots,a_n>1$,

$a_1=p_1^{a_{11}}p_2^{a_{12}}\ldots p_s^{a_{1s}},\ldots,\ a_n=p_1^{a_{n1}}p_2^{a_{n2}}\ldots p_s^{a_{ns}}$, где

$p_1,\ldots,p_s$ есть множество всех различных простых делителей чисел $a_1,\ldots,a_n$, то

$$(a_1,\ldots,a_n) = p_1^{\min(a_{11},\ldots,a_{11})}\cdot\ldots\cdot p_s^{\min(a_{n1},\ldots,a_{ns})}.$$

2. $(a_1,\ldots,a_{n-1},a_n) = ((a_1,\ldots,a_{n-1}),a_n)$.

1.4.1. Алгоритм Евклида вычисления наибольшего общего делителя

Пусть a и b есть натуральные числа и $a\geqslant b$. Деление с остатком дает следующую последовательность равенств.

$$\begin{aligned}
a &= bq_1 + r_2,\ 0<r_2<b,\\
b &= r_2q_2 + r_3,\ 0<r_3<r_2,\\
r_2 &= r_3q_3 + r_4,\ 0<r_4<r_3,\\
r_3 &= r_4q_4 + r_5,\ 0<r_5<r_3,\\
&\ldots\\
r_{n-2} &= r_{n-1}q_{n-1} + r_n,\\
r_{n-1} &= r_nq_n\ (\text{здесь } r_{n+1}=0).
\end{aligned}$$

Так как последовательность остатков $r_2,r_3,\ldots$ строго убывает, то $r_{n+1}=0$ при некотором n. Пусть $d=\text{нод}(a,b)$. Тогда из первого равенства получаем, что $d\,|\,a$, $d\,|\,b$, $d\,|\,r_2$, откуда $d = \text{нод}(b,r_2)$, ибо если $d'\,|\,b$, $d'\,|\,r_2$ для некоторого $d'>d$, то $d'\,|\,a$ и $d\neq\text{нод}(a,b)$. Аналогичные рассуждения, примененные к выше написанным равенствам, последовательно дают:

$$d = (a,b) = (b,r_2) = (r_2,r_3) =\ldots= (r_{n-1},r_n) = r_n.$$

Следовательно, $\text{нод}(a,b) = r_n$.

Замечание. Множество общих делителей для a и b совпадает с множеством делителей для $d=(a,b)$.

1.4.1.1. Алгоритм Евклида вычисления нод(a,b)

ВХОД. Натуральные числа a и b, $a\geqslant b$.

ВЫХОД. нод(a,b).

1. Пока $b\neq0$, выполнять следующее.
   1.1. $q:=\lfloor a/b\rfloor$, $r:=a-qb$, $a:=b$, $b:=r$.
2. Вернуть a.

Пример. Найти $(1050,231)$.

$1050 = 231\cdot4 + 126$, остаток $r=126$.

$231 = 126\cdot1 + 105$, остаток $r=105$.

$126 = 105\cdot1 + 21$, остаток $r=21$.

$105 = 21\cdot5$, остаток $r=0$.

$d = (1050,231) = 21$.

Утверждение. $\forall\, m\in\mathbb{N}\ (am,bm) = (a,b)m$.

Доказательство. Умножим равенства алгоритма Евклида на m и получим $(am,bm)=r_nm$. Так как $(a,b)=r_n$, то $(am,bm)=(a,b)m$.

Замечание. Теорема верна для нескольких чисел.

Утверждение. Если $d=\text{cd}(a,b)$, то $\left(\frac{a}{d},\frac{b}{d}\right) = \frac{(a,b)}{d}$.

Доказательство.

$(a,b) = \left(\frac{a}{d}\cdot d,\frac{b}{d}\cdot d\right) = \left(\frac{a}{d},\frac{b}{d}\right)d$, откуда $\left(\frac{a}{d},\frac{b}{d}\right) = \frac{(a,b)}{d}$.

Следствие. $\left(\frac{a}{(a,b)},\frac{b}{(a,b)}\right) = 1$.

Замечание. Теорема верна для нескольких чисел.

Утверждение. Если $(a,b)=1$, то $(ac,b) = (c,b)$.

Доказательство. (ac,b) делит ac,b и ac,bc. Тогда (ac,b) делит $(ac,bc)=(a,b)c=c$, то есть (ac,b) делит c. Но (ac,b)

делит b Поэтому(ac,b) делит (c,b).

(c,b) делит c,b и ac,b. Тогда (c,b) делит (ac,b).

(ac,b) и (c,b) делят друг друга. Тогда $(ac,b) = (c,b)$.

Утверждение. Если $(a,b)=1$ и $b \mid ac$, то $b \mid c$.

Доказательство. Из $(a,b)=1$ следует $(ac,b)=(c,b)$. Так как $b \mid ac$, то $(c,b)=(ac,b)=b$ делит c.

Теорема. Если каждое из $a_1,\ldots,a_m$ взаимно просто с каждым из $b_1,\ldots,b_n$, то произведение $a_1\cdot\ldots\cdot a_m$ взаимно просто с произведением $b_1\cdot\ldots\cdot b_n$.

Доказательство. Пусть $k=1,2,\ldots n$. Тогда

$(a_1,b_k)=1 \rightarrow (a_1a_2,b_k)=(a_2,b_k)=1$.

$(a_1a_2,b_k)=1 \rightarrow (a_1a_2a_3,b_k)=(a_3,b_k)=1$.

$\ldots$

$(a_1a_2\ldots a_{n-1},b_k)=1 \rightarrow (a_1a_2\ldots a_{n-1}a_n,b_k)=(a_n,b_k)=1$.

Пусть $A=a_1a_2\ldots a_n$. Тогда $(A,b_k)=(b_k,A)=1$, $k=1,2,\ldots,n$. Далее, $(b_1,A)=1 \rightarrow (b_1b_2,A)=(b_2,A)=1$.

$(b_1b_2,A)=1 \rightarrow (b_1b_2b_3,A)=(b_1b_2,A)=1$.

$\ldots$

$(b_1\ldots b_{n-1},A)=1 \rightarrow (b_1\ldots b_{n-1}b_n,A)=(b_1\ldots b_{n-1},A)=1$.

$(a_1\cdot\ldots\cdot a_m, b_1\cdot\ldots\cdot b_n)=1$.

Замечание. 1. Если $(a,b)=1$, то $\forall n,m\in\mathbb{N}$ $(a^n,b^m)=1$.

2. Если для некоторых положительных натуральных n и m $(a^n,b^m)=1$, то $(a,b)=1$. В самом деле, если $(a,b)=d$, то $d \mid a$, $d \mid b$, $d \mid a^n$, $d \mid b^m$, $d \mid (a^n,b^m)$, $d \mid 1$, $d=1$.

3. Если p есть простое число, $(a,p^m)\neq 1$, то $(a,p)\neq 1$, $p \mid a$.

4. Если $(a_1,a_2)=d_2,(d_2,a_3)=d_3$, $\ldots$, $(d_{n-1},a_n)=d_n$, то $(a_1, a_2,\ldots,a_n)=d_n$. В самом деле, множество общих делителей для a_1,a_2 совпадает с множеством делителей для $d_2=(a_1,a_2)$. Множество общих делителей для d_2,a_3 (множество общих делителей для a_1,a_2,a_3) совпадает с множеством делителей для $d_3=(d_2, a_3)$. И так далее. Множество общих делителей для a_1,a_2,a_3 совпадает с множеством делителей для $d_n=(d_{n-1},a_n)$. Так как d_n есть наибольший делитель для d_n, то $(a_1,\ldots,a_n) = d_n$.

Теорема. $\forall a_1,\ldots,a_n\in\mathbb{Z}$ $\exists \lambda_1,\ldots,\lambda_n\in\mathbb{Z}$

$$\text{нод}(a_1,\ldots,a_n) = \sum_{i=1}^{n}\lambda_i a_i.$$

Доказательство. Пусть $S = \{\sum_{i=1}^{n}\mu_i a_i$: все $\mu_i\in\mathbb{Z}\}$. Пусть d

$= \sum_{i=1}^{n}\lambda_i a_i$ есть наименьшее положительное целое из S. Покажем, что $d = \text{нод}(a_1,\ldots,a_n)$. Так как $d\neq 0$, то каждое $a_i = q_i d + r_i$ с $0\leqslant r_i<d$. Покажем, что все $r_i=0$. Пусть для простоты $i=1$. Допустим противное: $r_1\neq 0$. Целое $r_1 = a_1 - q_1 d = a_1 - q_1(\lambda_1 a_1+\ldots +\lambda_n a_n) = (1-\lambda_1 q_1)a_1 - q_1\lambda_2 a_2 - \ldots - q_1\lambda_n a_n \in S$ и $0<r_1<d$. Противоречие с минимальностью d. Следовательно все $r_i=0$, $a_i = q_i d$, $d \mid a_i$, $d=\text{cd}(a_1,\ldots,a_n)$. Если s есть любой другой од$(a_1,\ldots,a_n)$, то $a_i=h_i s$, $d = \sum_{i=1}^{n}\lambda_i a_i = \sum_{i=1}^{n}\lambda_i h_i s = s\sum_{i=1}^{n}\lambda_i h_i$, и $s \mid d$.

Тогда $d = \text{нод}(a_1,\ldots,a_n)$.

Следствие. 1. $\forall a_1,a_2\in\mathbb{Z}$ $\exists \lambda_1,\lambda_2\in\mathbb{Z}$ $\text{нод}(a_1,a_2) = \lambda_1 a_1+\lambda_2 a_2$.

2. Если $a_1,a_2\in\mathbb{Z}$ и $1 = \text{нод}(a_1,a_2)$, то $\exists\mu_1,\mu_2\in\mathbb{Z}$ $(1 = \mu_1 a_1 + \mu_2 a_2)$.

1.4.2. Расширенный алгоритм Евклида вычисления наибольшего общего делителя

Алгоритм Евклида может быть описан следующим образом.

$a = bq_1 + r_2$, $0<r_2<b$, $q_1=\lfloor a/b\rfloor$, $r_2=a-bq_1$,
$b = r_2q_2 + r_3$, $0<r_3<r_2$, $q_2=\lfloor b/r_2\rfloor$, $r_3=b-r_2q_2$,
$r_2 = r_3q_3 + r_4$, $0<r_4<r_3$, $q_3=\lfloor r_2/r_3\rfloor$, $r_4=r_2-r_3q_3$,
$r_3 = r_4q_4 + r_5$, $0<r_5<r_3$, $q_4=\lfloor r_3/r_4\rfloor$, $r_5=r_3-r_4q_4$,

$\ldots$

$r_{n-2}=r_{n-1}q_{n-1}+r_n$, $q_{n-1}=\lfloor r_{n-2}/r_{n-1}\rfloor$, $r_n=r_{n-2}-r_{n-1}q_{n-1}$,
$r_{n-1} = r_nq_n$ (here $r_{n+1}=0$), $d=r_n$.

Тогда

$r_2=a-bq_1=a\cdot 1+b(-q_1)=au_1+bv_1$, $u_1=1$, $v_1=-q_1$,
$r_3=b-r_2q_2=b-(au_1+bv_1)=b(1-v_1)+a(-u_1)=au_2+bv_2$,
$u_2=1-v_1$, $v_2=-u_1$,
$r_4=r_2-r_3q_3=(au_1+bv_1)-(au_2+bv_2)q_3=a(u_1-u_2q_3)+b(v_1-v_2q_3) =$
au_3+bv_3, $u_3=u_1-u_2q_3$, $v_3=v_1-v_2q_3$,
$r_5=r_3-r_4q_4=(au_2+bv_2)-(au_3+bv_3)q_3=a(u_2-u_3q_4)+b(v_2-v_3q_4) =$
au_4+bv_4, $u_4=u_2-u_3q_4$, $v_4=v_2-v_3q_4$,

$\ldots$

$d=r_n=r_{n-2}-r_{n-1}q_{n-1}=(au_{n-3}+bv_{n-3})-(au_{n-2}+bv_{n-2})q_{n-1} =$
$a(u_{n-3}-u_{n-2}q_{n-1})+b(v_{n-3}-v_{n-2}q_{n-1}) =$
$au_{n-1}+bv_{n-1}$, $u_{n-1}=u_{n-3}-u_{n-2}q_{n-1}$, $v_{n-1}=v_{n-3}-v_{n-2}q_{n-1}$,
Получили: $d=r_n$, $u=u_{n-1}$, $v=v_{n-1}$.

1.4.2.1. Расширенный алгоритм Евклида вычисления d=нод(a,b), $a\geqslant b$, и чисел u,v, для которых $d = ua+vb$

ВХОД. Натуральные числа a и b, $a\geqslant b$.
ВЫХОД. d=нод(a,b) и целые u,v, для которых $d = ua+vb$.

1. Если b=0, то $d:=a$, $u:=1$, $v:=0$ и return (d,u,v).
2. $u_2:=1$, $u_1:=0$, $v_2:=0$, $v_1:=1$.
3. Пока $b>0$ выполнять следующее.
   3.1. $q:=\lfloor a/b\rfloor$, $r:=a-qb$, $u:=u_2-qu_1$, $v:=v_2-qv_1$.
   3.2. $a:=b$, $b:=r$, $u_2:=u_1$, $u_1:=u$, $v_2:=v_1$, $v_1:=v$.
4. $d:=a$, $u:=u_2$, $v:=v_2$, вернуть (d,u,v).

Пример. Найти d=нод(a,b) и целые u,v, для которых $d=au+bv$. Целые a=5187, b=1520.

Решение. Вычисления приведены в следующей таблице.

n	q	r	u	v	a	b	u_2	u_1	v_2	v_1
0	–	–	–	–	5187	1520	1	0	0	1
1	3	627	1	–3	1520	627	0	1	1	–3
2	2	266	–2	7	627	266	1	–2	–3	7
3	2	95	5	–17	266	95	–2	5	7	–17
4	2	76	–12	41	95	76	5	–12	–17	41
5	1	19	17	–58	76	19	–12	17	41	–58
6	4	0	–80	273	19	0	17	–80	–58	273

Ответ. d = нод(a,b) = нод$(3549,1040)$ = 19, u=17, v=–58.

Теорема. нод(s,t) = нод$(s,t-rs)$ $\forall s,t,r\in\mathbb{N}_+$, $s\leqslant t$.

Доказательство. Если $d\mid s$, $d\mid t$, то $d\mid(t-rs)$. Поэтому всякий од(s,t) есть также од$(s,t-rs)$. Аналогично, всякий од$(s,t-rs)$ есть также од(s,t), ибо $t = (t-rs)+rs$. Получено, что множество всех од(s,t) совпадает с множеством всех од$(s,t-rs)$. Следовательно нод(s,t) = нод$(s,t-rs)$.

Замечание. Эта теорема дает алгоритм вычисления нод(s,t) последовательным вычитанием меньшего из большего, пока получающиеся два целых числа не совпадут. Эти равные целые есть нод(s,t).

Теорема. Если t,m,n есть положительные целые, то

нод$(t^n-1,t^m-1) = t^{\text{нод}(n,m)} - 1$.

Доказательство. Индукция по $\max(n,m)$. Если $\max(n,m)$=1 или $n=m$, то результат тривиален. Иначе допустим $m<n$ и заметим,

что $(t^n-1)-t^{n-m}(t^m-1) = t^{n-m}-1$. Тогда по предыдущей теореме

нод(t^n-1,t^m-1) = ⌈теорема⌉ =

нод$(\underbrace{t^m-1}_{s},\underbrace{t^n-1}_{t})$ = нод$(\underbrace{t^m-1}_{s},\underbrace{t^n-1}_{t}-\underbrace{t^{n-m}}_{r}\underbrace{(t^m-1)}_{s})$ =

нод$(t^m-1,\ t^{n-m}-1)$ = ⌈индукция⌉ =

$t^{\text{нод}(m,n-m)}-1$ = ⌈теорема⌉ = $t^{\text{нод}(n,m)}-1$.

Следствие. При тех же допущениях

нод$(x^{q^n}-x,\ x^{q^d}-x) = x^{q^{\text{нод}(n,d)}}$.

1.5. Наименьшее общее кратное

Определение. *Общее кратное* ок$(a,b,\ldots,l)$ целых $a,b,\ldots,l$ есть всякое целое, которое кратно каждому из $a,b,\ldots,l$.

Определение. *Наименьшее общее кратное* нок$(a,b\ldots,l)$ или $[a,b,\ldots,l]$ целых $a,b,\ldots,l$ есть наименьшее неотрицательное целое среди всех общих кратных для $a,b\ldots,l$.

Замечание. 1. нок$(a,b\ldots,l)$ есть наименьшее неотрицательное целое, которое делится на каждое целое из $a,b,\ldots,l$.

2. d = нок(a,b) если 1) $a\mid d$, $b\mid d$, 2) если $a\mid c$, $b\mid c$, то $d\mid c$.

Пример. ок$(18,21) = 18\cdot 21 = 378$, ок$(6,12,18) = 72$; нок$(18,21) = 126$, $[6,12,18] = 36$.

Замечание. 1. $[a_1,\ldots,a_n]=m$ $\longleftrightarrow$ 1) $0<m\in\mathbb{Z}$, 2) $a_1\mid m,\ldots,a_n\mid m$, 3) если $0<M\in\mathbb{Z}$ и $a_1\mid M,\ldots,a_n\mid M$, то $m\leqslant M$.

2. Если a_n=1, то $[a_1,\ldots,a_{n-1},a_n]=[a_1,\ldots,a_{n-1}]$.

Теорема. Если целые $a>1$, $b>1$ и их факторизации $a = p_1^{a_1}p_2^{a_2}\ldots p_s^{a_s}$, $b = p_1^{b_1}p_2^{b_2}\ldots p_s^{b_s}$, где

$p_1,\ldots,p_s$ есть все различные простые делители a или b, то

$$m = [a,b] = p_1^{\max(a_1,b_1)}\cdot p_2^{\max(a_2,b_2)}\cdot\ldots\cdot p_s^{\max(a_s,b_s)}.$$

Доказательство. Пусть $a>1$, $b>1$ и $p_1,\ldots,p_s$ есть все простых чисел, которые делят хотя бы одно из a,b. Если простое p из $p_1,\ldots,p_s$ отсутствует в канонической факторизации a, то добавим к ней множитель p^0. Аналогично для b. Далее имеем следующее.

1. $m>0$ есть целое число.

2. $a_1\leqslant\max(a_1,b_1),\ldots,a_s\leqslant\max(a_s,b_s)$. Поэтому $a_1\mid m,\ldots,a_n\mid m$.

3. Пусть $M>0$ есть целое число и $a_1\mid M,\ldots,a_n\mid M$. Целое M =

$p_1^{l_1}p_2^{l_2}\dots p_s^{l_s}\cdot N$ (все $l_i\geqslant 0$). Так как

$a_1\mid M,\dots,\ a_n\mid M$, то $a_1\leqslant l_1,\ b_1\leqslant l_1,\ \dots,\ a_s\leqslant l_s, b_s\leqslant l_s$, откуда $l_1\geqslant\max(a_1,b_1),\dots,l_s\geqslant\max(a_s,b_s)$. Поэтому $m\mid M$, откуда $m\leqslant M$. Следовательно, m есть наименьшее общее кратное для a и b.

Замечание. 1. Если $a_1>1,\dots,a_n>1$,

$a_1=p_1^{c_1}p_2^{c_2}\dots p_s^{c_s},\ \dots,\ a_n=p_1^{d_1}p_2^{d_2}\dots p_s^{d_s}$, где $p_1,\dots,p_s$

есть все различные простые числа, каждое из которых делит хотя бы одно из $a_1,\dots,a_n$, то

$$[a_1,\dots,a_n]=p_1^{\max(c_1,\dots,d_1)}\cdot\dots\cdot p_s^{\max(c_s,\dots,d_s)}.$$

2. $[a_1,\dots,a_{n-1},a_n]=[[a_1,\dots,a_{n-1}],a_n]$.

Теорема. Пусть $a\geqslant 1$, $b\geqslant 1$ есть натуральные числа, $d=(a,b)$, $m=[a,b]$. Тогда $dm=ab$.

Доказательство. Пусть $a>1$, $b>1$ и

$a=p_1^{a_1}p_2^{a_2}\dots p_s^{a_s},\ b=p_1^{b_1}p_2^{b_2}\dots p_s^{b_s}$,

где $p_1,\dots,p_s$ есть все различные простые делители чисел a или b. Если простое p из $p_1,\dots,p_s$ отсутствует в канонической факторизации a, то добавим к ней множитель p^0. Аналогично для b. Далее имеем следующее.

$$d=(a,b)=p_1^{\min(a_1,b_1)}\cdot p_2^{\min(a_2,b_2)}\cdot\dots\cdot p_s^{\min(a_s,b_s)},$$
$$m=[a,b]=p_1^{\max(a_1,b_1)}\cdot p_2^{\max(a_2,b_2)}\cdot\dots\cdot p_s^{\max(a_s,b_s)},$$
$$\min(a_1,b_1)+\max(a_1,b_1)=a_1+b_1,$$
$$\dots$$
$$\min(a_s,b_s)+\max(a_s,b_s)=a_s+b_s,\quad\text{откуда } dm=ab.$$

Следствие. $[a,b]=\dfrac{a\cdot b}{(a,b)}$.

Замечание. 1. Всякое $\text{ок}(a,b)=[a,b]\cdot t$ для некоторого натурального t.

2. $[a_1,\dots,a_n]=\dfrac{a_1\cdot\dots\cdot a_n}{(a_1,\dots,a_n)}$.

3. Наименьшее общее кратное взаимно простых чисел равно их произведению.

4. Если $m_1\mid a,\dots,\ m_k\mid a$, то $\text{нок}(m_1,\dots,m_k)\mid a$.

Теорема. Пусть натуральные числа $a\geqslant 2$, $b\geqslant 2$. Тогда a,b взаимно просты, если и только если канонические факторизации

для a,b не имеют общих простых множителей.

Доказательство. Если $(a,b)=1$, то канонические факторизации a,b не имеют общих простых множителей, иначе $(a,b)>1$.

Пусть канонические факторизации для a,b не имеют общих простых множителей. Тогда $a=p_1^{a_1}p_2^{a_2}\dots p_s^{a_s}$, $b=p_1^{b_1}p_2^{b_2}\dots p_s^{b_s}$, где $c_i=\min(\alpha_i,\beta_i)=0$, $i=1,2,\dots,s$. Поэтому

$$d=(a,b)=p_1^{\min(a_1,b_1)}\cdot p_2^{\min(a_2,b_2)}\cdot\dots\cdot p_s^{\min(a_s,b_s)}=1.$$

Следствие. Если p простое число, то верно следующее.

1. $p\nmid a\ \longleftrightarrow$ каноническая факторизация a нет множителя с p.
2. $p\nmid a\ \longleftrightarrow\ (a,p)=1$.

1.6. Непрерывные (цепные) и подходящие дроби

Пусть c есть вещественное число. Пусть q_1 есть наибольшее целое не больше чем c. При нецелом c имеем

$$c=q_1+\frac{1}{c_2},\ c_2>1.$$

Аналогично

$$c_2=q_2+\frac{1}{c_3},\ c_3>1,$$

$$c_3=q_3+\frac{1}{c_4},\ c_4>1,$$

$$\dots$$

$$c_{s-1}=q_{s-1}+\frac{1}{c_s},\ c_s>1.$$

Получили представление c в виде *непрерывной (цепной) дроби*:

$$c=q_1+\cfrac{1}{q_2+\cfrac{1}{q_3+\cdot\cdot\cdot+\cfrac{1}{q_{s-1}+\cfrac{1}{c_s}}}}.$$

Если число c иррационально, то всякое c_s иррационально и дробь продолжается до бесконечности. Если число c рационально, то $c=a/b$ для некоторых целых a,b с $(a,b)=1$, $b>0$. Тогда

непрерывная дробь будет конечной, и с помощью алгоритма Евклида ее можно получить следующим образом.

$$a = bq_1 + r_2, \quad \frac{a}{b} = q_1 + \frac{r_2}{b} = q_1 + \frac{1}{b/r_2},$$

$$b = r_2q_2 + r_3, \quad \frac{b}{r_2} = q_2 + \frac{1}{r_2/r_3},$$

$$r_2 = r_3q_3 + r_4, \quad \frac{r_2}{r_3} = q_3 + \frac{1}{r_3/r_4},$$

$$\ldots$$

$$r_{n-2} = r_{n-1}q_{n-1} + r_n, \quad \frac{r_{n-2}}{r_{n-1}} = q_{n-1} + \frac{1}{r_{n-1}/r_n},$$

$$r_{n-1} = r_nq_n, \quad r_{n+1}=0, \quad \frac{r_{n-1}}{r_n} = q_n.$$

Тогда непрерывная дробь

$$\frac{a}{b} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdot \cdot \cdot + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}}.$$

Дроби $\delta_1=q_1$, $\delta_2 = q_1 + \frac{1}{q_2}$, $\delta_3 = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3}}, \ldots$

называются *подходящими дробями*.

1.6.1. Вычисление подходящих дробей

δ_s можно получить из δ_{s-1} заменой q_{s-1} в δ_{s-1} на $q_{s-1} + 1/q_s$. Получим $P_0=1$, $Q_0=0$. Тогда

$$\delta_1 = \frac{q_1}{1} = \frac{P_1}{Q_1}, \quad \begin{cases} P_1=q_1, \\ Q_1=1, \end{cases} \quad \delta_2=\delta_1(q_1)\Big|_{q_1:=q_2+1/q_2},$$

$$\delta_2 = \frac{q_1+\frac{1}{q_2}}{1} = \frac{q_1q_2+1}{q_2\cdot 1+0} = \frac{q_2P_1+P_0}{q_2Q_1+Q_0} = \frac{P_2}{Q_2},$$

$$\delta_3 = \frac{P_1\left(q_2+\frac{1}{q_3}\right)+P_0}{Q_1\left(q_2+\frac{1}{q_3}\right)+Q_0} = \frac{(P_1q_2+P_0)q_3+P_1}{(Q_1q_2+Q_0)q_3+Q_1} = \frac{q_3P_2+P_1}{q_3Q_2+Q_1} = \frac{P_3}{Q_3},$$

$$\ldots$$

$$\delta_s = \frac{q_sP_{s-1}+P_{s-2}}{q_sQ_{s-1}+Q_{s-2}} = \frac{P_s}{Q_s},$$

$$\ldots$$

1.6.2. Алгоритм вычисления подходящих дробей

$P_0=1$, $Q_0=0$, $P_1=q_1$, $Q_1=1$, $\delta_1 = \frac{P_1}{Q_1}$,

$$\delta_s = \frac{P_s}{Q_s}, \quad \text{где} \begin{cases} P_s=q_sP_{s-1}+P_{s-2}, \\ Q_s=q_sQ_{s-1}+Q_{s-2}, \end{cases} \quad s=2,3,4,\ldots,$$

Пример. Найдем непрерывную дробь для числа 105/38.

$$\begin{aligned} 105 &= 38\cdot 2 + 29, & q_1&=2,\\ 38 &= 29\cdot 1 + 9, & q_2&=1,\\ 29 &= 9\cdot 3 + 2, & q_3&=3,\\ 9 &= 2\cdot 4 + 1, & q_4&=4,\\ 2 &= 1\cdot 2, & q_5&=2. \end{aligned}$$

$$\frac{105}{38} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{1}{q_4 + \cfrac{1}{q_5}}}} = 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{4 + \cfrac{1}{2}}}}.$$

Подходящие дроби. $P_0=1$, $Q_0=0$, $P_1=q_1=2$, $Q_1=1$,

$$\begin{cases} P_0=1, \quad P_1=q_1=2, \\ Q_0=0, \quad Q_1=1, \end{cases} \qquad \delta_1 = \frac{P_1}{Q_1} = \frac{2}{1} = 2,$$

$$\begin{cases} P_2=q_2P_1+P_0=1\cdot 2+1=3, \\ Q_2=q_2Q_1+Q_0=1\cdot 1+0=1, \end{cases} \qquad \delta_2 = \frac{P_2}{Q_2} = \frac{3}{1} = 3,$$

$$\begin{cases} P_3=q_3P_2+P_1=3\cdot 3+2=11, \\ Q_3=q_3Q_2+Q_1=3\cdot 1+1=4, \end{cases} \qquad \delta_3 = \frac{P_3}{Q_3} = \frac{11}{4},$$

$$\begin{cases} P_4=q_4P_3+P_2=4\cdot 11+3=47, \\ Q_4=q_4Q_3+Q_2=4\cdot 4+1=17, \end{cases} \qquad \delta_4 = \frac{P_4}{Q_4} = \frac{47}{17},$$

$$\begin{cases} P_5=q_5P_4+P_3=2\cdot 47+11=105, \\ Q_5=q_5Q_4+Q_3=2\cdot 17+4=38, \end{cases} \qquad \delta_5 = \frac{P_5}{Q_5} = \frac{105}{38}.$$

Теорема. Подходящие дроби δ_s, $s>1$, несократимы.

Доказательство.

$$\delta_s-\delta_{s-1} = \frac{P_s}{Q_s} - \frac{P_{s-1}}{Q_{s-1}} = \frac{P_sQ_{s-1}-Q_sP_{s-1}}{Q_sQ_{s-1}} = \frac{h_s}{Q_sQ_{s-1}};$$

$h_s = P_sQ_{s-1}-Q_sP_{s-1} =$
$(q_sP_{s-1}+P_{s-2})Q_{s-1}-(q_sQ_{s-1}+Q_{s-2})P_{s-1} =$
$q_sP_{s-1}Q_{s-1}+P_{s-2}Q_{s-1} - q_sQ_{s-1}P_{s-1}-Q_{s-2}P_{s-1} =$
$P_{s-2}Q_{s-1} - Q_{s-2}P_{s-1} = -(P_{s-1}Q_{s-2}-Q_{s-1}P_{s-2}) = -h_{s-1}.$

Аналогично получаем $h_s = (-1)h_{s-1} = (-1)^2h_{s-2} =$
$(-1)^3h_{s-3} = \ldots = (-1)^{s-1}h_{s-(s-1)} = (-1)^{s-1}h_1 =$
$(-1)^{s-1}(P_1Q_0-Q_1P_0)=(-1)^{s-1}(q_1\cdot 0-1\cdot 1)=(-1)^{s-1}\cdot(-1) =$
$(-1)^s$. Тогда

$$P_sQ_{s-1}-Q_sP_{s-1} = (-1)^s, \ s>1, \tag{1.3}$$

$$\delta_s-\delta_{s-1} = (-1)^s/(Q_sQ_{s-1}), \ s>1. \tag{1.4}$$

Так как (P_s,Q_s) делит P_s,Q_s и левая часть в (1.3), то (P_s, Q_s) делит правую часть в (1.3). Поэтому $(P_s,Q_s)=1$. Следовательно подходящие дроби $\delta_s=P_s/Q_s$, $s>1$, несократимы.

Замечание. Если c есть вещественное число, $s\geqslant 2$, $c\neq\delta_s$ то по (1.4) c лежит между δ_{s-1},δ_s и

$$|c - \delta_{s-1}| \leqslant |\delta_s - \delta_{s-1}| \leqslant 1/(Q_sQ_{s-1}).$$

2. ФУНКЦИИ МЕБИУСА И ЭЙЛЕРА

2.1. Функции $\lfloor x\rfloor$, $\lceil x\rceil$, $\{x\}$ для вещественного x

Определение. Функция $\lfloor x\rfloor$ (целая часть x) есть наибольшее целое не превосходящее x. Функция $\lceil x\rceil$ есть наименьшее целое не меньшее x. Функция $\{x\} = x - \lfloor x\rfloor$ есть дробная часть x.

Пример. $\lfloor 7\rfloor=7$, $\lfloor 3.6\rfloor=3$, $\lfloor -6.74\rfloor=-7$; $\lceil 7\rceil=7$, $\lceil 3.6\rceil=4$,

$\lceil -6.74\rceil=-6$; $\{7\}=0$, $\{3.6\}=0.6$, $\{-6.74\}=0.26$.

Теорема. Показатель степени простого p в факторизации $n!$ равно $\left\lfloor\frac{n}{p}\right\rfloor + \left\lfloor\frac{n}{p^2}\right\rfloor + \ldots + \left\lfloor\frac{n}{p^m}\right\rfloor + \ldots$

Доказательство. Число множителей в произведении

$n!=1\cdot 2\cdot 3\cdot\ldots\cdot p\cdot\ldots\cdot 2p\cdot\ldots\cdot 3p\cdot\ldots\cdot p^2\cdot\ldots\cdot p^m\cdot\ldots\cdot n$

приводится в табл.2.1. Суммируем числа правого столбца в табл.2.1 и получаем нужный показатель.

Пример. Показатель простого $p=7$ в факторизации 348! равно $\left\lfloor\frac{348}{7}\right\rfloor + \left\lfloor\frac{348}{49}\right\rfloor + \left\lfloor\frac{348}{343}\right\rfloor = 49+7+1 = 57.$

Таблица 2.1

Множество L_1 из кратных p множителей в $n!$	имеет мощность	$\lfloor n/p\rfloor$
Множество L_2 из кратных p^2 множителей в L_1	имеет мощность	$\lfloor n/p^2\rfloor$
...	...	
Множество L_m из кратных p^m множителей в L_{m-1}	имеет мощность	$\lfloor n/p^m\rfloor$
...	...	

2.2. Мультипликативные функции

Определение. Функция $\theta(a)$ называется *мультипликативной функцией*, если она удовлетворяет следующим условиям.

1. $\theta(a)$ определена для всех положительных целых a и не равна нулю хотя бы для одного такого a.

2. Для всяких взаимно простых a_1,a_2

$$\theta(a_1a_2) = \theta(a_1)\theta(a_2).$$

Пример. Функция $\theta(a)=a^s$ мультипликативна, ибо

$$\theta(ab) = (ab)^s = a^sb^s = \theta(a)\theta(b).$$

Замечание. 1. $\theta(a)=\theta(1\cdot a)=\theta(1)\theta(a)$ влечет $\theta(1)=1$.

2. Если функции $\theta_1(a)$, $\theta_2(a)$ мультипликативны, то функция $\theta(a)=\theta_1(a)\theta_2(a)$ тоже мультипликативна, ибо

$\theta(1) = \theta_1(1)\theta_2(1) = 1\cdot 1=1 \neq 0;$
$\theta(ab) = \theta_1(ab)\theta_2(ab) = \theta_1(a)\theta_1(b)\,\theta_2(a)\theta_2(b) =$
$\theta_1(a)\theta_2(a)\,\theta_1(b)\theta_2(b) = \theta(a)\theta(b).$

Теорема. Если функция $\theta(a)$ мультипликативна и натуральные числа $a_1,\ldots,a_s$ попарно взаимно просты, то

$$\theta(a_1a_2\ldots a_s) = \theta(a_1)\theta(a_2)\ldots\theta(a_s).$$

Доказательство. Так как $(a_i,a_j)=1\ \forall i\neq j$, то $(a_1a_2\ldots a_{m-1}, a_m)=1$, $m=2,3,\ldots,s$. По определению мультипликативной функции $\theta(a_1a_2\ldots a_{s-1},a_s) = \theta(a_1a_2\ldots a_{s-1})\theta(a_s) =$

$$\theta(a_1a_2\ldots a_{s-2})\theta(a_{s-1})\theta(a_s) =\ldots= \theta(a_1)\theta(a_2)\ldots\theta(a_s).$$

Теорема. Пусть функция $\theta(a)$ мультипликативна и $a = p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}$ есть каноническая факторизация a. Тогда

$$\sum_{d\mid a}\theta(d) = (1+\theta(p_1)+\theta(p_1^2)+\ldots+\theta(p_1^{a_1}))\cdot$$
$$(1+\theta(p_2)+\theta(p_2^2)+\ldots+\theta(p_2^{a_2}))\cdot$$
$$\ldots$$
$$(1+\theta(p_k)+\theta(p_k^2)+\ldots+\theta(p_k^{a_k})) =$$
$$\prod_{n=1}^{k}(1+\theta(p_n)+\theta(p_n^2)+\ldots+\theta(p_n^{a_n})).$$

Если $a=1$, то считаем правую часть единицей.

Доказательство. $d=p_1^{d_1}p_2^{d_2}\ldots p_k^{d_k} \mid a=p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k} \longleftrightarrow$ $0\leqslant d_1\leqslant a_1,\ 0\leqslant d_2\leqslant a_2,\ \ldots,\ 0\leqslant d_k\leqslant a_k$.

Правая часть равенства есть сумма всех возможных слагаемых, каждое из которых есть произведение k множителей, первое из которых берется из первой скобки, второй множитель из второй скобки, и так далее, k-й множитель из k-й скобки:

$$\prod_{n=1}^{k}(1+\theta(p_n)+\theta(p_n^2)+\ldots+\theta(p_n^{a_n})) =$$
$$\sum_{\substack{(d_1,d_2,\ldots,d_k)\\(d_1\leqslant a_1,\ldots,d_k\leqslant a_k}}\theta(p_1^{d_1})\theta(p_2^{d_2})\ldots\theta(p_k^{d_k}) =$$
$$\sum_{\substack{(d_1,d_2,\ldots,d_k)\\(d_1\leqslant a_1,\ldots,d_k\leqslant a_k}}\theta(\underbrace{p_1^{d_1}p_2^{d_2}\ldots p_k^{d_k}}_{d}) = \sum_{d\mid a}\theta(d),$$

ибо целые $d = p_1^{d_1}p_2^{d_2}\ldots p_k^{d_k}$ при всех возможных наборах $(d_1,\ldots,d_k)$, $d_1\leqslant a_1,\ldots,d_k\leqslant a_k$, пробегают множество всех делителей целого a.

Замечание. Если $\theta(a) = a^s$, то

$$\sum_{d\mid a}d^s = (1+p_1^s+p_1^{2s}+\ldots+p_1^{a_1\cdot s})\cdot$$
$$(1+p_2^s+p_2^{2s}+\ldots+p_2^{a_2\cdot s})\cdot \qquad (2.1)$$
$$\ldots$$
$$(1+p_k^s+p_k^{2s}+\ldots+p_k^{a_k\cdot s}).$$

В частности, левая часть равенства (2.1) при $s=1$ есть сумма $S(a)$ всех делителей целого a и тогда

$$S(a) = \sum_{d\mid a}d = (1+p_1^1+p_1^2+\ldots+p_1^{a_1})\cdot$$
$$(1+p_2^1+p_2^2+\ldots+p_2^{a_2})\cdot$$
$$\ldots$$
$$(1+p_k^1+p_k^2+\ldots+p_k^{a_k}) =$$

$$\left[\text{сумма членов геометрической прогрессии равна } \frac{a_nq-a_1}{q-1}\right]$$

$$\frac{p_1^{a_1+1}-1}{p_1-1}\cdot\frac{p_2^{a_2+1}-1}{p_2-1}\cdot\ldots\cdot\frac{p_k^{a_k+1}-1}{p_k-1}.$$

Левая часть равенства (2.1) при $s=0$ есть число $\tau(a)$ всех делителей целого a и тогда

$$\tau(a) = (a_1+1)(a_2+1)\ldots(a_k+1).$$

Пример. $S(720) = S(2^4\cdot 3^2\cdot 5)=\frac{2^{4+1}-1}{2-1}\cdot\frac{3^{2+1}-1}{3-1}\cdot\frac{5^{1+1}-1}{5-1} = 2418.$

$\tau(720) = (4+1)(2+1)(1+1) = 30.$

2.3. Функция Мебиуса и формула обращения Мебиуса

Определение. Если натуральное число $a>1$ и $a = p_1^{a_1}p_2^{a_2}\cdot\ldots\cdot p_k^{a_k}$ есть каноническая факторизация a, то *функция Мебиуса* $\mu(a) = 0$, если a делится на квадрат простого числа; $\mu(a) = (-1)^k$, если a не делится на квадрат никакого простого числа;

$\mu(1)=1$, то есть

$$\mu(a) = \begin{cases} 1, & \text{если } a=1, \\ 0, & \text{если } a>1 \text{ и } \exists i\in\{1,2,\ldots,k)\ a_i\geqslant 2, \\ (-1)^k, & \text{если } a>1 \text{ и } \forall i\in\{1,\ldots,k\}\ a_i=1, \\ & \text{то есть } a = p_1p_2\ldots p_k. \end{cases}$$

Пример. $\mu(1)=1$; $\mu(2)=(-1)^1=-1$; $\mu(3)=(-1)^1=-1$; $\mu(4)=\mu(2^2)=0$; $\mu(5)=(-1)^1=-1$; $\mu(6)=\mu(2^1\cdot 3^1)=(-1)^2=1$, $\mu(7)=(-1)^1=-1$; $\mu(8)=0$, ибо $2^2\,|\,8$; $\mu(9)=0$, ибо $3^2\,|\,9$; $\mu(10)=\mu(2^1\cdot 5^1)=(-1)^2=1$; $\mu(11)=(-1)^1=-1$; $\mu(12)=0$, ибо $2^2\,|\,12$; $\mu(144)=0$, ибо $3^2\,|\,144$; $\mu(2^1\cdot 7^1\cdot 11^3)=0$, ибо $11^2\,|\,(2^1\cdot 7^1\cdot 11^3)$; $\mu(3^1\cdot 7^1\cdot 11^1)=(-1)^3=-1$; $\mu(3^1\cdot 7^1\cdot 11^1\cdot 17^1)=(-1)^4=1$.

Теорема. $\sum\limits_{d\,|\,n}\mu(d) = \begin{cases} 1, & \text{если } n=1, \\ 0, & \text{если } n>1. \end{cases}$

Доказательство. *Случай* 1. $n=1$. $\mu(1)=1$ по определению функции $\mu(n)$.

Случай 2. $n>1$. Пусть $n = p_1^{a_1}p_2^{a_2}\ldots p_s^{a_s}$ есть каноническая факторизация числа n. Если некоторый делитель d для n делится на квадрат некоторого простого числа p^2, то $\mu(d)=0$. Поэтому в сумме можно оставить только те слагаемые $\mu(d)$, для которых $d\,|\,p_1p_2\ldots p_s$. Тогда

$$\sum_{d\,|\,n}\mu(d) = \sum_{d\,|\,p_1p_2\ldots p_s}\mu(d) = \mu(1) + \sum_{1\leqslant i\leqslant s}\mu(p_i) +$$

$$\sum_{1\leqslant i<j\leqslant s}\mu(p_ip_j) + \ldots + \sum_{1\leqslant i_1<i_2<\ldots<i_s\leqslant s}\mu(p_{i_1}p_{i_2}\ldots p_{i_s}) =$$

$$1 - C_s^1 + C_s^2 - \ldots + (-1)^sC_s^s =$$
$$1^s + C_s^1\cdot 1^{s-1}(-1)^1 + C_s^2\cdot 1^{s-2}(-1)^2 + \ldots + C_s^s(-1)^s =$$
$$(1 + (-1))^s = 0.$$ Теорема доказана.

Теорема (Формула обращения Мебиуса). Если $a(n)$, $b(n)$ есть две функции $\mathbb{N}_+\to G$, где $\mathbb{N}_+ = \{1,2,3,\ldots\}$, G есть коммутативная аддитивная группа (в частности, G может быть группой вещественных или комплексных чисел), то равенство

$$a(n) = \sum_{d\,|\,n} b(d) \text{ справедливо } \forall n\in\mathbb{N}_+ \longleftrightarrow$$

формула обращения Мебиуса

$$b(n) = \sum_{d\,|\,n} a(d)\,\mu\left(\frac{n}{d}\right) \text{ справедлива } \forall n\in\mathbb{N}_+.$$

Доказательство. *Достаточность*. $\leftarrow$. Пусть

$$b(d) = \sum_{e\,|\,d} a(e)\,\mu\left(\frac{d}{e}\right).$$

Суммируем это равенство по $d\,|\,n$. Тогда получаем

$$\sum_{d\,|\,n} b(d) = \sum_{d\,|\,n}\ \sum_{e\,|\,d} a(e)\,\mu\left(\frac{d}{e}\right).$$

Покажем ход доказательства для $n=12$. Все делители $n=12$ есть $1,2,3,4,6,12$. Тогда

$$\sum_{d\,|\,12} b(d) = \sum_{d\,|\,12}\ \sum_{e\,|\,d} a(e)\,\mu\left(\frac{d}{e}\right) =$$

$$\sum_{e\,|\,1} a(e)\,\mu\left(\frac{1}{e}\right) + \sum_{e\,|\,2} a(e)\,\mu\left(\frac{2}{e}\right) + \sum_{e\,|\,3} a(e)\,\mu\left(\frac{3}{e}\right) +$$
$$\sum_{e\,|\,4} a(e)\,\mu\left(\frac{4}{e}\right) + \sum_{e\,|\,6} a(e)\,\mu\left(\frac{6}{e}\right) + \sum_{e\,|\,12} a(e)\,\mu\left(\frac{12}{e}\right) =$$
$$a(1)\mu\left(\frac{1}{1}\right) +$$
$$a(1)\mu\left(\frac{2}{1}\right) + a(2)\mu\left(\frac{2}{2}\right) +$$
$$a(1)\mu\left(\frac{3}{1}\right) + a(3)\mu\left(\frac{3}{3}\right) +$$
$$a(1)\mu\left(\frac{4}{1}\right) + a(2)\mu\left(\frac{4}{2}\right) + a(4)\mu\left(\frac{4}{4}\right) +$$
$$a(1)\mu\left(\frac{6}{1}\right) + a(2)\mu\left(\frac{6}{2}\right) + a(3)\mu\left(\frac{6}{3}\right) + a(6)\mu\left(\frac{6}{6}\right) +$$

$$a(1)\mu\left(\frac{12}{1}\right)+a(2)\mu\left(\frac{12}{2}\right)+a(3)\mu\left(\frac{12}{3}\right)+a(4)\mu\left(\frac{12}{4}\right)+a(6)\mu\left(\frac{12}{6}\right)+a(12)\mu\left(\frac{12}{12}\right)$$

= [группируем по столбцам] =

$$a(1)\left\{\mu\left(\frac{1}{1}\right)+\mu\left(\frac{2}{1}\right)+\mu\left(\frac{3}{1}\right)+\mu\left(\frac{4}{1}\right)+\mu\left(\frac{6}{1}\right)+\mu\left(\frac{12}{1}\right)\right\}+$$

$$a(2)\left\{\mu\left(\frac{2}{2}\right)+\mu\left(\frac{4}{2}\right)+\mu\left(\frac{6}{2}\right)+\mu\left(\frac{12}{2}\right)\right\}+$$

$$a(3)\left\{\mu\left(\frac{3}{3}\right)+\mu\left(\frac{6}{3}\right)+\mu\left(\frac{12}{3}\right)\right\}+$$

$$a(4)\left\{\mu\left(\frac{4}{4}\right)+\mu\left(\frac{12}{4}\right)\right\}+$$

$$a(6)\left\{\mu\left(\frac{6}{6}\right)+\mu\left(\frac{12}{6}\right)\right\}+$$

$$a(12)\left\{\mu\left(\frac{12}{6}\right)\right\}=$$

$$a(1)\sum_{f\left|\frac{12}{1}\right.}\mu(f)+a(2)\sum_{f\left|\frac{12}{2}\right.}\mu(f)+a(3)\sum_{f\left|\frac{12}{3}\right.}\mu(f)+$$

$$a(4)\sum_{f\left|\frac{12}{4}\right.}\mu(f)+a(6)\sum_{f\left|\frac{12}{6}\right.}\mu(f)+a(12)\sum_{f\left|\frac{12}{12}\right.}\mu(f)=$$

$$\sum_{d|12}a(e)\sum_{f\left|\frac{12}{e}\right.}\mu(f)=a(12)\sum_{f|1}\mu(f)=a(12),\ \text{ибо}$$

$$\sum_{f\left|\frac{12}{e}\right.}\mu(f)=\begin{cases}1, & \text{если } e=12,\\ 0, & \text{если } e\mid 12 \text{ и } e\neq 12.\end{cases}$$

Обобщая, имеем $\sum_{d|n} b(d) = a(n)$.

Необходимость. $\rightarrow$. Пусть $a(d) = \sum_{e|d} b(e)$. Тогда

$$\sum_{d|n}\mu\left(\frac{n}{d}\right)a(d)=\sum_{d|n}\mu\left(\frac{n}{d}\right)\underbrace{\sum_{e|d}b(e)}_{a(d)}=\sum_{d|n}\sum_{e|d}b(e)\,\mu\left(\frac{n}{d}\right).$$

Покажем ход доказательства для n=12. Все делители n=12 есть те же 1,2,3,4,6,12. Тогда

$$\sum_{d|12}\mu\left(\frac{n}{d}\right)a(d)=\sum_{d|12}\sum_{e|d}b(e)\,\mu\left(\frac{n}{d}\right)=$$

$$\sum_{e|1}b(e)\,\mu\left(\frac{12}{1}\right)+\sum_{e|2}b(e)\,\mu\left(\frac{12}{2}\right)+\sum_{e|3}b(e)\,\mu\left(\frac{12}{3}\right)+$$

$$\sum_{e|4}b(e)\,\mu\left(\frac{12}{4}\right)+\sum_{e|6}b(e)\,\mu\left(\frac{12}{6}\right)+\sum_{e|12}b(e)\,\mu\left(\frac{12}{12}\right)=$$

$$b(1)\mu\left(\frac{12}{1}\right)+$$

$$b(1)\mu\left(\frac{12}{2}\right)+b(2)\mu\left(\frac{12}{2}\right)+$$

$$b(1)\mu\left(\frac{12}{3}\right)+\quad+b(3)\mu\left(\frac{12}{3}\right)+$$

$$b(1)\mu\left(\frac{12}{4}\right)+b(2)\mu\left(\frac{12}{4}\right)+\quad+b(4)\mu\left(\frac{12}{4}\right)+$$

$$b(1)\mu\left(\frac{12}{6}\right)+b(2)\mu\left(\frac{12}{6}\right)+b(3)\mu\left(\frac{12}{6}\right)+\quad+b(6)\mu\left(\frac{12}{6}\right)+$$

$$b(1)\mu\left(\frac{12}{12}\right)+b(2)\mu\left(\frac{12}{12}\right)+b(3)\mu\left(\frac{12}{12}\right)+b(4)\mu\left(\frac{12}{12}\right)+b(6)\mu\left(\frac{12}{12}\right)+b(12)\mu\left(\frac{12}{12}\right)$$

= [группируем по стобцам] =

$$b(1)\left\{\mu\left(\frac{12}{1}\right)+\mu\left(\frac{12}{2}\right)+\mu\left(\frac{12}{3}\right)+\mu\left(\frac{12}{4}\right)+\mu\left(\frac{12}{6}\right)+\mu\left(\frac{12}{12}\right)\right\}+$$

$$b(2)\left\{\mu\left(\frac{12}{2}\right)+\quad+\mu\left(\frac{12}{4}\right)+\mu\left(\frac{12}{6}\right)+\mu\left(\frac{12}{12}\right)\right\}+$$

$$b(3)\left\{\quad \mu\left(\frac{12}{3}\right) \quad +\mu\left(\frac{12}{6}\right)+\mu\left(\frac{12}{12}\right)\right\} +$$

$$b(4)\left\{\quad \mu\left(\frac{12}{4}\right) \quad +\mu\left(\frac{12}{12}\right)\right\} +$$

$$b(6)\left\{\quad \mu\left(\frac{12}{6}\right)+\mu\left(\frac{12}{12}\right)\right\} +$$

$$b(12)\left\{\quad \mu\left(\frac{12}{12}\right)\right\} =$$

$$b(1)\sum_{f|12}\mu(f) + b(2)\sum_{f|12/2}\mu(f) + b(3)\sum_{f|12/3}\mu(f) +$$

$$b(4)\sum_{f|12/4}\mu(f) + b(6)\sum_{f|12/6}\mu(f) + b(12)\sum_{f|12/12}\mu(f) = b(12),$$

то есть $\sum_{d|12}\mu\left(\frac{n}{d}\right)a(d) = b(12)$.

Обобщая, имеем $b(n) = \sum_{d|12} a(d)\,\mu\left(\frac{n}{d}\right)$.

Необходимость установлена. Теорема доказана.

Замечание 1. Если $d_1,d_2,\ldots,d_s$ есть все делители для n в порядке возрастания, то

$d_1 = \frac{n}{d_s}$, $d_2 = \frac{n}{d_{s-1}}$, $\ldots$, $d_s = \frac{n}{d_1}$. Поэтому

$$b(n) = \sum_{d|n} a(d)\,\mu\left(\frac{n}{d}\right) =$$

$$a(d_1)\,\mu\left(\frac{n}{d_1}\right) + a(d_2)\,\mu\left(\frac{n}{d_2}\right) + \ldots + a(d_s)\,\mu\left(\frac{n}{d_s}\right) =$$

$$a\left(\frac{n}{d_s}\right)\mu(d_s) + a\left(\frac{n}{d_{s-1}}\right)\mu(d_{s-1}) + \ldots + a\left(\frac{n}{d_1}\right)\mu(d_1) =$$

$$\sum_{d|n} a\left(\frac{n}{d}\right)\mu(d).$$

Тогда $b(n) = \sum_{d|n} a(d)\,\mu\left(\frac{n}{d}\right) = \sum_{d|n} a\left(\frac{n}{d}\right)\mu(d)$ есть другой вид формулы обращения Мебиуса.

2. Для мультипликативной группы G (в частности, G может быть группой вещественных или комплексных чисел):

$$a(n) = \prod_{d|n} b(d) \longleftrightarrow b(n) = \prod_{d|n} a(n/d)^{\mu(d)} = \prod_{d|n} a(d)^{\mu(n/d)}.$$

2.4. Функция Эйлера

Определение. *Функция Эйлера* $\varphi(a)$ есть функция, определенная для всех положительных натуральных чисел a и равная числу всех целых в ряду $1,2,\ldots,a$, взаимно простых с a.

Пример. $\varphi(1)=1$, $\varphi(2)=1$, $\varphi(3)=2$, $\varphi(4)=2$, $\varphi(5)=4$, $\varphi(6)=2$.

Теорема. Если p есть простое число, $a\geqslant 1$ есть натуральное число, то $\varphi(p^a) = p^{a-1}(p-1)$.

Доказательство. Покажем справедливость следующего утверждения A:

$x \leqslant p^a$ и нод$(x,p^a)\neq 1 \longleftrightarrow p\,|\,x$ и $x \in T = \{p,2p,$
$3p,\ldots,kp,(k+1)p,\ldots,sp\}$, где $s = p^a/p = p^{a-1}$.

В самом деле, пусть левая часть в A истинна. Так как d = нод$(x,p^a)\neq 1$, то $d=p^c$ при некотором c, $1\leqslant c\leqslant a$, откуда $d=p^c\,|\,x$ и $p\,|\,x$. Далее, $d=p^c=kp$ при $k=p^{c-1}\leqslant p^{a-1}=s$. Поэтому $x\in T$. Получили, что правая часть в A тоже истинна.

Пусть теперь правая часть в A истинна. Так как $p\,|\,x$, $p\,|\,p^a$, то нод$(x,p^a)\neq 1$. Так как $x\in T$, то $x \leqslant sp = p^{a-1}\cdot p = p^a$. Левая часть в A тоже истинна.

Множество T исчерпывает все целые, которые не есть взаимно просты с p^a. Поэтому все другие p^a-p^{a-1} целых между 1 и p^a взаимно просты с p^a. Поэтому функция Эйлера $\varphi(p^a) = p^a-p^{a-1} = p^{a-1}(p-1)$.

Теорема. Функция Эйлера мультипликативна:
$\varphi(a\cdot b) = \varphi(a)\cdot\varphi(b)$ при $(a,b)=1$.

Доказательство. Рассмотрим табл.2.1.

Таблица 2.1

1	2	3	...	r	...	b
b+1	b+2	b+3	...	b+r	...	2b
2b+1	2b+2	2b+3	...	2b+r	...	3b
					...	
(a−1)b+1	(a−1)b+2	(a−1)b+3	...	(a−1)b+r	...	ab

Найдем число целых в табл. 2.1, взаимно простых с ab. Заметим, что $(kb+r,b)=1 \leftrightarrow (r,b)=1$. Поэтому если столбец r в табл.2.1 таков, что $(r,b)=1$, то все целые $kb+r$ столбца r взаимно просты с b. Число таких столбцов равно $\varphi(b)$. Всякий такой столбец r состоит из целых r, $b+r$, $2b+r,\ldots,(a-1)b+r$, то есть числа вида $bx+r$, где x пробегает полную систему вычетов $0,1,\ldots,a-1$ по модулю a. Так как $(a,b)=1$, то столбец r имеет $\varphi(a)$ чисел, взаимно простых с a. Если некоторое целое взаимно просто с a и b, то оно взаимно просто с ab. Поэтому табл.2.1 состоит из $\varphi(a)\varphi(b)$ чисел, взаимно простых с ab. Так как табл.2.1 содержит все целые между 1 и ab, то она имеет $\varphi(ab)$ целых, взаимно простых с ab. Поэтому

$\varphi(ab) = \varphi(a)\varphi(b)$.

Теорема. Пусть $a = p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}$ есть факторизация числа a. Тогда $\varphi(a) = p_1^{a_1-1}p_2^{a_2-1}\ldots p_k^{a_k-1}(p_1-1)(p_2-1)\ldots(p_k-1)$.

Доказательство. Так как $p_1,\ldots,p_k$ есть различные простые числа, то числа $p_1^{a_1},\ldots,p_k^{a_k}$ попарно взаимно просты. Так как функция $\varphi(x)$ мультипликативна (по предыдущей теореме), то

$$\varphi(a) = \varphi(p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}) = \varphi(p_1^{a_1})\varphi(p_2^{a_2})\ldots\varphi(p_k^{a_k}) =$$
$$(p_1^{a_1}-p_1^{a_1-1})(p_2^{a_2}-p_2^{a_2-1})\ldots(p_k^{a_k}-p_k^{a_k-1}) =$$
$$p_1^{a_1-1}p_2^{a_2-1}\ldots p_k^{a_k-1}(p_1-1)(p_2-1)\ldots(p_k-1).$$

Замечание.

1. $\varphi(a)=p_1^{a_1-1}p_2^{a_2-1}\ldots p_k^{a_k-1}(p_1-1)(p_2-1)\ldots(p_k-1)$.
2. $\varphi(a) = (p_1^{a_1}-p_1^{a_1-1})(p_2^{a_2}-p_2^{a_2-1})\ldots(p_k^{a_k}-p_k^{a_k-1})$.
3. $\varphi(a) = p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k}(1-1/p_1)(1-1/p_2)\ldots(1-p_k)$.

Пример. $\varphi(60)=\varphi(2^2\cdot 3\cdot 5)=60(1-1/2)(1-1/3)(1-1/5)=16$, $\varphi(80)=\varphi(2^4\cdot 5)=(2^4-2^3)(5-1)=8\cdot 4=32$, $\varphi(81)=\varphi(3^4)=(3^4-3^3)=81-27=54$, $\varphi(23)=23-1=22$, $\varphi(405)=\varphi(81)\varphi(5)=54\cdot 4=216$.

Замечание. $\varphi(a) > \dfrac{a}{6\ln\ln a}$ $\forall a \geqslant 5$.

Теорема. $a = \sum\limits_{d|a} \varphi(d)$.

Доказательство. Пусть $\theta(a) = \varphi(a)$. По теореме предыдущего параграфа $\sum\limits_{d|a} \varphi(d) =$

$$(1 + \varphi(p_1) + \varphi(p_1^2) + \ldots + \varphi(p_1^{a_1}))\cdot$$
$$\ldots$$
$$(1 + \varphi(p_k) + \varphi(p_k^2) + \ldots + \varphi(p_k^{a_k})) =$$
$$(1 + (p_1-1) + (p_1^2-p_1^1) + \ldots + (p_1^{a_1}-p_1^{a_1-1}))\cdot$$
$$\ldots$$
$$(1 + (p_k-1) + (p_k^2-p_k^1) + \ldots + (p_k^{a_k}-p_k^{a_k-1})) = p_1^{a_1}\ldots p_k^{a_k} = a.$$

Пример. $\sum\limits_{d|12} \varphi(d) = \varphi(1)+\varphi(2)+\varphi(3)+\varphi(4)+\varphi(6)+\varphi(12) =$
$1+1+2+2+2+4 = 12$.

3. СРАВНЕНИЯ

3.1. Сравнение целых чисел

Определение. Целые числа a и b *сравнимы по модулю* m (обозначение $a \equiv b \pmod m$), если $(a-b) \vdots m$. (Модуль m есть положительное целое число.

Теорема. Следующие утверждения эквивалентны.

1. $a \equiv b \pmod m$.
2. $a - b$ делится на m.
3. $a = b + mt$ при некотором целом t.
4. a и b при делении на m дают один и тот же остаток.

Доказательство. $1 \leftrightarrow 2$ верно по определению сравнения.

$2 \leftrightarrow 3$. Разность $(a-b) \vdots m \leftrightarrow a-b=mt$ при некотором целом $t \leftrightarrow a=b+mt$ при некотором целом t.

$2 \leftrightarrow 4$. Пусть $a=q_am+r_1$, $b=q_bm+r_2$, где $0\leqslant r_1,r_2<m$. Тогда $0 \leqslant |r_1-r_2|<m$, $a-b=(q_a-q_b)m+(r_1-r_2)$. Так как $(a-b) \vdots m$, $(q_a-q_b)m \vdots m$, то $(r_1-r_2) \vdots m$. Так как $|r_1-r_2|<m$, то $r_1-r_2=0$, откуда $r_1=r_2$. То есть a и b при делении на m дают один и тот же остаток. Пусть теперь $a=q_am+r$, $b=q_bm+r$, где $0\leqslant r<m$. Тогда $a-b=(q_a-q_b)m$, откуда $(a-b) \vdots m$.

Доказаны эквивалентности 

,

из которых следует теорема.

Теорема. Сравнение есть отношение эквивалентности.

Доказательство. Покажем истинность следующих сравнений.

1. $a \equiv a \pmod m$.
2. $a \equiv b \pmod m$ влечет $b \equiv a \pmod m$.
3. $a \equiv b \pmod m$, $b \equiv c \pmod m$ влечет $a \equiv c \pmod m$.

Первые два сравнения очевидны. Покажем третье. Так как $b=a+mt$, $c=b+ms$ при некоторых $t,s \in \mathbb{Z}$, то $c = a+mt+ms = a+(t+s)m$, откуда $a \equiv c \pmod m$.

3.2. Свойства сравнений

Свойство 1. Сравнения по одному модулю можно почленно складывать. В самом деле, если $a_1 \equiv b_1 \pmod m$, $a_2 \equiv b_2 \pmod m$, то $a_1=b_1+mt_1$, $a_2=b_2+mt_2$ и $a_1+a_2 = b_1+b_2 + m(t_1+t_2)$, откуда $a_1+a_2 \equiv b_1+b_2 \pmod m$.

Замечание. Свойство 1 верно для нескольких сравнений.

Свойство 2. Слагаемые можно переносить из одной части сравнения в другую с обратным знаком. В самом деле, сложим почленно $a+b \equiv c \pmod m$, $-b \equiv -b \pmod m$. Получим $a \equiv c-b \pmod m$.

Свойство 3. К каждой части сравнения можно прибавить (или вычесить) кратное модулю. В самом деле, сложим почленно $a \equiv b \pmod m$ с $mk \equiv 0 \pmod m$ и получим $a+mk \equiv b \pmod m$.

Свойство 4. Сравнения по одному модулю можно почленно перемножать. В самом деле, пусть $a_1 \equiv b_1 \pmod m$, $a_2 \equiv b_2 \pmod m$. Так как $a_1=b_1+mt_1$, $a_2=b_2+mt_2$ и $a_1a_2=(b_1+mt_1)\cdot(b_2+mt_2) = b_1b_2 + m(b_1t_2+b_2t_1+mt_1t_2)=b_1b_2+mt$, $t = b_1t_2+b_2t_1+mt_1t_2$, откуда $a_1a_2 \equiv b_1b_2 \pmod m$.

Замечание. 1. Свойство 4 верно для нескольких сравнений.

2. Обе части сравнения можно возвысить в натуральную степень. Обе части сравнения можно умножить на любое целое.

Свойство 5. Если

1) $S = \sum_{(a_1,\ldots,a_k)} A_{a_1\ldots a_k} x_1^{a_1}\ldots x_k^{a_k}$, где $A_{a_1\ldots a_k}$ есть

целые коэффициенты, а сумма $\sum$ берется по конечному множеству наборов $(a_1,\ldots,a_k)$ с натуральными $a_1,\ldots,a_k$,

2) $A_{a_1\ldots a_k} \equiv B_{a_1\ldots a_k} \pmod m$,

3) $x_i \equiv y_i \pmod m$, $i=1,2,\ldots,k$,

4) $S_1 = \sum_{(a_1,\ldots,a_k)} B_{a_1\ldots a_k} y_1^{a_1}\ldots y_k^{a_k}$,

то $S \equiv S_1 \pmod m$.

Доказательство следует из свойств 1 – 4.

Следствие. Если $x \equiv y \pmod m$, $a_i \equiv b_i \pmod m$, $i=0,1,\ldots,n$, то $a_nx^n+a_{n-1}x^{n+1}+\ldots+a_1x+a_0 \equiv b_ny^n+b_{n-1}y^{n+1}+\ldots+b_1y+b_0 \pmod m$.

Свойство 6. Обе части сравнения $a \equiv b \pmod m$ можно разделить на всякое $d=\text{од}(a,b)$, если $(d,m)=1$. В самом деле, $a=a_1d$, $b=b_1d$, $a-b = a_1d-b_1d = (a_1-b_1)d \vdots m$. Так как d делится на m, то $a_1-b_1 \vdots m$. Поэтому $a_1 \equiv b_1 \pmod m$.

Свойство 7. Сравнение $a \equiv b \pmod m$ влечет $ak \equiv bk \pmod{mk}$ для всякого целого k, ибо $a=b+mt$, $ak=bk+mkt$.

Свойство 8. Сравнение $a \equiv b \pmod m$ можно сократить на всякое $d=\text{од}(a,b,m)$. В самом деле, $a=b+mt$ влечет $a/d = b/d + (m/d)t$ и $a/d \equiv b/d \pmod{(m/d)}$.

Свойство 9. Если $a \equiv b$ по нескольким модулям, то $a \equiv b$ по их наименьшему общему кратному. В самом деле, пусть $a \equiv b \pmod{m_i}$, $i=1,\ldots,k$. $a-b = m_1t_1$, $\ldots, a-b=m_kt_k$. Тогда $a-b$ делится на $m_1,\ldots,m_k$ и потому на $m=\text{нок}(m_1,\ldots,m_k)$ в соответствии с параграфом 1.3. Так как $(a-b) \vdots m$, то $a \equiv b \pmod m$.

Свойство 10. Если $a \equiv b \pmod m$, то $a \equiv b \pmod d$ для всякого (положительного) делителя d модуля m. В самом деле, если $a-b \vdots m \vdots d$, то $a-b \vdots d$, откуда $a \equiv b \pmod d$.

Свойство 11. Если одна часть сравнения и модуль делятся на некоторое целое d, то другая часть сравнения делится на d. В самом деле, пусть $a \equiv b \pmod m$, $a \vdots d$, $m \vdots d$. Тогда $a = b+mt$ и $b \vdots d$.

Свойство 12. Если $a \equiv b \pmod m$, то $(a,m)=(b,m)$, свойство 11.

3.3. Полная система вычетов

Определение. *Класс вычетов (класс сравнимости) по модулю* m есть множество всех целых чисел, которые попарно сравнимы по модулю m (то есть всех целых чисел, которые при делении на m дают один и тот же остаток).

Замечание. Класс вычетов по модулю m $C_r = \{a=mq+r : q\in\mathbb{Z}, 0\leqslant r<m\}$ соответствует остатку r при делении a на m. Существует m классов вычетов C_r, $r = 0,1,\ldots,m-1$, по модулю m.

Определение. *Вычет по модулю m* есть всякое число из класса сравнимости по модулю m. Неотрицательный вычет r из класса вычетов $C_r = \{a=mq+r : q\in\mathbb{Z}, 0\leqslant r<m\}$ по модулю m при $q=0$ в $mq+r$ называется *наименьшим неотрицательным вычетом* класса C_r. *Абсолютно наименьший вычет* класса есть вычет ρ с наименьшим значением $|\rho|$.

Замечание. $\rho=r$, если $r<m/2$ и $\rho=r-m$, если $r>m/2$. Если m четно и $r=m/2$, то ρ есть $m/2$ или $m/2-m$.

Определение. Возьмем только одно число из каждого класса вычетов по модулю m. Полученное множество называется *полной системой вычетов* по модулю m.

Замечание. 1. Всякие m попарно несравнимых целых чисел образуют полную систему вычетов. В самом деле, будучи попарно несравнимы, все m таких целых принадлежат различным классам. Следовательно они образуют полную систему вычетов.

2. Множество целых $\mathbb{Z}_m = \{0,1,\ldots,m-1\}$ образует *наименьшую неотрицательную полную систему вычетов*. Целые

$-(m/2)+1,\ldots,-1,0,1,\ldots,m/2$ при четном m,
$-(m-1)/2,\ldots,-1,0,1,\ldots,(m-1)/2$ при нечетном m

образуют *обсолютно наименьшую полную систему вычетов*.

Теорема. Если $(a,m)=1$ и x пробегает полную систему вычетов по модулю m, то $ax+b$ со всяким целым b тоже пробегает полную систему вычетов по модулю m.

Доказательство. Система целых $ax+b$ имеет m чисел. Допустим, что пара целых ax_1+b, ax_2+b с $x_1\not\equiv x_2 (\mathrm{mod}\, m)$ сравнимы: $ax_1+b \equiv ax_2+b \ (\mathrm{mod}\, m)$. Тогда $ax_1 \equiv ax_2 \ (\mathrm{mod}\, m)$. Так как $(a, m)=1$, то $x_1 \equiv x_2 (\mathrm{mod}\, m)$. Противоречие. Следовательно, $ax_1 + b$, ax_2+b несравнимы. Поэтому система $ax+b$ пробегает полную систему вычетов.

3.3.1. Операции над классами

Пусть a есть целое число. Класс всех целых, сравнимых с a по модулю m, обозначим через $\bar{a}$. Тогда $\bar{a}$ есть класс всех x, для которых $x\equiv a(\mathrm{mod}\, m)$. Множество $\bar{a}$ есть класс всех целых, дающих при делении на m один и тот же остаток. Например, по модулю 10 имеем: $73 \in \overline{13}$, $-17 \in \overline{3}$, $8 \in \overline{-2}$.

Если $a \equiv b(\mathrm{mod}\, m)$, то $\bar{a} = \bar{b}$. Множество всех классов по модулю m есть $\{\bar{0},\bar{1},\ldots,\overline{m-1}\}$. Если $a_i\equiv i(\mathrm{mod}\, m)$, $i = 0,1,\ldots, m-1$, то $C = \{\bar{a}_i : i = 0,1,\ldots,m-1\}$.

Определение. $\bar{a}+\bar{b} = \overline{a+b}$, $\bar{a}\cdot\bar{b} = \overline{a\cdot b}$.

Замечание. 1. $\bar{a}+\bar{0}=\overline{a+0}=\bar{a}$, $\bar{a}\cdot\bar{0}=\overline{a\cdot 0}=\bar{0}$, $\bar{a}\cdot\bar{1}=\overline{a\cdot 1}=\bar{a}$.

2. $-\bar{a} = \bar{0}-\bar{a} = \overline{0-a} = \overline{-a}$.

Отметим следующие свойства сложения и умножения классов.

1. Сумма $\bar{a}+\bar{b}$ и произведение $\bar{a}\cdot\bar{b}$ классов $\bar{a}$ и $\bar{b}$ не зависит от выбора представителей в $\bar{a}$ и $\bar{b}$. В самом деле, если $a'\in\bar{a}$, $b'\in\bar{b}$, то $a'\equiv a(\mathrm{mod}\, m)$, $b'\equiv b(\mathrm{mod}\, m)$, $a'+b'\equiv a+b(\mathrm{mod}\, m)$, $a'\cdot b'\equiv a\cdot b(\mathrm{mod}\, m)$, $\overline{a'+b'}=\overline{a+b}$, $\overline{a'\cdot b'}=\overline{a\cdot b}$.

2. Всякое целое $c\in\bar{a}+\bar{b}$ можно представить в виде $c=a'+b'$, где $a'\in\bar{a}$, $b'\in\bar{b}$. В самом деле, $c\in\bar{a}+\bar{b}=\overline{a+b}$ означает, что $c \equiv a+b \ (\mathrm{mod}\, m)$, $c-a\equiv b(\mathrm{mod}\, m)$, $c-a\in\bar{b}$, откуда $c=a+(c-a)$, где $a \in \bar{a}$, $c-a\in\bar{b}$.

3. Не всякое целое в $\bar{a}\cdot\bar{b}$ можно представить как произведение двух целых из $\bar{a}$ и $\bar{b}$. В самом деле, пусть $m=7$. Тогда $\bar{5}\cdot\bar{3} = \bar{1}$, но 1 не может быть представлена как произведение $a'\cdot b'$, где $a'\in\bar{5}$, $b'\in\bar{3}$.

4. Множество C классов по данному модулю замкнуто относительно сложения и умножения.

Определение. Элемент $g \in C$ есть *генератор* для C по сложению, если элементы $\bar{0}+g$, $\bar{1}+g$, $\bar{2}+g,\ldots,(\overline{m-1})+g$ исчерпывают все элементы из C.

Замечание. 1. Элемент $\bar{1}$ есть генератор для C по сложению.

2. Не всегда существует генератор по умножению для множества $C-\{\bar{0}\}$ (тем более для C).

Справедливы следующие равенства.

1. $\bar{a}+(\bar{b}+\bar{c}) = \bar{a}+\overline{(b+c)}=\overline{a+(b+c)}=\overline{(a+b)+c}=\overline{(a+b)}+\bar{c} = (\bar{a}+\bar{b})+\bar{c}$.

2. $\bar{a}+\bar{0} = \overline{a+0} = \bar{a}$.

3. $\bar{a}+\overline{-a} = \overline{a+(-a)} = \bar{0}$.

4. $\bar{a} + \bar{b} = \overline{a+b} = \overline{b+a} = \bar{b} + \bar{a}$.

5. $\bar{a}\cdot(\bar{b}\cdot\bar{c}) = \bar{a}\cdot\overline{(b\cdot c)}=\overline{a\cdot(b\cdot c)}=\overline{(a\cdot b)\cdot c}=\overline{(a\cdot b)}\cdot\bar{c} = (\bar{a}\cdot\bar{b})\cdot\bar{c}$.

6. $\bar{a} \cdot \bar{b} = \overline{a\cdot b} = \overline{b\cdot a} = \bar{b} \cdot \bar{a}$.

7. $(\bar{a}+\bar{b})\cdot\bar{c} = \overline{(a+b)}\cdot\bar{c}=\overline{(a+b)\cdot c}=\overline{ac+bc}=\overline{ac}+\overline{bc} = \bar{a}\cdot\bar{c}+\bar{b}\cdot\bar{c}$.

8. $\bar{1}\cdot\bar{a} = \bar{a}\cdot\bar{1}=\overline{a\cdot 1} = \bar{a}$.

Тогда

1. $\bar{a}+(\bar{b}+\bar{c}) = (\bar{a}+\bar{b})+\bar{c}$. 2. $\bar{a} + \bar{0} = \bar{a}$.

3. $\bar{a} + \overline{-a} = \bar{0}$. 4. $\bar{a} + \bar{b} = \bar{b} + \bar{a}$.

5. $\bar{a}\cdot(\bar{b}\cdot\bar{c}) = (\bar{a}\cdot\bar{b})\cdot\bar{c}$. 6. $\bar{a} \cdot \bar{b} = \bar{b} \cdot \bar{a}$.

7. $(\bar{a}+\bar{b})\cdot\bar{c} = \bar{a}\cdot\bar{c}+\bar{b}\cdot\bar{c}$. 8. $\bar{1}\cdot\bar{a} = \bar{a}$.

Замечание. 1. Множество C классов по данному модулю с операцией сложения образует конечную аддитивную циклическую абелеву группу с генератором $\bar{1}$.

2. Множество C классов по данному модулю с операцией сложения и умножения классов образует конечное коммутативное кольцо с нулем $\bar{0}$ и единицей $\bar{1}$.

Определение. Пусть $n\in\mathbb{N}_+$. Тогда $n\cdot\bar{a} = \bar{a}+\ldots+\bar{a}$ (n раз), $(\bar{a})^n = \bar{a}\cdot\ldots\cdot\bar{a}$ (n раз), $-n\cdot\bar{a}=n\cdot\overline{-a}$, $0\cdot\bar{a}=\bar{0}$.

Замечание. Для всякого целого n и всякого класса $\bar{a}$ по модулю m справедливо равенство: $n\bar{a}=\overline{na}$. Для всякого $n\in\mathbb{N}$ $(\bar{a})^n = \overline{a^n}$. В самом деле,

$n\bar{a} = \bar{a}+\ldots+\bar{a} = \overline{a+\ldots+a} = \overline{na}$,

$(\bar{a})^n = \bar{a}\cdot\ldots\cdot\bar{a} = \overline{a\cdot\ldots\cdot a} = \overline{a^n}$.

Определение. Пусть $f(x)=c_0+c_1x+c_2x^2+\ldots+c_nx^n$ есть *полином* с целыми коэффициентами и $\bar{a}$ есть класс по модулю m. Положим $f(\bar{a}) = c_0+c_1\bar{a}+c_2(\bar{a})^2+\ldots+c_n(\bar{a})^n$.

Теорема. $f(\bar{a}) = \overline{f(a)}$.

Доказательство. Пусть $f(x)=c_0+c_1x+c_2x^2+\ldots+c_nx^n$ есть полином с целыми коэффициентами, $\bar{a}$ есть класс по модулю m. Тогда $f(\bar{a}) = c_0+c_1\bar{a}+c_2(\bar{a})^2+\ldots+c_n(\bar{a})^n =$

$\overline{c_0}+\overline{c_1a}+\overline{c_2a^2}+\ldots+\overline{c_na^n} = \overline{c_0+c_1a+c_2a^2+\ldots+c_na^n} = \overline{f(a)}$.

Определение. Ненулевые классы $\bar{a}\neq\bar{0}$, $\bar{b}\neq\bar{0}$ по модулю m есть *делители нуля*, если $\bar{a}\cdot\bar{b}=\bar{0}$.

Замечание. 1. Множество классов по составному модулю имеет делители нуля. В самом деле, пусть модуль $m=a\cdot b$, $1<a$, $b<m$. Тогда $a\not\equiv 0\pmod m$, $\bar{a}\neq\bar{0}$. Аналогично $\bar{b}\neq\bar{0}$. Тогда $\bar{a}\cdot\bar{b}=\overline{m}=\bar{0}$ и потому $\bar{a}$ и $\bar{b}$ есть делители нуля.

2. Множество классов по простому модулю p не имеет делителей нуля. В самом деле, пусть $\bar{a}\neq\bar{0}$, $\bar{b}\neq\bar{0}$. Тогда $a\not\equiv 0\pmod p$, $b\not\equiv 0\pmod p$, $p\nmid a$, $p\nmid b$, $p\nmid ab$, $ab\not\equiv 0\pmod p$, $\bar{a}\cdot\bar{b} = \overline{ab} \neq \bar{0}$.

Замечание. Множество C классов по простому модулю с операциями сложения и умножения классов образует конечное поле.

Определение. Класс $(\bar{a})^{-1}$ *обратен к классу* $\bar{a}$ по модулю m, если $(\bar{a})^{-1}\cdot\bar{a} = \bar{1}$. Пишут также $(\bar{a})^{-1} = 1/\bar{a}$. Деление определяется как $\bar{a}/\bar{b} = \bar{a}\cdot(\bar{b})^{-1}$.

Замечание. Не всякий класс $\bar{a}$ имеет обратный класс $(\bar{a})^{-1}$ по модулю m.

3.4. Приведенная система вычетов

Определение. *Приведенная система вычетов* по модулю m есть множество целых из полной системы вычетов, взаимно простых с модулем m.

Замечание. Обычно приведенная система вычетов $\mathbb{Z}_m^*$ образуется из наименьших неотрицательных вычетов $0,1,\ldots,m-1$. Число элементов в приведенной системе вычетов равно $\varphi(m)$, то есть числу положительных целых, взаимно простых с m.

Пример. Приведенная система вычетов по модулю 42 есть множество $\mathbb{Z}_{42}^* = \{1,3,11,13,17,19,23,25,29,31,37,41\}$.

Теорема 1. Всякое множество S из $\varphi(m)$ целых чисел, которые попарно несравнимы по модулю m и *взаимно просты с* m, образуют приведенную систему вычетов по модулю m.

Доказательство. Будучи несравнимы и взаимно просты с

модулем, эти $\varphi(m)$ целых чисел из S лежат в различных $\varphi(m)$ классах целых чисел, взаимно простых с модулем. Так как существует $\varphi(m)$ таких целых и $\varphi(m)$ таких классов, то каждый класс имеет только одного представителя в S и потому S есть приведенная система вычетов по модулю m.

Теорема 2. Если $(a,m)=1$ и x пробегает приведенную систему вычетов по модулю m, то ax тоже пробегает приведенную систему вычетов по модулю m.

Доказательство. Число целых чисел ax равно $\varphi(m)$. По предыдущей теореме, остается показать, что все такие целые попарно несравнимы и взаимно просты с m. Первое доказано в предыдущей теореме для целых более общего вида $ax+b$. Второе следует из $(a,x)=1$, $(x,m)=1$, ибо $(ax,m)=1$.

3.5. Теоремы Эйлера и Ферма

Теорема Эйлера. Если $m>1$ и $(a,m)=1$, то
$a^{\varphi(m)} \equiv 1 \pmod m$.

Доказательство. Если x пробегает приведенную наименьшую неотрицательную систему вычетов $r_1,\ldots,r_c$, $c = \varphi(m)$, то ax тоже пробегает приведенную систему вычетов $ar_1,\ldots,ar_c$ из классов $C_{\rho_1},\ldots,C_{\rho_c}$, где $\rho_1,\ldots,\rho_c$ есть некоторая перестановка элементов $r_1,\ldots,r_c$. Тогда $ar_1 \equiv \rho_1 \pmod m$, $\ldots$, $ar_c \equiv \rho_c \pmod m$. Перемножим эти сравнения почленно и получим $a^c r_1 \ldots r_c \equiv \rho_1 \ldots \rho_c \pmod m$, откуда $a^c \equiv 1 \pmod m$, ибо $r_1 \ldots r_c = \rho_1 \ldots \rho_c$.

Замечание. (Теорема Ферма). Если число p просто и p не делит a, то $a^{p-1} \equiv 1 \pmod p$. Теорема Ферма есть частный случай теоремы Эйлера при $m = p$.

Следствие. $a^p \equiv a \pmod p$ для всякого целого a.

Доказательство. Если p не делит a, то $a^p \equiv a \pmod p$ есть результат умножения $a^{p-1} \equiv 1 \pmod p$ на a. Если $p \mid a$, то тривиально $\frac{a^p}{p} \equiv \frac{a}{p} \pmod 1$, откуда $a^p \equiv a \pmod p$.

3.6. Классы целых чисел по модулю m, взаимно простых с модулем m

Пусть $G = \{\overline{r_1},\overline{r_2},\ldots,\overline{r_{\varphi(m)}}\}$ есть множество классов целых чисел, которые взаимно просты с модулем m. Множество G замкнуто по умножению. В самом деле, если $(\overline{r_i},m)=1$, $(\overline{r_j},m)=1$, то

$(\overline{r_i}\cdot\overline{r_j},m)=(\overline{r_i r_j},m)=1$ и $\overline{r_i}\cdot\overline{r_j} \in G$.

Определение. Элемент $g \in G$ есть *генератор* для G, если $G = \{r^0, r^1, r^2, \ldots, r^{\varphi(m)-1}\}$.

Замечание. Множество G не замкнуто по сложению.

Определение. Класс $(\overline{a})^{-1}$ *мультипликативно обратен* классу $\overline{a}$ по модулю m, если $(\overline{a})^{-1}\cdot\overline{a} = \overline{1}$.

Замечание. Если $\overline{r}\in G$, то $(\overline{r})^{-1} = \overline{r^{\varphi(m)-1}}$ because $\overline{r}\cdot\overline{r^{\varphi(m)-1}} = \overline{r\cdot r^{\varphi(m)-1}} = \overline{r^{\varphi(m)}} = \overline{1}$.

Следующие равенства верны $\forall i=1,\ldots,\varphi(m)$.

1. $\overline{r_i}\cdot(\overline{r_j}\cdot\overline{r_k}) = (\overline{r_i}\cdot\overline{r_j})\cdot\overline{r_k}$.
2. $\overline{r_i}\cdot\overline{r_j} = \overline{r_j}\cdot\overline{r_j}$.
3. $\overline{1}\cdot\overline{r_i} = \overline{r_i}$.
4. $(\overline{r_i})^{-1}\cdot\overline{r_i} = \overline{1}$.

G есть мультипликативная абелева группа. Число элементов в G равно $\varphi(m)$.

3.7. Модулярные арифметические операции

Пусть $\mathbb{Z}$ есть множество целых чисел; $a,b \in \mathbb{Z}$, $m \in \mathbb{N}_+$, $m\geq 2$, и пусть выражение $\mathrm{rest}(a,m)$ означает остаток от деления a на m. Определим целое число, сложение, умножение по модулю m в $\mathbb{Z}$ следующим образом.

$a \pmod m = \mathrm{rest}(a,m)$,
$a + b \pmod m = \mathrm{rest}(a+b,\ m)$,
$a \cdot b \pmod m = \mathrm{rest}(a\cdot b,\ m)$.

Множество $\mathbb{Z}_m = \{0,1,2,\ldots,m-1\}$ есть множество *целых чисел по модулю* m. Множество $\mathbb{Z}_m^* = \{a\in\mathbb{Z}_m : \text{нод}(a,m)=1\}$. Сложение и умножение в $\mathbb{Z}_m$ выполняются по модулю m.

Пример. $\mathbb{Z}_{25} = \{0,1,2,\ldots,24\}$, Тогда $13+16 \pmod{25} = 29 \pmod{25} = 4$. Аналогично, $13\cdot 16 \pmod{25} = 208 \pmod{25} = 8$.

Определение. Пусть $a \in \mathbb{Z}_m$. *Мультипликативно обратное* для a по модулю m есть целое $a^{-1} \in \mathbb{Z}_m$, для которого $a\cdot a^{-1} \pmod m = 1$. Если такое a^{-1} существует, то оно единственно и a называется *обратимым* элементом.

Определение. Пусть $a,b \in \mathbb{Z}_m$. *Деление* a на b по модулю m есть произведение a и b^{-1} по модулю m. Деление определено, если элемент b обратим по модулю m.

Замечание. Пусть $a \in \mathbb{Z}_m$. Тогда a обратимо, если и только если $\text{нод}(a,m) = 1$.

Пример. Элементы 1,2,4,5,7,8 в $\mathbb{Z}_9$ обратимы. Например, $4^{-1} \pmod 9 = 7$, ибо $4 \cdot 7 \pmod 9 = 1$.

Замечание. Множество обратимых элементов в $\mathbb{Z}_m$ с умножением по модулю m есть (мультипликативная) группа.

Определение. *Порядок* $\mathrm{ord}(a)$ *элемента* a *по модулю* m есть наименьшее положительное целое t, для которого $a^t \equiv 1 \pmod m$.

Определение. Элемент a из $\mathbb{Z}_m - \{0\}$ есть *генератор* для $\mathbb{Z}_m - \{0\}$, если его порядок $\mathrm{ord}(a) = m-1$, то есть если множество $\{a^i : i=0,1,\ldots,m-2\} = \mathbb{Z}_m - \{0\}$.

Пример. В следующей таблице приведены степени $a^i \pmod m$ элементов a из $\mathbb{Z}_{15}$ и их порядки.

Exponent i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2		4	9	1	10	6	4	4	6	10	1	9	4	1
3		8	12		5	6	13	2	9	10		3	7	
4		1	6		10	6	1	1	6	10		6	1	
5			3		5	6			9	10		12		
6			9		10	6			6	10		9		
7			12		5	6			9	10		3		
8			6		10	6			6	10		6		
$\mathrm{ord}(a)$	1	4	–	2	–	–	4	4	–	–	2	–	4	2

Множество $\mathbb{Z}_{15} - \{0\}$ генератора (по умножению) не имеет.

Из следующей таблицы видно, что элемент $a=6$ есть генератор для $\mathbb{Z}_{13} - \{0\}$, ибо множество его степеней $\{6^i : i=0,1,\ldots, 11\} = \mathbb{Z}_{13} - \{0\}$.

i	0	1	2	3	4	5	6	7	8	9	10	11
$6^i \pmod{13}$	1	6	10	8	9	2	12	7	3	5	4	11

Замечание. Множество $\mathbb{Z}_p^* = \mathbb{Z}_p - \{0\}$ с умножением по простому модулю p есть (мультипликативная) абелева группа.

3.7.1. Алгоритм вычисления мультипликативно обратного элемента $a^{-1} \pmod n$ *в* $\mathbb{Z}_n$

1. С помощью расширенного алгоритма Евклида найти $d = \mathrm{нод}(a,n)$ и те целые x и y, для которых $ax+ny = d$.
2. Если $d>1$, то $a^{-1} \pmod n$ не существует. Иначе обратный элемент $a^{-1} = x$.

Пример. 1. $a=533$, $n=770$, $d=\mathrm{нод}(a,n)=1$, $x=-13=757 \pmod n$, $y=9$, $a^{-1} \pmod n = 757$;

2. $a=748$, $n=770$, $d=\mathrm{нод}(a,n)=22 \neq 1$, $x=-1=769 \pmod n$, $y=1$, $a^{-1} \pmod n$ не существует.

3.7.2. Алгоритм вычисления модулярной степени в $\mathbb{Z}_n$

Если $k = k_t k_{t-1} \ldots k_1 k_0 = \sum_{i=0}^{t} k_i 2^i$ есть бинарное представление натурального числа k, то $a^k = a^{\sum_{i=0}^{t} k_i 2^i} =$

$$a^{k_0 2^0 + k_1 2^1 + \ldots + k_t 2^t} = a^{k_0 2^0} \cdot a^{k_1 2^1} \cdot \ldots \cdot a^{k_t 2^t} =$$

$$(a^{2^0})^{k_0} (a^{2^1})^{k_1} \ldots (a^{2^t})^{k_t} = \prod_{i=0}^{t} (a^{2^i})^{k_i}.$$

ВХОД. Натуральные числа a,k,n.
ВЫХОД. Степень $a^k \pmod n$

1. $b:=1$. Если $k=0$, то вернуть b.
2. $A:=a$.
3. Если $k_0=1$, то $b:=a$.
4. Для i от 1 до t выполнить следующее:
   4.1. $A:=A^2 \pmod n$.
   4.2. Если $k_i=1$, то $b:=A \cdot b \pmod n$.
5. Вернуть b.

Пример. Найти $a^k = 7^{951} \pmod{1374}$, $a=7$, $k = 951_{10} = (k_9 k_8 \ldots k_1 k_0)_2 = 1110110111_2$.
В следующей таблице приведены шаги вычисления степени.

i	0	1	2	3	4	5	6	7	8	9
k_i	1	1	1	0	1	1	0	1	1	1
A	7	49	1027	871	193	151	817	1099	55	277
b	7	343	517	517	853	1021	1021	895	1135	1123

Ответ. $7^{951} \pmod{1374} = 1123$.

3.7.3. Алгоритм вычисления генератора мультипликативной циклической группы $\mathbb{Z}_p^*$ *при простом* p *(перебор)*

ВХОД. Простое число p.

ВЫХОД. Генератор циклической группы $\mathbb{Z}_p^*$

1. Выбрать случайный элемент a в $\mathbb{Z}_p^*$.
2. $b := a$, $k := 1$.
3. Пока $b \neq 1$ и $k \leqslant p$, выполнить следующее.
   3.1. $b := b \cdot a \pmod p$.
   3.2. $k := k+1$.
4. Если $b=1$ и $k=p-1$, вернуть a. Иначе перейти к шагу 1.

4. СРАВНЕНИЯ С ОДНОЙ ПЕРЕМЕННОЙ

4.1. Решение сравнения с переменными

Определение. Выражение $F(x_1,\ldots,x_s) \equiv 0 \pmod m$, где F есть функция от переменных $x_1,\ldots,x_s$, определенная на множестве $\mathbb{Z}$ целых чисел, называется *сравнением с s переменными*.

Будем рассматривать сравнения вида

$$f(x) \equiv 0 \pmod m, \quad f(x)=a_n x^n+a_{n-1}x^{n-1}+\ldots+a_1 x+a_0. \qquad (4.1)$$

Определение. Если a_n в формуле (4.1) не делится на m, то n есть *степень сравнения* (4.1).

Определение. *Решение сравнения* (4.1) есть множество всех значений переменной x, удовлетворяющих (4.1). Два сравнения эквивалентны, если они имеют одинаковые решения (они удовлетворяются одними и теми же значениями x).

Замечание. 1. Если целое a удовлетворяет сравнению (4.1), то всякое целое x, для которого $x \equiv a \pmod m$, то есть всякое x из класса $\bar{a}$ удовлетворяет сравнению (4.1), ибо $f(x) \equiv f(a) \equiv 0 \pmod m$ по свойству 5 в 3.2. Этот класс $\bar{a}$ целых чисел рассматривается как одно решение. Заметим, что $f(\bar{a}) = \overline{f(a)} = \bar{0}$.

2. Число решений сравнения (4.1) есть число классов, удовлетворяющих (4.1).

Пример. 1. Сравнение $x^5+x+1 \equiv 0 \pmod 7$ удовлетворяется целыми $x=2$ и $x=4$ из полной системы вычетов $0,1,\ldots,6$. Тогда данное сравнение имеет два решения:

$x \equiv 2 \pmod 7$, $x \equiv 4 \pmod 7$.

2. $x^3-2x+6 \equiv 0 \pmod{11}$. Полная система вычетов есть $0,1,\ldots,10$. Одно только целое $x=5$ удовлетворяет сравнению. Тогда единственное решение есть $x \equiv 5 \pmod{11}$.

3. $x^4+2x^3+6 \equiv 0 \pmod 8$. Ни одно целое число из полной системы вычетов $0,1,\ldots,7$ не удовлетворяет сравнению. Данное сравнение не имеет решений.

Более общей задачей является задача решения системы сравнений с одним неизвестным

$$\begin{cases} f_1(x) \equiv 0 \pmod{m_1}, \\ f_2(x) \equiv 0 \pmod{m_2}, \\ \ldots \\ f_s(x) \equiv 0 \pmod{m_s}, \end{cases} \qquad (4.2)$$

где $f_1(x),\ldots,f_s(x)$ есть данные полиномы с целыми коэффициентами.

Пусть $M=[m_1,\ldots,m_s]$ есть наименьшее общее кратное чисел $m_1,\ldots,m_s$. Если целое a удовлетворяет системе (4.2) и $b \equiv a \pmod M$, то $m_1 \mid f_1(a),\ldots,\ m_s \mid f_s(a)$, $f_i(b) \equiv f_i(a) \pmod M$, $i = 1,2,\ldots,s$. Поэтому $f_i(b) \equiv f_i(a) \pmod{m_i}$ и $f_i(b) \equiv 0 \pmod{m_i}$, $i=1,2,\ldots,s$. Итак, если число a удовлетворяет системе (4.2), то системе (4.2) удовлетворяет всякое число из класса $\bar{a}$ по модулю M. Этот класс $\bar{a}$ считается одним из решений системы (4.2).

Определение. *Решение системы сравнений* (4.2) есть класс целых чисел по модулю $M=[m_1,\ldots,m_s]$, состоящих из чисел, удовлетворяющих всем сравнениям этой системы.

Число решений системы (4.2) есть число классов по модулю M, удовлетворяющих всем сравнениям в (4.2).

Пример. 1. $\begin{cases} x^2+x+7 \equiv 0 \pmod 9, \\ x^3-x+3 \equiv 0 \pmod 9. \end{cases}$ Полная система вычетов по модулю 9 есть $0,1,\ldots,8$. Только $x=4$ удовлетворяет обоим сравнениям. Решение есть класс $x \equiv 4 \pmod 9$.

2. $\begin{cases} x^2-3x+2 \equiv 0 \pmod 6, \\ 2x^2+x+2 \equiv 0 \pmod 4. \end{cases}$ $M=[6,4]=12$. Полная система вычетов по модулю 12 есть $0,1,\ldots,11$. Целые $x=2$, $x=-2$ удовлетворяют обоим сравнениям. Решение есть два класса

$x \equiv 2 \pmod{12}$, $x \equiv -2 \pmod{12}$.

Еще более общей является задача решения системы сравнений с несколькими переменными

$$\begin{cases} f_1(x_1,\ldots,x_t)\equiv 0 \pmod{m_1}, \\ f_2(x_1,\ldots,x_t)\equiv 0 \pmod{m_2}, \\ \ldots \\ f_s(x_1,\ldots,x_t)\equiv 0 \pmod{m_s}, \end{cases} \tag{4.3}$$

где $f_1(x_1,..,x_t),\ldots,f_s(x_1,..,x_t)$ есть данные полиномы с целыми коэффициентами. Если

1) $f_i(a_1,\ldots,a_t) \equiv 0 \pmod{m_i}$, $i=1,2,\ldots,s$,
2) $M = [m_1,\ldots,m_s]$,
3) $b_i \equiv a_i \pmod{M}$, $i=1,2,\ldots,t$,

то набор $(b_1,\ldots,b_t)$ из $(\overline{a_1},\ldots,\overline{a_t})=\overline{a_1}\times\ldots\times\overline{a_t}$ удовлетворяет всем сравнениям системы (4.3).

Определение. *Решение системы* (4.3) есть комплекс классов $(\overline{a_1},\ldots,\overline{a_t})=\overline{a_1}\times\ldots\times\overline{a_t}$ по модулю $M=[m_1,\ldots,m_s]$, каждый набор которого удовлетворяет системе (4.3). Число решений есть число таких различных комплексов.

Пример. Найти все решения системы сравнений

$$\begin{cases} x^2-y^2+2\equiv 0 \pmod 6, \\ x^3+x+y+1\equiv 0 \pmod 3. \end{cases}$$

$M=[6,3]=6$. Число наборов (a,b), $0\leqslant a,b\leqslant 6$, равно 36. Наборы (1,3) и (4,0) удовлетворяют данной системе. Она имеет два решения:

1) $x\equiv 1 \pmod 6$, $y\equiv 3 \pmod 6$, 2) $x\equiv 4 \pmod 6$, $y\equiv 0 \pmod 6$.

4.2. Сравнения первой степени

Сравнение первой степени имеет вид

$$ax \equiv b \pmod m. \tag{4.4}$$

1. Пусть $(a,m)=1$. Если x пробегает полную систему вычетов по модулю m, то ax тоже пробегает полную систему вычетов по модулю m. Только одно целое число из полной системы вычетов сравнимо с b. Поэтому сравнение (4.4) имеет только одно решение.

2. Пусть $(a,m)=d>1$. Если $d \nmid b$, то сравнение (4.4) не имеет решений. Пусть b делится на d и $a=a_1d$, $b=b_1d$, $m=m_1d$. Тогда сравнение (4.4) эквивалентно сравнению $a_1x \equiv b_1 \pmod{m_1}$ и $(a_1,m_1)=1$. Последнее имеет только одно решение по модулю m_1. Пусть x_1 есть наименьший неотрицательный вычет этого решения. Тогда его можно записать в виде

$$x \equiv x_1 \pmod{m_1}. \tag{4.5}$$

Так как $m=m_1d$, то имеем решения

$$x_1,\ x_1+m_1,\ x_1+2m_1,\ \ldots,\ x_1+(d-1)m_1$$

по модулю m. Итак, сравнение (4.4) имеет d решений:

$$x_1+i\,m_1 \pmod m,\ i=0,1,\ldots,d-1.$$

Замечание. Решение сравнения (4.4) можно найти с помощью непрерывных дробей. Разложим m/a в непрерывную дробь:

$$\frac{m}{a} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdot \cdot \cdot + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}}.$$

Две последних подходящих дроби есть $\frac{P_{n-1}}{Q_{n-1}}$ и $\frac{P_n}{Q_n} = \frac{m}{a}$. Тогда имеем (по 1.4):

$mQ_{n-1} - aP_{n-1} = (-1)^n$, $-aP_{n-1} = -mQ_{n-1} + (-1)^n$,
$aP_{n-1} = mQ_{n-1} + (-1)^{n-1}$, $aP_{n-1} \equiv (-1)^{n-1} \pmod m$,
$a\cdot(-1)^{n-1}P_{n-1}b \equiv b \pmod m$.

Сравнение (4.4) имеет решение $x \equiv (-1)^{n-1}P_{n-1}b \pmod m$.

Пример. Решить сравнение.

$$111x \equiv 75 \pmod{321}. \tag{4.6}$$

Здесь $(111,321) = 3$ и $75 \vdots 3$. Поэтому сравнение (4.6) имеет три решениеs. Сократим сравнение на 3 и получим эквивалентное сравнение

$$37x \equiv 25 \pmod{107}. \tag{4.7}$$

Разложим дробь $m/a = 107/37$ в непрерывную дробь.

$107 = 37\cdot 2 + 33$, $q_1=2$,
$37 = 33\cdot 1 + 4$, $q_2=1$,
$33 = 4\cdot 8 + 1$, $q_3=8$,
$4 = 1\cdot 4$, $q_4=4$, $n=4$.

Подходящие дроби. $P_0=1$, $Q_0=0$, $P_1=q_1=2$, $Q_1=1$,

$$\delta_1 = \frac{P_1}{Q_1} = \frac{2}{1} = 2,$$

$$\delta_2 = \frac{q_2P_1+P_0}{q_2Q_1+Q_0} = \frac{1\cdot 2+1}{1\cdot 1+0} = \frac{3}{1}, \qquad \begin{matrix} P_2 = 3, \\ Q_2 = 1, \end{matrix}$$

$$\begin{array}{r|l} 107 & 37 \\ \hline 74 & 2 \\ \hline 33 & 1 \\ \hline 32 & 8 \\ \hline 4 & \\ \hline 0 & \end{array}$$

$$\frac{107}{37} = 2 + \cfrac{1}{1 + \cfrac{1}{8 + \cfrac{1}{4}}}.$$

$$\delta_3 = \frac{q_3P_2+P_1}{q_3Q_2+Q_1} = \frac{8\cdot 3+2}{8\cdot 1+1} = \frac{26}{9}, \qquad \begin{matrix} P_3 = 26, \\ Q_3 = 9, \end{matrix}$$

$$\delta_4 = \frac{q_4P_3+P_2}{q_4Q_3+Q_2} = \frac{4\cdot 26+3}{4\cdot 9+1} = \frac{107}{37}, \qquad \begin{matrix} P_4 = 107, \\ Q_4 = 37, \end{matrix}$$

В этом случае, $n=4$, $P_{n-1}=P_3=26$, $b=25$. Решение сравнения (4.7) есть $x \equiv -26\cdot 25 \equiv 99 \pmod{107}$, откуда

$x \equiv 99,\ 99+107,\ 99+2\cdot 107 \pmod{321}$,

то есть имеем три решения сравнения (4.6):

$x \equiv 99 \pmod{321}$, $x \equiv 206 \pmod{321}$, $x \equiv 313 \pmod{321}$.

4.3. Система сравнений первой степени

4.3.1. Попарно взаимно простые модули

Рассмотрим простейшую систему сравнений

$$\begin{cases} x \equiv c_1 \pmod{m_1} \\ x \equiv c_2 \pmod{m_2} \\ \dots \\ x \equiv c_k \pmod{m_k} \end{cases} \qquad (4.8)$$

с неизвестным x и с попарно взаимно простыми модулями.

Теорема (Китайская теорема об остатках). Пусть $M = m_1m_2\cdot\dots m_k$, $M_s=M/m_s$ и N_s определяется из условия $M_sN_s \equiv 1 \pmod{m_s}$, то есть $N_s = M_s^{-1} \pmod{m_s}$, $s=1,2,\dots,k$. Пусть

$x_0 = M_1N_1c_1 + M_2N_2c_2 + \dots + M_kN_kc_k$.

Тогда единственное решение системы (4.8) есть

$$x \equiv x_0 \pmod{m_1m_2\dots m_k}. \qquad (4.9)$$

Доказательство (Гаусса). $M_s = m_1\dots m_{s-1}m_{s+1}\dots m_k$, N_s есть решение сравнения $M_sN_s \equiv 1 \pmod{m_s}$, $s=1,2,\dots,k$. Пусть

$$1 \equiv M_1N_1 \pmod{m_1}, \dots, 1 \equiv M_kN_k \pmod{m_k}. \qquad (4.10)$$

Перемножим (4.8) и (4.10) почленно:

$$x \equiv M_1N_1c_1 \pmod{m_1}, \ \dots, \ x \equiv M_kN_kc_k \pmod{m_k}.$$

Добавим к правой части этих сравнений кратное модулей соответствующих сравнений:

$$x \equiv M_1N_1c_1 + \underbrace{M_2N_2c_2 + \dots + M_kN_kc_k}_{\vdots\, m_1} = x_0 \pmod{m_1},$$

$$\dots$$

$$x \equiv \underbrace{M_1N_1c_1 + \dots + M_{k-1}N_{k-1}c_{k-1}}_{\vdots\, m_k} + M_kN_kc_k = x_0 \pmod{m_k}.$$

Тогда $x \equiv x_0 \pmod{[m_1,\dots,m_k]}$, или $x \equiv x_0 \pmod{m_1\dots m_k}$.

4.3.2. Алгоритм Гаусса для системы сравнений $x \equiv c_1 \pmod{m_1}, \dots, x \equiv c_k \pmod{m_k}$ с попарно взаимно простыми модулями

$$x = \left(\sum_{s=1}^{k} c_sM_sN_s \right) \pmod{M},$$

где $M=m_1m_2\dots m_k$, $M_s=M/m_s$, $N_s \equiv M_s^{-1} \pmod{m_s}$, $s=1,2,\dots,k$.

Пример. Решить систему сравнений

$x \equiv 1 \pmod 4$, $x \equiv 3 \pmod 5$, $x \equiv 2 \pmod 7$.

$c_1=1$, $c_2=3$, $c_3=2$, $m_1=4$, $m_2=5$, $m_3=7$.

Решение. $M=m_1m_2m_3=4\cdot 5\cdot 7=140$, $M_1=5\cdot 7=35$, $M_2=4\cdot 7=28$, $M_3=M_1=4\cdot 5=20$. Сравнения $M_1N_1 \equiv 1 \pmod{m_1}$, $M_2N_2 \equiv 1 \pmod{m_2}$, $M_3N_3 \equiv 1 \pmod{m_3}$ есть $35N_1 \equiv 1 \pmod 4$, $28N_2 \equiv 1 \pmod 5$, $20N_3 \equiv 1 \pmod 7$ и они удовлетворяются целыми $N_1=3$, $N_2=2$, $N_3=6$. Тогда x_0 =

$M_1N_1c_1 + M_2N_2c_2 + M_3N_3c_3 = 35\cdot 3\cdot 1+28\cdot 2\cdot 3+20\cdot 6\cdot 2=513$. Решение $x\equiv 513 (\mathrm{mod}\, 140)$, или $x\equiv 93(\mathrm{mod}\, 140)$. *Ответ*. $x\equiv 93(\mathrm{mod}\, 140)$.

4.3.3. Произвольные модули

Рассмотрим систему сравнений (S) $\begin{cases} x\equiv c_1(\mathrm{mod}\, m_1), \\ x\equiv c_2(\mathrm{mod}\, m_2). \end{cases}$ Пусть $d = (m_1,m_2)$, $M=[m_1,m_2]$. Для каждого целого t значение $x=c_1+m_1t$ удовлетворяет первому сравнению из (S). Надо найти найти такое t, при котором x удовлетворяет второму сравнению из (S), то есть $c_1+m_1t \equiv c_2 (\mathrm{mod}\, m_2)$. Задача сведена к решению сравнения $m_1t\equiv c_2-c_1(\mathrm{mod}\ m_2)$. Если $d \nmid (c_2-c_1)$, то это сравнение (и сравнение (S)) не имеет решений. Если $d \mid (c_2- c_1)$, то сравнение имеет одно решение $t\equiv a (\mathrm{mod}\, m_2/d)$, или $t=a+(m_2/d)t_1$ при некотором целом a. Поэтому $x=c_1+m_1(a+(m_2/d)t_1) = c_1 + m_1a+(m_1m_2)/d)t_1) = a_1+Mt_1$ с $a_1 = c_1+m_1a$, откуда $x\equiv a_1 (\mathrm{mod}\, M)$ есть решение сравнения (S).

Рассмотрим систему сравнений

$$a_1x \equiv b_1 (\mathrm{mod}\, m_1), \ \ldots, \ a_sx \equiv b_s(\mathrm{mod}\, m_s). \qquad (4.10')$$

Если $(a_i,m_i)=d_i$, $d_i \nmid b_i$ для некоторого $1\leqslant i\leqslant s$, то система $(4.10')$ не имеет решений. Если $\forall i=1,\ldots,s$ целое $d_i \mid b_i$, то каждое сравнение можно решить относительно x, и система $(4.10')$ эквивалентна системе

$$x \equiv c_1 (\mathrm{mod}\, m_1/d_1), \ \ldots, \ x \equiv c_s (\mathrm{mod}\, m_s/d_s). \qquad (4.11)$$

Система (4.11) или не имеет решений, или если решения существуют, то можно найти решение системы (4.11), последовательно решая системы двух сравнений, в результате чего получим решение для (4.11), которое образует класс по модулю $[m_1/d_1,\ldots,m_s/d_s]$.

Пример. Система

$$\begin{cases} 6x\equiv 5 (\mathrm{mod}\, 7), \\ 7x\equiv 8 (\mathrm{mod}\, 9), \\ 2x\equiv 7 (\mathrm{mod}\, 15) \end{cases} \text{ эквивалентна системе } \begin{cases} x\equiv 2 (\mathrm{mod}\, 7), \\ x\equiv 5 (\mathrm{mod}\, 9), \\ x\equiv 11 (\mathrm{mod}\, 15). \end{cases}$$

Решение. $c_1=2$, $c_2=5$, $c_3=11$. Модули $m_1=7$, $m_2=9$, $m_3=15$ не являются попарно взаимно простыми:

$d=(9,15)=3\neq 1$. Тогда $x=5+9t$, $x=5+9t\equiv 2(\mathrm{mod}\, 7)$, $9t\equiv -3(\mathrm{mod}\, 7)$, $9t-7t\equiv -3(\mathrm{mod}\, 7)$, $2t\equiv -3(\mathrm{mod}\ 7)$, $2t\equiv -3+7(\mathrm{mod}\, 7)$, $2t\equiv 4(\mathrm{mod}\, 7)$, $t\equiv 2(\mathrm{mod}\, 7)$, $t=2+7y$,

$x = 5+9t = 5+9(2+7y) = 23+63y\equiv 23(\mathrm{mod}\, 63)$.

Данная система эквивалентна системе $\begin{cases} x\equiv 23(\mathrm{mod}\, 63), \\ x\equiv 11(\mathrm{mod}\, 15). \end{cases}$ Здесь $d = (63,15)=3$ и $3 \mid (23-11)$. Система совместна. Тогда $x = 23+63y \equiv 11(\mathrm{mod}\, 15)$, $63y-15y\cdot 4 \equiv -12(\mathrm{mod}\, 15)$, $3y\equiv 3(\mathrm{mod}\, 15)$, $y\equiv 1(\mathrm{mod}\, 5)$, $y=1+5z$, $x=23+63(1+5z)=86+315z$. *Ответ*. $x\equiv 86 (\mathrm{mod}\, 315)$.

4.4. Сравнения любой степени с простым модулем

Пусть p есть простое число. Будем рассматривать сравнения

$$f(x) \equiv 0 (\mathrm{mod}\, p), \ f(x)=ax^n+a_{n-1}x^{n-1}+\ldots+a_0. \qquad (4.12)$$

Теорема 1. Сравнение (4.12) эквивалентно сравнению степени не больше $p-1$.

Доказательство. Разделим $f(x)$ на x^p-x и получим $f(x) = (x^p-x)Q(x)+R(x)$. Так как теореме Ферма $x^p-x\equiv 0 (\mathrm{mod}\, p)$, то $f(x) \equiv R(x) (\mathrm{mod}\, p)$, где степень $R(x)$ не больше $p-1$.

Теорема 2. Если сравнение (4.12) имеет больше n решений, то все коэффициенты в $f(x)$ кратны p.

Доказательство. Пусть сравнение (4.12) имеет, например, $n+1$ решений. Пусть $x_1,\ldots,x_n,x_{n+1}$ есть вычеты этих решений. Тогда $f(x)$ можно представвить в виде

$$\begin{aligned} f(x) &= ax^n+a_{n-1}x^{n-1}+\ldots+a_0 = \qquad (4.13)\\ &a(x-x_1)(x-x_2)\ldots(x-x_{n-2})(x-x_{n-1})(x-x_n) + \\ &b(x-x_1)(x-x_2)\ldots(x-x_{n-2})(x-x_{n-1}) + \\ &c(x-x_1)(x-x_2)\ldots(x-x_{n-2}) + \\ &\quad\ldots \\ &k(x-x_1)(x-x_2) + \\ &l(x-x_1) + \\ &m, \end{aligned}$$

где коэффициенты $b,\ldots,l,m$ пока не определены. Производим перемножения в правой части равенства (4.13). Собираем слагаемые с одинаковыми степенями x.

Первое слагаемое есть ax^n.

Второе слагаемое $f_{n-1}(a,b)x^{n-1}$ есть результат перемножения скобок в слагаемых с a и b в (4.13). Поэтому коэффициент $f_{n-1}(a,b)$ зависит только от a,b и является суммой a и b с некоторыми коэффициентами.

Третье слагаемое $f_{n-2}(a,b,c)x^{n-2}$ есть результат перемно-

жения скобок в слагаемых с a,b,c в (4.13). Поэтому коэффициент $f_{n-2}(a,b,c)$ зависит только от a,b,c и является суммой a,b,c с некоторыми коэффициентами.

И так далее.

В результате получим равенство двух многочленов:

$$ax^n+a_{n-1}x^{n-1}+\ldots+a_0 =$$
$$ax^n + f_{n-1}(a,b)x^{n-1} + f_{n-2}(a,b,c)x^{n-2} + \ldots +$$
$$f_2(a,b,c,\ldots,k)x^2 + f_1(a,b,c,\ldots,k,l)x +$$
$$f_0(a,b,c,\ldots,k,l,m).$$

Коэффициент a известен. Коэффициент b получаем из равенства $f_{n-1}(a,b)=a_{n-1}$. Коэффициент c получаем из равенства $f_{n-2}(a,b,c)=a_{n-2}$. И так далее. Наконец, m получаем из равенства $f_0(a,b,c,\ldots,k,l,m)=a_0$.

Пусть $x=x_1$ в (4.13). Тогда $m=f(x_1)\equiv 0(\mathrm{mod}\, p)$ и m кратно p. Пусть $x=x_2$ в (4.13). Тогда $m+l(x_2-x_1)=f(x_2) \equiv 0 \pmod p$. Целые x_1,x_2 дают различные остатки при делении на p. Поэтому x_1-x_2 не делится на p. Тогда l делится на p, то есть l кратно p. Аналогично получаем, что коэффициенты $a,b,\ldots,k$ кратны p. Так как каждое $a_i=f_i$ и f_i есть сумма $a,b,\ldots,m$ с некоторыми коэффициентами, то каждое a_i кратно p.

Следствие. Сравнение (с простым модулем) порядка n имеет не больше n решений.

Теорема 3 (Вильсон). Если p есть простое число, то

$$1\cdot 2\cdot\ldots\cdot(p-1) + 1 \equiv 0 \pmod p. \qquad (4.14)$$

Доказательство. (4.14) очевидно при $p=2$. Пусть $p>2$. Сравнение $f(x) = (x-1)(x-2)\ldots(x-(p-1)) - (x^{p-1}-1) \equiv 0 \pmod p$ степени не больше $p-2$ имеет $p-1$ решений, а именно, решения с вычетами $1,2,\ldots,p-1$. По теореме 2, все коэффициенты полинома $f(x)$ кратны p, в том числе свободный член. Поэтому $1\cdot 2\cdot\ldots\cdot(p-1)+1$ делится на p.

Пример. $1\cdot 2\cdot\ldots\cdot 6 + 1 = 721 \equiv 0 (\mathrm{mod}\, 7)$.

4.5. Сравнения произвольной степени по составному модулю

Теорема. Если $m_1,m_2,\ldots,m_k$ попарно взаимно просты, то сравнение

$$f(x) \equiv 0 \pmod{m_1m_2\ldots m_k},\ f(x)=ax^n+a_1x^{n-1}+\ldots+a_n. \qquad (4.15)$$

эквивалентно системе сравнений

$$f(x)\equiv 0(\mathrm{mod}\, m_1),\ f(x)\equiv 0(\mathrm{mod}\, m_2),\ldots,f(x)\equiv 0(\mathrm{mod}\, m_k). \qquad (4.15')$$

Если $T_1,T_2,\ldots,T_k$ есть число решений сравнений (4.15′) соответственно, то число решений сравнения (4.15) $T = T_1T_2\ldots T_k$.

Доказательство. Первая часть теоремы следует из параграфа 3.2, свойства 9,10. Докажем вторую часть. Пусть сравнение $f(x) \equiv 0 \pmod{m_s}$ имеет множество B_s вычетов для T_s решений $x \equiv b_s \pmod{m_s}$, $b_s\in B_s$. Целое x удовлетворяет (4.15′) $\longleftrightarrow$ x удовлетворяет одному из $T=T_1T_2\ldots T_k$ систем сравнений

$$\begin{cases} x \equiv b_1 (\mathrm{mod}\, m_1),\ b_1 \text{ пробегает } B_1, \\ \ldots \\ x \equiv b_s (\mathrm{mod}\, m_s),\ b_k \text{ пробегает } B_k. \end{cases} \qquad (4.16)$$

T различных решений систем (4.16) есть все T различных решений систем (4.16) и потому сравнения (4.15).

Пример. Сравнение

$$f(x) \equiv 0 \pmod{35},\ f(x)=x^4+2x^3+8x+9 \qquad (4.17)$$

эквивалентно системе сравнений

$$f(x) \equiv 0 (\mathrm{mod}\, 5),\ f(x) \equiv 0 (\mathrm{mod}\, 7).$$

Первое сравнение имеет два решения: $x\equiv 1; 4 (\mathrm{mod}\, 5)$, второе имеет три решения: $x\equiv 3; 5; 6 (\mathrm{mod}\, 7)$. Сравнение (4.17) имеет $2\cdot 3=6$ решений. Чтобы их найти, надо решить шесть систем вида

$$x \equiv b_1 \pmod 5,\ x \equiv b_2 \pmod 7, \qquad (4.18)$$

где b_1 пробегает 1,4 и b_2 пробегает 3,5,6. По 4.3, $m_1m_2=5\cdot 7 = 35$, $M_1=7$, $M_2=5$. Сравнения $M_1N_1 \equiv 1 (\mathrm{mod}\, m_1)$, $M_2N_2 \equiv 1 (\mathrm{mod}\, m_2)$ есть $7N_1 \equiv 1 (\mathrm{mod}\, 5)$, $5N_2 \equiv 1 (\mathrm{mod}\, 5)$. Они удовлетворяются числами $N_1=3$, $N_2=3$. Тогда $x_0 = M_1N_1b_1 + M_2N_2b_2 = 7\cdot 3\cdot b_1 + 5\cdot 3\cdot b_2 = 21b_1+15b_2$ и решение системы (4.18) есть $x \equiv 21b_1 + 15b_2 (\mathrm{mod}\, 35)$. Возьмем (b_1,b_2) последовательно равными (1,3), (1,5),(1,6),(4,3),(4,5),(4,6) и найдем шесть значений выражения $21b_1+15b_2$, соответственно равных 66,96,111,129,159, 174. Решения для (4.17) есть $x \equiv 66;96;111;129;159;174 \pmod{35}$, или $x \equiv 31; 26; 6; 24; 19; 34 (\mathrm{mod}\, 35)$.

Замечание. По предыдущей теореме решение сравнения вида $f(x) \equiv 0 (\mathrm{mod}\, p_1^{a_1}p_2^{a_2}\ldots p_k^{a_k})$ можно свести к решению сравнений вида

$$f(x) \equiv 0 \pmod{p^a}, \qquad (4.19)$$

а решение (4.19) можно свести к решению сравнения

$$f(x) \equiv 0 \pmod{p}. \qquad (4.20)$$

Теорема. Всякое решение $x \equiv x_1 \pmod{p}$ сравнения (4.20) при $p \nmid f'(x_1)$ дает решение сравнения $x \equiv x_a \pmod{p^a}$ при некотором $x_a \equiv x_1 \pmod{p}$.

Доказательство. Всякое решение сравнения (4.19) есть решение (4.20), ибо если $f(x) \vdots p^a$, то $f(x) \vdots p$. Пусть теперь $x \equiv x_1 \pmod{p}$ есть решение (4.20). Тогда $x = x_1 + pt_1$, где $t_1 \in \mathbb{Z}$ (то есть переменная t_1 пробегает множество целых чисел). Подставим это x в сравнение $f(x) \equiv 0 \pmod{p^2}$. Разложим $f(x) = f(x_1 + pt_1)$ по формуле Тейлора в точке x_1 и получим

$$f(x_1+pt_1) = f(x_1) + \frac{f'(x_1)}{1!}pt_1 + \frac{f''(x_1)}{2!}(pt_1)^2 + \ldots + \frac{f^{(n)}(x_1)}{n!}(pt_1)^n \equiv 0 \pmod{p^2}.$$

Так как полином $f(x_1+pt_1)$ относительно pt_1 имеет целые коэффициенты, то все $\frac{f^{(k)}(x_1)}{k!}$, $k = 1, 2, \ldots, n$, есть целые числа. Удалим все кратные p^2 и получим

$$f(x_1) + pt_1 f'(x_1) \equiv 0 \pmod{p^2}$$

Так как модуль и второе слагаемое делятся на p, то $f(x_1)$ тоже делится на p. Сократим сравнение на p и получим

$$\frac{f(x_1)}{p} + t_1 f'(x_1) \equiv 0 \pmod{p}.$$

Так как $f'(x_1)$ не делится на p, то имеем одно решение $t_1 \equiv q_1 \pmod{p}$ or $t_1 = q_1 + pt_2$, откуда

$x = x_1 + pt_1 = x_1 + p(q_1 + pt_2) = x_1 + pq_1 + p^2 t_2 = x_2 + p^2 t_2$,

$x = x_2 + p^2 t_2$, $t_2 \in \mathbb{Z}$, где $x_2 = x_1 + pq_1$. Поэтому

$x \equiv x_2 \pmod{p^2}$ есть решение для $f(x) \equiv 0 \pmod{p^2}$ и

$x_2 \equiv x_1 \pmod{p}$.

Подставим это $x = x_2 + p^2 t_2$ в сравнение $f(x) \equiv 0 \pmod{p^3}$ и аналогично получим

$f(x_2) + p^2 t_2 f'(x_2) \equiv 0 \pmod{p^3}$, откуда

$$\frac{f(x_2)}{p^2} + t_2 f'(x_2) \equiv 0 \pmod{p}. \qquad (4.21)$$

$f'(x_2)$ не делится на p, ибо

$x_2 \equiv x_1 \pmod{p}$, $f(x_2) \equiv f(x_1) \pmod{p}$.

Поэтому сравнение (4.21) имеет одно решение

$t_2 \equiv q_2 \pmod{p}$, или $t_2 = q_2 + pt_3$, $t_3 \in \mathbb{Z}$,

и выражение для x имеет вид

$x = x_3 + p^3 t_3$, $t_3 \in \mathbb{Z}$, $x_3 \equiv x_2 \equiv x_1 \pmod{p}$, где $x_3 = x_2 + pq_2$.

И так далее. Наконец, будет найдено решение для (4.19):

$x = x_a + p^a t_a$ or $x \equiv x_a \pmod{p^a}$. Поэтому

$x \equiv x_a \pmod{p^a}$ есть решение для $f(x) \equiv 0 \pmod{p^a}$ и

$x_a \equiv x_1 \pmod{p}$.

Итак, каждое решение $x \equiv x_1 \pmod{p}$ для (4.20) дает решение $x \equiv x_a \pmod{p^a}$ для (4.19) и $x_a \equiv x_1 \pmod{p}$.

Следствие. Пусть $p \mid f'(a)$ и

$$x \equiv a \pmod{p^k} \text{ (то есть } x = a + p^k t \text{ при некотором } t) \qquad (4.22)$$

удовлетворяют сравнению $f(x) \equiv 0 \pmod{p^k}$.

1) если $p^{k+1} \nmid f(a)$, то (4.22) не имеет целых чисел, удовлетворяющих сравнению

$$f(x) \equiv 0 \pmod{p^{k+1}} \text{ (то есть } p^{k+1} \nmid f(x)). \qquad (4.23)$$

2) если $p^{k+1} \mid f(a)$, то все числа в (4.22) удовлетворяют (4.23).

Доказательство. По формуле Тейлора

$$f(a+p^k t) = f(a) + f'(a)p^k t + c_2(p^k t)^2 + \ldots + c_n(p^k t)^n, \qquad (4.24)$$

где n есть степень $f(x)$, $c_s = \frac{f^{(s)}(a)}{s!}$ есть целые $(2 \leqslant s \leqslant n)$. Так как $p \mid f'(a)$, то все слагаемые в (4.24), кроме первого, делятся на p^{k+1}.

1) Если $p^{k+1} \nmid f(a)$, то $p^{k+1} \nmid f(a+p^k t)$ при любом t. Поэтому (4.22) не удовлетворяет (4.23).

2) Если $p^{k+1} \mid f(a)$, то $p^{k+1} \mid f(a+p^k t)$ при любом t. Поэтому (4.22) удовлетворяет (4.23).

Замечание. Если $p \mid f'(a)$, то множество целых, которые есть решения $f(x) \equiv 0 \pmod{p}$, могут не иметь целых, удовлетворяющих сравнению $f(x) \equiv 0 \pmod{p^k}$, но могут иметь несколько классов по модулю p^{k+1}, которые есть решения для $f(x) \equiv 0 \pmod{p^{k+1}}$.

4.5.1. Алгоритм решения сравнения $f(x) \equiv 0 \pmod{p^a}$

1. Решить сравнение $f(x) \equiv 0 \pmod{p}$.

2. Для каждого решения $x \equiv x_1 \pmod{p}$ (то есть $x = x_1 + pt_1$) для $f(x) \equiv 0 \pmod{p}$ с $p \nmid f'(x_1)$ выполнить следующее.

3. Для $k = 1, 2, \ldots, a-1$:

3.1. Решить сравнение $\frac{f(x_k)}{p^k}+t_k f'(x_k)\equiv 0 \pmod p$ и получить его решение $t_k \equiv q_k \pmod p$, или $t_k = q_k + p t_{k+1}$, $t_{k+1} \in \mathbb{Z}$.

3.2. Вычислить

$x_{k+1} = x_k + p^k q_k$,

$x = x_{k+1} + p^{k+1} t_{k+1}$, где $t_{k+1} \in \mathbb{Z}$, $x_{k+1} = x_k + p^k q_k$.

4. $x \equiv x_a \pmod{p^a}$ есть решение для $f(x) \equiv 0 \pmod{p^a}$.

Пример. 1. Решить сравнение

$$f(x) \equiv 0 \pmod{3^3},\ f(x)=x^4+7x+4.$$

Решение. Сравнение $f(x)\equiv 0 \pmod 3$ имеет одно решение $x\equiv 1 \pmod 3$. Производная $f'(1)\equiv 2 \pmod 3$ не делится на 3. Находим: $x=1+3t_1$, $f(1)+3t_1 f'(1)\equiv 0 \pmod{3^2}$, $3+3t_1 \cdot 2\equiv 0 \pmod{3^2}$, $2t_1+1 \equiv 0 \pmod 3$, $t_1\equiv 1 \pmod 3$, $t_1=1+3t_2$, $x=1+3t_1=1+3(1+3t_2)=4+9t_2$, $f(4)+9t_2 f'(4) \equiv 0 \pmod{3^3}$, $18+9t_2 \cdot 2 \equiv 0 \pmod{3^3}$, $2t_2+2 \equiv 0 \pmod 3$, $t_2 \equiv 2 \pmod 3$, $t_2=2+3t_3$, $x=4+9t_2=4+9(2+3t_3)=22+27t_3$, $x\equiv 22 \pmod{3^3}$.

Ответ. $x \equiv 22 \pmod{3^3}$.

2. Решить сравнение $f(x)=x^3-2x^2-30x+41\equiv 0 \pmod{5^3}$.

Решение. $f'(x)=3x^2-4x-30$. Сравнение $f(x)\equiv 0 \pmod 5$ эквивалентно $x^3-2x^2+1\equiv 0 \pmod 5$, решение которого есть $x \equiv 1 = a = b_1 \pmod 5$.

Возьмем сравнение $f'(1)t+f(1)/5\equiv 0 \pmod 5$, то есть $-31t+2 \equiv 0 \pmod 5$, или $t\equiv 2 \pmod 5$. Возьмем $t_1=2$. Тогда $\gamma_1=b_1+p^k t_1=1+2\cdot 5=11$, $k=1$, и $x\equiv\gamma_1=11=b_2 \pmod{5^2}$ есть решение для $f(x)\equiv 0 \pmod{5^2}$. Возьмем сравнение $f'(11)t+f(11)/5^2\equiv 0 \pmod 5$, то есть $289t+32\equiv 0 \pmod 5$. Его решение есть $t\equiv 2 \pmod 5$. Пусть $t_2=2$. Тогда $\gamma_2 = b_2+p^k t_2=11+2\cdot 5^2 \pmod{5^3}=61$, $k=2$, и $x\equiv\gamma_2=61 = b_3 \pmod{5^3}$ есть решение для $f(x)\equiv 0 \pmod{5^3}$.

Ответ. $x\equiv 61 \pmod{5^3}$.

5. СРАВНЕНИЯ ВТОРОЙ СТЕПЕНИ

5.1. Квадратичные вычеты по простому модулю

Рассмотрим сравнение

$$x^n \equiv a \pmod m,\ n\geqslant 2,\ (a,m)=1. \qquad (5.1)$$

Если сравнение (5.1) имеет решение, то число a называется *вычетом степени n по модулю m*, иначе a называется *невычетом степени n по модулю m*. При $n=2$ вычеты и невычеты называются *квадратичными*, при $n=3$ *кубичными*, при $n=4$ *биквадратичными*.

Принято обозначение $x = \sqrt[n]{a} \pmod m$, если $x \in \mathbb{Z}_m$; это (дискретный) корень степени n по модулю m.

Рассмотрим случай квадратичного сравнения ($n=2$) по простому нечетному модулю p:

$$x^2 \equiv a \pmod p,\ p\geqslant 3,\ (a,p)=1. \qquad (5.2)$$

Замечание. Если x_1 удовлетворяет (5.2), то x_1^2 не делится на p, ибо a не делится на p.

Теорема 1. Если a есть квадратичный вычет по модулю p, то сравнение (5.2) имеет два решения.

Доказательство. Пусть $a\neq 0$ есть квадратичный вычет. Тогда сравнение (5.2) имеет хотя бы одно решение $x\equiv x_1 \pmod p$. Так как $(-x_1)^2 = x_1^2$, то $x\equiv -x_1 \pmod p$ есть второе решение (5.2). Сравнение $x_1 \equiv -x_1 \pmod p$ невозможно, ибо $2x_1\equiv 0 \pmod p$, $x_1\equiv 0 \pmod p$, $x_1^2\equiv 0 \pmod p$, $x_1^2 \vdots p$, что невозможно. Сравнение (5.2) не имеет других решений по теореме 2 в 4.4.

Теорема 2. Приведенная система вычетов по модулю p

$$-\frac{p-1}{2},\ldots,-2,-1,1,2,\ldots,\frac{p-1}{2} \qquad (5.3)$$

содержит $\frac{p-1}{2}$ квадратичных вычетов, сравнимых с числами

$$1^2,\ 2^2,\ \ldots,\ \left(\frac{p-1}{2}\right)^2, \qquad (5.4)$$

и $\frac{p-1}{2}$ квадратичных невычетов.

Доказательство. Все числа в (5.4) попарно несравнимы. В самом деле, если $k^2\equiv l^2 \pmod p$, $0<k<l\leqslant(p-1)/2$, то сравнение $x^2 \equiv l^2 \pmod p$ имеет четыре решения $x=-l,-k,k,l$. Поэтому числа в (5.4) лежат в различных классах вычетов с представителями из (5.4). Если одно из чисел некоторого класса есть квадратичный вычет, то все числа этого класса есть квадратичные вычеты. Поэтому $(p-1)/2$ чисел в (5.3) есть квадратичные вычеты. Все другие числа в (5.3) есть квадратичные невычеты.

Теорема 3. Если a есть квадратичный вычет по модулю p, то

$a^{\frac{p-1}{2}} \equiv 1 \pmod p$, то есть $a^{\frac{p-1}{2}} - 1 \equiv 0 \pmod p$. (5.5)

Если a есть квадратичный невычет по модулю p, то

$a^{\frac{p-1}{2}} \equiv -1 \pmod p$, то есть $a^{\frac{p-1}{2}} + 1 \equiv 0 \pmod p$. (5.6)

Доказательство. По теореме Ферма $a^{p-1} \equiv 1 \pmod p$, откуда

$$\left(a^{\frac{p-1}{2}} - 1\right)\left(a^{\frac{p-1}{2}} + 1\right) \equiv 0 \pmod p.$$

Только один из множителей делится на p, иначе их разность, равная 2, тоже делится на p. Поэтому только одно из сравнений (5.5) и (5.6) имеет место. Пусть a есть квадратичный вычет. Тогда сравнение $a \equiv x_0^2 \pmod p$ удовлетворяется при некотором $x=x_0$. Возвысим сравнение в степень $(p-1)/2$. Тогда

$a^{\frac{p-1}{2}} \equiv x_0^{p-1} \equiv 1 \pmod p$ и (5.5) верно. $(p-1)/2$ значений квадратичных вычетов a исчерпывают все решения для (5.5), ибо (5.5) не может иметь более $(p-1)/2$ решений. Поэтому квадратичные невычеты a удовлетворяют сравнению (5.6).

5.2. Символ Лежандра

Определение. Пусть целое $a \in \mathbb{Z}$, p есть нечетное простое число и $p \nmid a$. *Символ Лежандра* (символ a по p)

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{если } a \text{ есть квадратичный вычет по модулю } p, \\ -1, & \text{если } a \text{ есть квадратичный невычет по модулю } p, \end{cases}$$

Целое a есть числитель, целое p есть знаменатель символа Лежандра.

Теорема 1. $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod p$. **Доказательство** следует по теореме 3 в 5.1.

Теорема 2. Если $a \equiv a_1 \pmod p$, то $\left(\frac{a}{p}\right) = \left(\frac{a_1}{p}\right)$.

Доказательство. Два целых из одного и того же класса есть или оба квадратичные вычеты, или оба квадратичные невычеты.

Теорема 3. $\left(\frac{1}{p}\right) = 1$. **Доказательство.** Сравнение $x^2 \equiv 1 \pmod p$ имеет два решения $x=\pm 1$. Поэтому 1 есть квадратичный вычет.

Теорема 4. $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

Доказательство. Пусть $(p-1)/2$ четно. Тогда по теореме 1 символ Лежандра $\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} = 1 \pmod p$. $\left(\frac{-1}{p}\right) = 1 = (-1)^{\frac{p-1}{2}}$, ибо если $\left(\frac{-1}{p}\right) = -1$, то $-1 \equiv 1 \pmod p$, $2 \equiv 0 \pmod p$, $p \mid 2$, что невозможно.

Пусть $(p-1)/2$ нечетно. Тогда по теореме 1 символ Лежандра $\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} = -1 \pmod p$. $\left(\frac{-1}{p}\right) = -1 = (-1)^{\frac{p-1}{2}}$, ибо если $\left(\frac{-1}{p}\right) = 1$, то $1 \equiv -1 \pmod p$ и опять $p \mid 2$, что невозможно. В обоих случаях $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

Замечание. Так как $(p-1)/2$ четно, если простое $p=4m+1$ при некотором m, и нечетно, если $p=4m+3$ (случаи $p=4m, 4m+2$ невозможны в силу простоты p), то -1 есть квадратичный вычет при $p=4m+1$ и -1 есть квадратичный невычет при $p=4m+3$.

Теорема 5. $\left(\frac{ab\ldots l}{p}\right) = \left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right) \cdot \ldots \cdot \left(\frac{l}{p}\right)$.

Доказательство. По теореме 1 $\left(\frac{ab\ldots l}{p}\right) \equiv (ab\ldots l)^{\frac{p-1}{2}} =$

$a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} \ldots l^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)\ldots\left(\frac{l}{p}\right) \pmod p$.

Замечание. $\left(\frac{ab^2}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{a}{p}\right)$, ибо $\left(\frac{b}{p}\right)\left(\frac{b}{p}\right)$ равно $1 \cdot 1 = 1$ или $(-1)\cdot(-1) = 1$.

Теорема 6. Пусть $p_1=\frac{p-1}{2}$ и пусть имеем систему сравнений

$$\begin{cases} a\cdot 1 \equiv \varepsilon_1 r_1 \pmod p, \\ a\cdot 2 \equiv \varepsilon_2 r_2 \pmod p, \\ \dots \\ a\cdot x \equiv \varepsilon_x r_x \pmod p, \\ \dots \\ a\cdot p_1 \equiv \varepsilon_{p_1} r_{p_1} \pmod p, \end{cases} \tag{5.7}$$

где $\varepsilon_x r_x$ (с $r_x \geqslant 1$, $\varepsilon_x = \pm 1$) есть абсолютно наименьший вычет для ax по модулю p. Тогда

$$\left(\frac{a}{p}\right) = \varepsilon_1\varepsilon_2\dots\varepsilon_{p_1}. \tag{5.8}$$

Доказательство. Целые $a\cdot 1$, $-a\cdot 1$, $a\cdot 2$, $-a\cdot 2,\dots,a\cdot p_1$, $-a\cdot p_1$ есть приведенная система вычетов по модулю p (теорема 2 в 3.4). Их абсолютно наименьшие вычеты есть $\varepsilon_1 r_1$, $-\varepsilon_1 r_1$, $\varepsilon_2 r_2$, $-\varepsilon_2 r_2,\dots,\varepsilon_{p_1} r_{p_1}$, $-\varepsilon_{p_1} r_{p_1}$. Положительные из них $r_1, r_2, \dots, r_{p_1}$ являются перестановкой целых чисел $1,2,\dots,p_1$ (параграф 3.3). Перемножим почленно сравнения в (5.7), сократим на $1\cdot 2\cdot\dots\cdot p_1 = r_1 r_2 \dots r_{p_1}$ и получим (с помощью теоремы 1) $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} = a^{p_1} \equiv \varepsilon_1\varepsilon_2\dots\varepsilon_{p_1} \pmod p$, откуда $\left(\frac{a}{p}\right) \equiv \varepsilon_1\dots\varepsilon_{p_1} \pmod p$. Так как $\left(\frac{a}{p}\right)$ и $\varepsilon_1\dots\varepsilon_{p_1}$ равны 1 или -1 и так как $-1 \not\equiv 1 \pmod p$, то $\left(\frac{a}{p}\right) = \varepsilon_1\dots\varepsilon_{p_1}$ (оба либо 1, либо -1).

Замечание. Напомним, что $\lfloor x \rfloor$ и $\{x\}$ есть целая часть и дробная часть вещественного числа x.

Теорема 7. $\left(\frac{a}{p}\right) = (-1)^{\sum_{x=1}^{p_1} \left\lfloor \frac{2ax}{p} \right\rfloor}$, $p_1 = \frac{p-1}{2}$.

Доказательство. Пусть неотрицательное вещественное число $\frac{ax}{p} = b.c_1c_2\dots$ Тогда

$$\left\lfloor 2\cdot\frac{ax}{p} \right\rfloor = \begin{cases} 2b = 2\cdot\left\lfloor \frac{ax}{p} \right\rfloor, & \text{если } \left\{\frac{ax}{p}\right\} = 0.c_1c_2\dots < 0.5, \\ 2b+1 = 2\cdot\left\lfloor \frac{ax}{p} \right\rfloor + 1, & \text{если } \left\{\frac{ax}{p}\right\} = 0.c_1c_2\dots \geqslant 0.5. \end{cases}$$

Поэтому неотрицательное целое

$$\left\lfloor \frac{2ax}{p} \right\rfloor = \left\lfloor 2\cdot\left\lfloor \frac{ax}{p} \right\rfloor + 2\cdot\left\{\frac{ax}{p}\right\} \right\rfloor = 2\cdot\left\lfloor \frac{ax}{p} \right\rfloor + \left\lfloor 2\cdot\left\{\frac{ax}{p}\right\} \right\rfloor$$

четно или нечетно, если наименьший неотрицательный вычет числа ax меньше $p/2$ и тогда $\left\{\frac{ax}{p}\right\} < 0.5$ или больше чем $p/2$ и тогда $\left\{\frac{ax}{p}\right\} > 0.5$, то есть если $\varepsilon_x = 1$ или $\varepsilon_x = -1$. Тогда $\varepsilon_x = (-1)^{\left\lfloor \frac{2ax}{p} \right\rfloor}$ и из формулы (5.8) имеем $\left(\frac{a}{p}\right) = (-1)^{\sum_{x=1}^{p_1} \left\lfloor \frac{2ax}{p} \right\rfloor}$.

Теорема 8. Для нечетного a

$$\left(\frac{2}{p}\right)\left(\frac{a}{p}\right) = (-1)^{\sum_{x=1}^{p_1} \left\lfloor \frac{ax}{p} \right\rfloor + \frac{p^2-1}{8}}. \tag{5.9}$$

Доказательство. Целое $a+p$ четно. Тогда $\left(\frac{2a}{p}\right) = \left(\frac{2a+2p}{p}\right) = \left(\frac{2\cdot 2\cdot\frac{a+p}{2}}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{2}{p}\right)\left(\frac{\frac{a+p}{2}}{p}\right) = \left(\frac{\frac{a+p}{2}}{p}\right) = (-1)^{\sum_{x=1}^{p_1} \left\lfloor \frac{(a+p)x}{p} \right\rfloor} = (-1)^{\sum_{x=1}^{p_1} \left\lfloor \frac{ax}{p} \right\rfloor + \sum_{x=1}^{p_1} x} = (-1)^{\sum_{x=1}^{p_1} \left\lfloor \frac{ax}{p} \right\rfloor + \frac{p^2-1}{8}}$.

Теорема 9. $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

Доказательство из теоремы 8 при $a=1$.

Замечание. p можно представить как $p=8m+s$ при $s=1,3,5,7$.

Целое $\frac{p^2-1}{8} = \frac{(8m+s)^2-1}{8} = 8m^2 + 2ms + \frac{s^2-1}{8}$ четно при $s=1,7$ и нечетно при $s=3,5$. Поэтому целое 2 есть квадратичный вычет по модулю p, если p есть $8m+1$ или $8m+7$ и целое 2 есть квадратичный невычет, если p есть $8m+3$ или $8m+5$.

Теорема 10 (Закон квадратичной взаимности).

$\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\left(\frac{p}{q}\right)$ для нечетных простых p и q.

Доказательство. Так как $\frac{p-1}{2}\cdot\frac{q-1}{2}$ нечетно при p и q лишь вида $4m+3$ и четно для p или q вида $4m+1$, то теорему можно сформулировать следующим образом.

Если p и q имеют вид $4m+3$, то $\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$.

Если p or q имеют вид $4m+1$, то $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$.

По теореме 8 формула (5.9) имеет вид

$$\left(\frac{a}{p}\right) = (-1)^{\sum_{x=1}^{p_1}\left[\frac{ax}{p}\right]} \qquad (5.10)$$

Пусть $q_1=\frac{q-1}{2}$. Пусть множество пар целых чисел есть

$$\{qx,py:\ x=1,2,\ldots,p_1;\ y=1,2,\ldots,q_1\}.$$

$qx=py$ невозможно, иначе py кратно q, что невозможно, ибо $0<y<q$ и $(p,q)=(y,q)=1$. Поэтому можно положить $p_1q_1=S_1+S_2$, где S_1 есть число пар с $qx<py$, и S_2 есть число пар с $py<qx$.

S_1 есть число пар при $x < \frac{p}{q}y$. При фиксированном y можно взять $x=1,2,\ldots,\left\lfloor\frac{p}{q}y\right\rfloor$, ибо $\frac{p}{q}y \leqslant \frac{p}{q}q_1 < \frac{p}{2}$ влечет $\left\lfloor\frac{p}{q}y\right\rfloor<p_1$.

Следовательно, $S_1 = \sum_{y=1}^{q_1}\left\lfloor\frac{p}{q}y\right\rfloor$. Аналогично $S_2 = \sum_{x=1}^{p_1}\left\lfloor\frac{q}{p}x\right\rfloor$.

Тогда равенство (5.10) дает $\left(\frac{p}{q}\right) = (-1)^{S_1}$, $\left(\frac{q}{p}\right) = (-1)^{S_2}$. Поэтому $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{S_1+S_2} = (-1)^{p_1q_1}$. Так как $\left(\frac{p}{q}\right)$ равно 1 или -1, то $\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\cdot\left(\frac{p}{q}\right)$.

5.3. Символ Якоби

Определение. Пусть $P>1$ есть нечетное число, $P=p_1p_2\ldots p_r$ есть факторизация P на простые множители (некоторые из них могут повторяться) и $(a,P)=1$. Тогда *символ Якоби*

$$\left(\frac{a}{P}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right).$$

Покажем, что символ Якоби свойства, аналогичные свойствам символа Лежандра.

Теорема 1. Если $a\equiv a_1 \pmod P$, то $\left(\frac{a}{P}\right) = \left(\frac{a_1}{P}\right)$.

Доказательство. $\left(\frac{a}{P}\right)=\left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right)=\left(\frac{a_1}{p_1}\right)\left(\frac{a_1}{p_2}\right)\cdots\left(\frac{a_1}{p_r}\right)=\left(\frac{a_1}{P}\right)$.

Теорема 2. $\left(\frac{1}{P}\right)=1$. **Доказательство**. $\left(\frac{1}{P}\right)=\left(\frac{1}{p_1}\right)\left(\frac{1}{p_2}\right)\cdots\left(\frac{1}{p_r}\right)=1$.

Теорема 3. $\left(\frac{-1}{P}\right) = (-1)^{\frac{P-1}{2}}$. **Доказательство**.

$$\left(\frac{-1}{P}\right) = \left(\frac{-1}{p_1}\right)\left(\frac{-1}{p_2}\right)\cdots\left(\frac{-1}{p_r}\right) = (-1)^{\frac{p_1-1}{2}+\frac{p_2-1}{2}+\ldots+\frac{p_r-1}{2}}. \qquad (5.11)$$

Но $\frac{P-1}{2} = \frac{p_1p_2\ldots p_r-1}{2} = \frac{\left(1+2\frac{p_1-1}{2}\right)\left(1+2\frac{p_2-1}{2}\right)\cdots\left(1+2\frac{p_r-1}{2}\right)}{2} =$

$\frac{p_1-1}{2} + \frac{p_2-1}{2} + \ldots + \frac{p_r-1}{2} + 2N$ при некотором N. Тогда по (5.11)

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

Теорема 4. $\left(\frac{ab\ldots l}{P}\right) = \left(\frac{a}{P}\right)\left(\frac{b}{P}\right)\ldots\left(\frac{l}{P}\right)$.

Доказательство. $\left(\frac{ab\ldots l}{P}\right) = \left(\frac{ab\ldots l}{p_1}\right)\left(\frac{ab\ldots l}{p_2}\right)\ldots\left(\frac{ab\ldots l}{p_r}\right) =$

$$\left(\frac{a}{p_1}\right)\ldots\left(\frac{l}{p_1}\right)\cdot\left(\frac{a}{p_2}\right)\ldots\left(\frac{l}{p_2}\right)\cdot\ldots\cdot\left(\frac{a}{p_r}\right)\ldots\left(\frac{l}{p_r}\right) =$$

$$\left(\frac{a}{p_1}\right)\ldots\left(\frac{a}{p_r}\right)\cdot\left(\frac{b}{p_1}\right)\ldots\left(\frac{b}{p_r}\right)\cdot\ldots\cdot\left(\frac{l}{p_1}\right)\ldots\left(\frac{l}{p_r}\right) =$$

$$\left(\frac{a}{P}\right)\left(\frac{b}{P}\right)\ldots\left(\frac{l}{P}\right).$$

Замечание. $\left(\frac{ab^2}{P}\right) = \left(\frac{a}{P}\right)$.

Теорема 5. $\left(\frac{2}{P}\right) = (-1)^{\frac{P^2-1}{8}}$. **Доказательство.**

$$\left(\frac{2}{P}\right) = \left(\frac{2}{p_1}\right)\left(\frac{2}{p_2}\right)\ldots\left(\frac{2}{p_r}\right) = (-1)^{\frac{p_1^2-1}{8}+\frac{p_2^2-1}{8}+\ldots+\frac{p_r^2-1}{8}}. \qquad (5.12)$$

Но $\frac{P^2-1}{2} = \frac{p_1^2p_2^2\ldots p_r^2-1}{2} =$

$$\frac{\left(1+8\cdot\frac{p_1^2-1}{8}\right)\left(1+8\cdot\frac{p_2^2-1}{8}\right)\ldots\left(1+8\cdot\frac{p_r^2-1}{8}\right)}{2} =$$

$\frac{p_1^2-1}{2} + \frac{p_2^2-1}{2} + \ldots + \frac{p_r^2-1}{2} + 2N$ при некотором N. Тогда по (5.12)

$$\left(\frac{2}{P}\right) = (-1)^{\frac{P_1^2-1}{8}+2N} = (-1)^{\frac{P^2-1}{8}}.$$

Теорема 6 (Закон квадратичной взаимности). Если P и Q есть положительные нечетные взаимно простые числа, то

$$\left(\frac{Q}{P}\right) = (-1)^{\frac{P-1}{2}\cdot\frac{Q-1}{2}}\left(\frac{P}{Q}\right).$$

Доказательство. Пусть $Q=q_1q_2\ldots q_s$ есть факторизация Q на простые множители (некоторые из них могут повторяться).

Тогда $\left(\frac{Q}{P}\right) = \left(\frac{Q}{p_1}\right)\left(\frac{Q}{p_2}\right)\ldots\left(\frac{Q}{p_r}\right) = \prod_{u=1}^{r}\prod_{v=1}^{s}\left(\frac{q_v}{p_u}\right) =$

$$(-1)^{\sum_{u=1}^{r}\sum_{v=1}^{s}\frac{p_u-1}{2}\cdot\frac{q_v-1}{2}}\prod_{u=1}^{r}\prod_{v=1}^{s}\left(\frac{p_u}{q_v}\right) =$$

$$(-1)^{\left(\sum_{u=1}^{r}\frac{p_u-1}{2}\right)\cdot\left(\sum_{v=1}^{s}\frac{q_v-1}{2}\right)}\left(\frac{P}{Q}\right).$$

По аналогии с теоремой 3 можно получить, что

$$\frac{P-1}{2} = \sum_{u=1}^{r}\frac{p_u-1}{2} + 2N, \quad \frac{Q-1}{2} = \sum_{v=1}^{s}\frac{q_v-1}{2} + 2N_1 \text{ при некоторых } N, N_1.$$

Тогда $\left(\frac{Q}{P}\right) = (-1)^{\frac{P-1}{2}\cdot\frac{Q-1}{2}}\left(\frac{P}{Q}\right)$.

5.3.1. Алгоритм вычисления символа Якоби (и символа Лежандра)

JACOBI(a,n)

ВХОД. Нечетное число $n\geqslant 3$ и число a, $0\leqslant a<n$.

ВЫХОД. Символ Якоби $\left(\frac{a}{n}\right)$ (и следовательно символ Лежандра

при простом n).

1. Если $a = 0$, то возвратить 0.
2. Если $a = 1$, то возвратить 1.
3. Записать a как $a = 2^e a_1$, где a_1 нечетно.
4. Если e четно, то $s := 1$. В противном случае $s := 1$, если $n \equiv 1$ или $7 \pmod 8$, или $s := -1$, если $n \equiv 3$ или $5 \pmod 8$.
5. Если $n \equiv 3 \pmod 4$ и $a_1 \equiv 3 \pmod 4$, то $s := -s$.
6. $n_1 := n \pmod{a_1}$.
7. Если $a_1 = 1$, то вернуть s; в противном случае вернуть $s \cdot \mathrm{JACOBI}(n_1, a_1)$.

Пример. Исследовать на разрешимость сравнение $x^2 \equiv 219 \pmod{383}$. *Решение*. Символ Якоби $\left(\frac{Q}{P}\right) =$

$$\left(\frac{219}{383}\right) = (-1)^{\frac{P-1}{2}\cdot\frac{Q-1}{2}}\left(\frac{383}{219}\right) = -\left(\frac{383-219}{219}\right) = -\left(\frac{164}{219}\right) =$$

$$-\left(\frac{2\cdot 2\cdot 41}{219}\right) = -\left(\frac{2}{219}\right)\left(\frac{2}{219}\right)\left(\frac{41}{219}\right) = -\left(\frac{41}{219}\right) = -\left(\frac{219}{41}\right) = -\left(\frac{14}{41}\right) =$$

$$-\left(\frac{2}{41}\right)\left(\frac{7}{41}\right) = -(-1)^{\frac{41-1}{2}}\left(\frac{7}{41}\right) = -\left(\frac{7}{41}\right) = -\left(\frac{41}{7}\right) = -\left(\frac{6}{7}\right) = -\left(\frac{-1}{7}\right) =$$

$$-(-1)^{\frac{7-1}{2}} = 1.$$

Следовательно, данное сравнение разрешимо и имеет два решения.

5.4. Квадратичные вычеты по составному модулю

Рассмотрим сравнения вида

$$x^2 \equiv a \pmod{p^\alpha}, \tag{5.13}$$

где $\alpha > 0$, $(a,p)=1$, $p>2$ есть нечетное простое число.

Теорема 1. Если сравнение

$$x^2 \equiv a \pmod p \tag{5.14}$$

имеет решение (то есть a есть квадратичный вычет по модулю p), то сравнение (5.13) имеет два решения. Если сравнение (5.14) не имеет решений (то есть a есть квадратичный невычет по модулю p), то сравнение (5.13) не имеет решений.

Доказательство. Пусть $f(x)=x^2-a$, $f'(x)=2x$. Если $x \equiv x_1 \pmod p$ есть решение сравнения (5.14), то условие $(a,p)=1$ влечет $(x_1,p)=1$. Так как простое p нечетно, то $(2x_1,p)=1$. Поэтому $f'(x_1)$ не делится на p. Применим рассуждения из 4.5 и получим, что каждое решение сравнения (5.14) даст одно решение сравнения (5.13).

Теорема 2. Если сравнение

$$x^2 \equiv a \pmod{2^\alpha},\ \alpha \geqslant 1,\ (a,2)=1, \tag{5.15}$$

разрешимо, то

$$\begin{aligned} &a \equiv 1 \pmod 2 \text{ при } \alpha=1,\\ &a \equiv 1 \pmod 4 \text{ при } \alpha=2,\\ &a \equiv 1 \pmod 8 \text{ при } \alpha \geqslant 3. \end{aligned} \tag{5.16}$$

Доказательство. Производная $f'(x_1)=2x_1$ делится на $p=2$. Поэтому рассуждения предыдущей теоремы неприменимы.

Пусть сравнение (5.15) разрешимо (имеет решение x). Тогда $(a,2)=1$ влечет $(x,2)=1$. Следовательно, x^2-1 делится на 8 (теорема 8 в 5.2) и поэтому x^2-1 на 4 и на 2. Сравнение (5.15) эквивалентно $(x^2-1)+1 \equiv a \pmod{2^\alpha}$ и $(x^2-1) \equiv a-1 \pmod{2^\alpha}$.

Пусть $\alpha=1$. Тогда $(x^2-1) \equiv a-1 \pmod 2$. Целые x^2-1 и 2 делятся на 2. Тогда $a-1$ тоже делится на 2. Поэтому $a-1 \equiv 0 \pmod 2$ и $a \equiv 1 \pmod 2$.

Пусть $\alpha=2$. Тогда $(x^2-1) \equiv a-1 \pmod 4$. Целые x^2-1 и 4 делятся на 4. Тогда $a-1$ тоже делится на 4. Поэтому $a-1 \equiv 0 \pmod 4$ и $a \equiv 1 \pmod 4$.

Пусть $\alpha \geqslant 3$. Тогда $(x^2-1) \equiv a-1 \pmod{2^\alpha}$. Целые x^2-1 и 2^α делятся на 8. Тогда $a-1$ тоже делится на 8. Поэтому $a-1 \equiv 0 \pmod 8$ и $a \equiv 1 \pmod 8$.

Получили, что разрешимость сравнения (5.15) влечет сравнения (5.16).

Следствие. 1. Если одно из условий (5.16) не верно, то сравнение (5.15) не имеет решений.

2. $x^2 \equiv a \pmod{2^\alpha}$ эквивалентно $\begin{cases} x^2 \equiv 1 \pmod{2^\alpha}, \\ 1 \equiv a \pmod{2^\alpha}, \end{cases}$ и

$$x^2-1 \equiv 0 \pmod{2^\alpha}. \tag{5.16'}$$

Теорема 3. Пусть условия (5.16) выполняются. Тогда сравнение (5.15) имеет один, два, четыре решения при $\alpha=1$, $\alpha=2$, $\alpha \geqslant 3$ соответственно.

Доказательство. Сравнение $x^2 \equiv a \pmod{2^\alpha}$ эквивалентно $x^2 \equiv 1 \pmod{2^\alpha}$ и потому $x^2-1 \equiv 0 \pmod{2^\alpha}$. Лишь нечетные числа x могут удовлетворять сравнениям (5.16′) и (5.15).

α=1. Только нечетное целое 1 полной системы вычетов 0,1 по модулю 2 удовлетворяет (5.16′) и $x \equiv 1 \pmod 2$ есть единственное решение (5.16′) и потому (5.15).

α=2. Нечетные целые 1,3 полной системы вычетов 0,1,2,3 по модулю 4 удовлетворяют (5.16′), и $x \equiv 1,3 \pmod 4$ есть два решения (5.16′) и потому (5.15).

α=3. Нечетные целые 1,3,5,7 полной системы вычетов 0, 1,3,...,7 по модулю 8 удовлетворяют (5.16′), и $x \equiv 1,3,5,7 \pmod 8$ есть четыре решения (5.16′) и потому (5.15).

$\alpha \geqslant 4$. Нечетные целые можно представить в виде двух арифметических прогрессий:

$$x = \pm(1+4t_3),\ t_3=0,\pm1,\pm2,\ldots \qquad (5.17)$$

Заметим, что $1+4t_3 \equiv 1 \pmod 4$, $-1-4t_3 \equiv -1 \equiv 3 \pmod 4$.

Заметим тоже, что все решения (5.15) при некотором $\alpha \geqslant 4$ удовлетворяют (5.15) при $\alpha-1$ тоже.

Рассмотрим сравнение $x^2 \equiv a \pmod{2^4}$ и найдем нечетные целые (5.17), которые удовлетворяют этому сравнению. Тогда $(1+4t_3)^2 \equiv a \pmod{16}$, откуда $1+8t_3+16t_3^2 \equiv a \pmod{16}$. Тогда $8t_3 \equiv a-1 \pmod{16}$, откуда $t_3 \equiv \frac{a-1}{8} \pmod 2$, $t_3=t_3'+2t_4$ при $t_3' \in \{0,1\}$.

$x = \pm(1+4t_3)=\pm(1+4(t_3'+2t_4)) = \pm(x_4+8t_4)$ при некотором x_4.

Найдем целые числа этого вида, удовлетворяющие сравнению $x^2 \equiv a \pmod{2^5}$. Имеем $(x_4+8t_4)^2 \equiv a \pmod{32}$, $t_4=t_4'+2t_5$, $x = \pm(x_5+16t_5)$. И так далее. Наконец, любое $x = \pm(x_c+2^{c-1}t_c$ удовлетворяет (5.15) при всяком $c \geqslant 4$. Эти x дают четыре различных решения сравнения (5.15):

$x \equiv x_\alpha,\ x_\alpha+2^{\alpha-1},\ -x_\alpha,\ -x_\alpha-2^{\alpha-1} \pmod{2^\alpha}$.

Пример. Решить сравнение

$$x^2 \equiv 57 \pmod{64}. \qquad (5.18)$$

Так как $57 \equiv 1 \pmod{64}$, то сравнение (5.18) имеет четыре решения. Представляя x в виде $x=\pm(1+4t_3)$, находим

$(1+4t_3)^2 \equiv 57 \pmod{16}$, $1+8t_3+16t_3^2 \equiv 57 \pmod{16}$,
$8t_3 \equiv 56 \pmod{16}$, $t_3 \equiv 7 \pmod 2$, $t_3 \equiv 1 \bmod 2$), $t_3=1+2t_4$,

$x=\pm(1+4t_3)=\pm(1+4(1+2t_4))=\pm(5+8t_4)$,

$(5+8t_4)^2 \equiv 57 \pmod{32}$, $25+80t_4+64t_4^2 \equiv 57 \pmod{32}$,
$5 \cdot 16t_4 \equiv 32 \pmod{32}$, $5 \cdot 16t_4 \equiv 0 \pmod{32}$, $t_4 \equiv 0 \pmod{32}$,
$t_4 \equiv 2t_5$, $x=\pm(5+8t_4)=\pm(5+16t_5)$

$(5+16t_5)^2 \equiv 57 \pmod{64}$, $5 \cdot 32t_5 \equiv 32 \pmod{64}$, $t_5 \equiv 1 \pmod 2$,
$t_9=1+2t_6$, $x=\pm(5+16t_5)=\pm(5+16(1+2t_6))=\pm(21+32t_6)$.

Четыре решения (5.18) есть $x \equiv \pm21, \pm53 \pmod 4$ при $t_6=0,1$.

Теорема 4. Если сравнение

$$x^2 \equiv a \pmod m,\ m=2^\alpha p_1^{\alpha_1}\ldots p_k^{\alpha_k},\ (a,m)=1, \qquad (5.19)$$

разрешимо, то

$$\begin{aligned} &a \equiv 1 \pmod 2 \text{ при } \alpha=1,\\ &a \equiv 1 \pmod 4 \text{ при } \alpha=2,\\ &a \equiv 1 \pmod 8 \text{ при } \alpha \geqslant 3,\\ &\left(\frac{a}{p_1}\right)=1, \left(\frac{a}{p_2}\right)=1, \ldots, \left(\frac{a}{p_k}\right)=1. \end{aligned} \qquad (5.20)$$

и число решений есть:

2^k при $\alpha=0$ и $\alpha=1$; 2^{k+1} при $\alpha=2$; 2^{k+2} при $\alpha \geqslant 3$.

Доказательство следует из теорем 1–4 этого параграфа и теорем параграфа 4.5.

Следствие. Если одно из условий (5.20) не верно, то сравнение (5.19) неразрешимо (не имеет решений).

Замечание. Факт разрешимости или неразрешимости сравнения $x^2 \equiv a \pmod m$ установивается легко. Именно, если символ Якоби $\left(\frac{a}{m}\right) = \begin{cases} 1, & \text{то сравнение } x^2 \equiv a \pmod m \text{ разрешимо,} \\ -1, & \text{то сравнение } x^2 \equiv a \pmod m \text{ не разрешимо.} \end{cases}$

Не известен алгоритм с полиномиальной временной сложностью решения сравнения $x^2 \equiv a \pmod m$. Есть предположение, что такой алгоритм не существует. Переборный алгоритм при больших m (с длиной десятиричной записи m порядка 120 цифр) практически неосуществим.

6. ПРИМИТИВНЫЕ КОРНИ И ИНДЕКСЫ

6.1. Экспонента, примитивные корни, индексы

Пусть $(a,m)=1$. Существуют положительные целые γ, например

$\varphi(m)$, для которых $a^\gamma \equiv 1 \pmod m$.

Определение. Наименьшее положительное δ, для которого $a^\delta \equiv 1 \pmod m$, называется *экспонентой, которой a принадлежит по модулю m* (обозначение: $a \in \exp\delta \pmod m$), или a имеет экспоненту δ по модулю m (обозначение: $E_m(a)=\delta$ или $E(a)=\delta$, если m известно из контекста).

Теорема 1. Если целое $a \in \exp\delta \pmod m$, то целые $1=a^0$, $a^1, a^2, \ldots, a^{\delta-1}$ не сравнимы по модулю m и взаимно просты с модулем m.

Доказательство. Пусть $a^l \equiv a^k \pmod m$, $0 \leqslant k < l < \delta$. Если $(e,m) = 1$, то обе части в $e \equiv f \pmod m$ можно разделить на од(e,f). Поэтому $a^{l-k} \equiv 1 \pmod m$, $0 < l-k < \delta$, что невозможно.

Так как $(a,m)=1$, то $(a^i,m)=1$ $\forall i \in \mathbb{N}_+$ (по параграфу 1.4).

Теорема 2. Если целое $a \in \exp\delta \pmod m$, то

$a^l \equiv a^k \pmod m \longleftrightarrow l \equiv k \pmod\delta$.

Доказательство. Пусть $a^\delta \equiv 1 \pmod m$.

Необходимость. Пусть $a^l \equiv a^k \pmod m$. Пусть r,s есть наименьшие неотрицательные вычеты целых l,k по модулю δ и $0<s \leqslant r<\delta$. Тогда $<r-s<\delta$, $l=\delta u+r$, $k=\delta v+s$ при некотором u,v. Так как $a^\delta \equiv 1 \pmod m$ и $a^l \equiv a^k \pmod m$, то

$(a^\delta)^u \equiv 1^u \pmod m$, $(a^\delta)^v \equiv 1^v \pmod m$,

$a^{\delta u} a^r \equiv 1 \cdot a^r \pmod m$, $a^{\delta v} a^s \equiv 1 \cdot a^s \pmod m$;

$a^{\delta u+r} \equiv a^r \pmod m$, $a^{\delta v+s} \equiv a^s \pmod m$;

$a^l \equiv a^r \pmod m$, $a^k \equiv a^s \pmod m$;

$a^r \equiv a^s \pmod m$, $a^{r-s} \equiv 1 \pmod m$;

$r=s$, иначе $0<r-s<\delta$ и $a^{r-s} \equiv 1 \pmod m$. Противоречие с $a \in \exp\delta \pmod m$. Тогда $l=\delta u+r$, $k=\delta v+r$, $l \equiv k \pmod\delta$.

Достаточность. Так как $l \equiv k \pmod\delta$, то $l=\delta u+r$, $k=\delta v+r$, $0 \leqslant r<\delta$, при некоторых u,v. Так как $a^\delta \equiv 1 \pmod m$, то

$(a^\delta)^u \equiv 1^u \pmod m$, $(a^\delta)^v \equiv 1^v \pmod m$;

$a^{\delta u} a^r \equiv 1 \cdot a^r \pmod m$, $a^{\delta v} a^r \equiv 1 \cdot a^r \pmod m$;

$a^{\delta u+r} \equiv a^r \pmod m$, $a^{\delta v+r} \equiv a^r \pmod m$;

$a^{\delta u+r} \equiv a^{\delta v+r} \equiv a^r \pmod m$, $a^l \equiv a^k \pmod m$.

Замечание. 1. Если $k=0$, то

$a^l \equiv a^0 = 1 \pmod m \longleftrightarrow l \equiv 0 \pmod\delta \longleftrightarrow \delta \mid l$.

2. Так как $a^{\varphi(m)} \equiv 1 \pmod m$, то $\delta \mid \varphi(m)$.

Определение. Целое a, принадлежащее экспоненте $\varphi(m)$ по модулю m, называется *примитивным корнем по модулю m*.

Замечание. Целые $a^0, a^1, \ldots, a^{\varphi(m)-1}$ образуют приведенную систему вычетов по модулю m.

Утверждение. Пусть g есть примитивный корень по модулю m.

Тогда 1) по модулю m для числа g его экспонента $\delta=\varphi(m)$,

2) $g^l \equiv g^k \pmod m \longleftrightarrow l \equiv k \pmod{\varphi(m)}$,

3) $g^l \equiv 1 \pmod m \longleftrightarrow l \equiv 0 \pmod{\varphi(m)} \longleftrightarrow \varphi(m) \mid l$.

Замечание. В теореме 3 в 6.7 мы покажем, что примитивные корни по модулю m существуют лишь при $m=2,4,p^\alpha,2p^\alpha$, где простое p нечетно и целое $\alpha \geqslant 1$.

Теорема 3. По модулю m: $E(a^s)=E(a) \longleftrightarrow (s,E(a))=1$.

Доказательство. Пусть $(s,E(a))=1$ и $y=E(a^s)$. Тогда y есть наименьшее положительное целое, для которого $(a^s)^y = a^{sy} \equiv 1 \pmod m$. Тогда $E(a) \mid sy$. Так как $(s,E(a))=1$, то $E(a) \mid y$. Так как y есть наименьшее целое с этим свойством, то $y=E(a)$, откуда $E(a^s)=E(a)$.

Пусть $(s,E(a))=d \neq 1$. Тогда $E(a)/d$ и s/d есть целые и $(a^s)^{E(a)/d} = (a^{E(a)})^{s/d} \equiv 1 \pmod m$, откуда $E(a^s) \leqslant E(a)/d < E(a)$. Поэтому $(s,E(a)) \neq 1$ влечет $E(a^s) \neq E(a)$.

Теорема 4. Если $a \equiv b \pmod m$, то $\exp a \pmod m = \exp b \pmod m$.

Доказательство. Пусть $E(a) = \exp a \pmod m$. Так как $a \equiv b \pmod m$, то $a^s \equiv b^s \pmod m$ для всякого натурального s, в том числе для $E(a)$. Так как $a^{E(a)} \equiv 1 \pmod m$ и $a^r \not\equiv 1 \pmod m$ для всякого r, $1 \leqslant r \leqslant E(a)$, то $b^{E(a)} \equiv 1 \pmod m$ и $b^r \not\equiv 1 \pmod m$ для $1 \leqslant r \leqslant E(a)$, откуда $E(b) = E(a)$.

Замечание. Все целые одного и того же класса вычетов $\bar a$ по модулю m принадлежат одной и той же экспоненте по модулю m.

Теорема 5. Если $E(a)=k$ по модулю m, то классы $\bar a, \bar a^2, \ldots, \bar a^k$ есть различные решения сравнения $x^k \equiv 1 \pmod m$.

Доказательство. Если $E(a)=k$, то $a^k \equiv 1 \pmod m$ и при всяком $s \geqslant 0$ $(a^k)^s = (a^s)^k \equiv 1 \pmod m$. Поэтому все целые a^s удовлетворяют сравнению $x^k \equiv 1 \pmod m$ и все классы $\bar a, \bar a^2, \ldots, \bar a^k$ есть различные решения сравнения $x^k \equiv 1 \pmod m$.

Замечание. Если $E(a)=k$ по простому модулю p, то классы $\bar a, \bar a^2, \ldots, \bar a^k$ исчерпывают все решения сравнения $x^k \equiv 1 \pmod p$, ибо сравнение степени k имеет не более k решений.

6.1.1. Число классов вычетов данной экспоненты

Рассмотрим взаимно простые с модулем m классы вычетов и пусть $\psi(k)$ есть число классов по модулю m, принадлежащих экспоненте k. Известно, что $k \mid \varphi(m)$. Если $k \nmid \varphi(m)$, то $\psi(k)=0$, ибо нет классов, принадлежащих экспоненте k.

Пример. 1. $m=11$. Тогда $\varphi(11)=10$. Возможные значения экс-

поненты находятся среди делителей 10, то есть среди 1,2,5, 10. Построим таблицу значений $E(a)$.

a	1	2	3	4	5	6	7	8	9	10
$k=E(a)$	1	10	5	5	5	10	10	10	5	2

Ряд 1 состоит из представителей всех классов, взаимно простых с модулем. Ряд 2 состоит из соответствующих значений экспонент $E(a)$. Тогда $\psi(1)=1$, $\psi(2)=1$, $\psi(5)=4$, $\psi(10)=4$.

2. $m=20$. Тогда $\varphi(20)=8$. Возможные значения экспонент есть 1,2,4,8. Построим таблицу значений $E(a)$.

a	1	2	7	9	11	13	17	19
$E(a)$	1	4	4	2	2	4	4	2

$\psi(1)=1$, $\psi(2)=3$, $\psi(4)=4$, $\psi(8)=0$.

Теорема 6. $\sum_{k\mid\varphi(m)} \psi(k) = \varphi(m)$.

Доказательство. Пусть $k_1=1,\ldots,k_s=\varphi(m)$ есть все положительные делители $\varphi(m)$. Существует $\varphi(m)$ взаимно простых по модулю m классов вычетов и каждое число из этих классов имеет экспонентой одно из чисел $k_1,k_1,\ldots,k_s$. Число классов, принадлежащих экспоненте k_i равно $\psi(k_i)$. Сумма $\psi(k_1)+\psi(k_2)+\ldots+\psi(k_s)$ равна числу $\varphi(m)$ всех классов, взаимно простых с m. Тогда $\sum_{k\mid\varphi(m)} \psi(k) = \psi(k_1)+\ldots+\psi(k_s) = \varphi(m)$.

Теорема 7. По простому модулю p, для всякого целого $k\geqslant 1$ верно неравенство $\psi(k) \leqslant \varphi(k)$.

Доказательство. Если $\psi(k)=0$, то $\psi(k)<\varphi(k)$. Если $\psi(k)>0$, то есть класс вычетов $\bar{a}$ с $E(a)=k$ существует, то классы

$$\bar{a},\bar{a}^2,\ldots,\bar{a}^k \tag{6.1}$$

исчерпывают все решения сравнения

$$x^k \equiv 1 \pmod p. \tag{6.2}$$

Ранее было $E(a^s) = E(a) \longleftrightarrow (s,k)=1$. The число таких s $(1\leqslant s\leqslant k)$ равно to $\varphi(k)$. Поэтому число классов (6.1) есть $\varphi(k)$. С другой стороны, всякий класс C с экспонентой $E(C)=k$ удовлетворяет (6.2). Поэтому этот класс есть в (6.1). Поэтому число $\psi(k)$ классов с экспонентой k равно $\varphi(k)$, откуда $\psi(k) = \varphi(k)$ и $\psi(k)=\varphi(k)$. Теорема доказана.

Следствие. По простому модулю: $\psi(k)=0$ либо $\psi(k)=\varphi(k)$.

Теорема 8. Если p есть простой модуль и k делит $p-1$, то

$\psi(k) = \varphi(k)$.

Доказательство. Так как $\sum_{k\mid\varphi(m)} \varphi(k)=\varphi(m)$, $\sum_{k\mid\varphi(m)} \psi(k)=\varphi(m)$, $\varphi(p)=p-1$, то $\sum_{k\mid p-1} \varphi(k)=p-1$, $\sum_{k\mid p-1} \psi(k)=\varphi(m)$. Вычтем второе равенство из первого. Тогда $\sum_{k\mid p-1} (\varphi(k)-\psi(k))=0$. Так как $\varphi(k)-\psi(k)\geqslant 0$, то $\varphi(k)-\psi(k)=0$, $\varphi(k)=\psi(k)$ для всякого $k\mid(p-1)$.

6.1.2. Индексы (дискретные логарифмы)

Определение. Пусть $(a,m)=1$, $(b,m)=1$. Число s есть *индекс (дискретный логарифм) числа a по основанию b по модулю m*, (обозначение: $s=\operatorname{ind}_b a \pmod m$), если $b^s\equiv a \pmod m$.

Замечание. 1. $s=\operatorname{ind}_b a \pmod m \longleftrightarrow b^s\equiv a \pmod m$.

2. Пишут также $s=\operatorname{ind}_b a$, $s=\operatorname{ind} a$, если m и b известны из контекста.

3. По определению индекса: $b^{\operatorname{ind}_b a} \equiv a \pmod m$.

4. Пусть $s=\operatorname{ind}_b a$, $b\equiv b_1 \pmod m$, $a\equiv a_1 \pmod m$. Тогда $b^s\equiv b_1^s \pmod m$, $b_1^s\equiv a_1 \pmod m$, $\bar{b}^s=\bar{a} \pmod m$.

5. Если b не есть примитивный корень по модулю m, то среди степеней a имеется только $k<\varphi(m)$ различных, и числа, принадлежащие остальным $\varphi(m)-k$ классам вычетов по модулю m, индексов не имеют.

6. Если g есть примитивный корень по модулю m, то любое число, взаимно простое с модулем, имеет бесконечное множество индексов.

7. Пусть g есть примитивный корень по модулю m и $(a,m)=1$, $(b,m)=1,\ldots,(l,m)=1$. Тогда

$$\operatorname{ind}_g(ab\ldots l) \equiv \operatorname{ind}_g a + \operatorname{ind}_g b + \operatorname{ind}_g l \pmod m.$$

В самом деле,

$a \equiv g^{\operatorname{ind} a} \pmod m$, $b \equiv g^{\operatorname{ind} b} \pmod m,\ldots,$
$l \equiv g^{\operatorname{ind} l} \pmod m$. Перемножим и получим
$ab\ldots l \equiv g^{\operatorname{ind} a+\operatorname{ind} b+\ldots+\operatorname{ind} l} \pmod m$,

откуда следует нужное сравнение.

Следствие. $\operatorname{ind}_g a^n \equiv n \operatorname{ind}_g a \pmod{\varphi(m)}$.

Заметим, что $\operatorname{ind} a + \operatorname{ind} b + \ldots + \operatorname{ind} l$ есть один из индексов произведения $ab\ldots l$.

6.2. Примитивные корни по модулям p^α и $2p^\alpha$

Пусть $p\geqslant 3$ есть нечетное простое число и $\alpha\geqslant 1$. Покажем су-

ществование примитивных кореней по модулям p^α и $2p^\alpha$.

Теорема 1. Если целое $x \in \exp ab \pmod m$, то целое $x^a \in \exp b \pmod m$.

Доказательство. Пусть x^a принадлежат экспоненте δ. Тогда $(x^a)^\delta \equiv 1 \pmod m$, ab делит $a\delta$, b делит δ. Далее, $x^{ab} \equiv 1 \pmod m$, $(x^a)^b \equiv 1 \pmod m$, δ делит b. Следовательно $\delta = b$.

Теорема 2. Если $x \in \exp a \pmod m$, $y \in \exp b \pmod m$ и $(a,b)=1$, то $xy \in \exp ab \pmod m$.

Доказательство. Пусть $xy \in \exp \delta \pmod m$. Тогда $(xy)^\delta \equiv 1 \pmod m$, $(xy)^{b\delta} \equiv 1 \pmod m$, $x^{b\delta} \cdot y^{b\delta} \equiv 1 \pmod m$, $x^{b\delta} \equiv 1 \pmod m$, $a \mid b\delta$, $a \mid \delta$, ибо $(a,b)=1$. Аналогично, $b \mid \delta$. Так как $(a,b)=1$, то $ab \mid \delta$. По условию, $x^a \equiv 1 \pmod m$, $y^b \equiv 1 \pmod m$. Тогда $x^a y^b = (xy)^{ab} \equiv 1 \pmod m$, откуда $\delta \mid ab$. Следовательно $\delta = ab$.

Теорема 3. Существуют примитивные корни по модулю p.

Доказательство. В самом деле, пусть

$$\delta_1, \delta_2, \ldots, \delta_r \tag{6.3}$$

есть все различные экспоненты, которым принадлежат целые 1, $2,3,\ldots,p-1$. Пусть целое $\tau = \text{нок}(\delta_1,\delta_2,\ldots,\delta_r)$. Если

$$\delta_1 = p_0^{a_0} \cdot p_1^{a_1} \cdot p_2^{a_2} \cdot \ldots \cdot p_k^{a_k},$$
$$\delta_2 = p_0^{b_0} \cdot p_1^{b_1} \cdot p_2^{b_2} \cdot \ldots \cdot p_k^{b_k},$$
$$\ldots$$
$$\delta_r = p_0^{c_0} \cdot p_1^{c_1} \cdot p_2^{c_2} \cdot \ldots \cdot p_k^{c_k},$$

то

$$\tau = \text{нок}(\delta_1,\delta_2,\ldots,\delta_r) = p_0^{\max(a_0,a_1,\ldots,a_k)} \cdot p_0^{\max(b_0,b_1,\ldots,b_k)} \cdot \ldots \cdot p_0^{\max(c_0,c_1,\ldots,c_k)} = p_0^{d_0} p_1^{d_1} \ldots p_k^{d_k},$$

где $d_0=\max(a_0,\ldots,c_0),\ldots,\ d_k=\max(a_k,\ldots,c_k)$. Каждый множитель $p_s^{d_s}$, $s=0,1,\ldots,k$, этого представления делит число δ_{j_s} ряда (6.3). Тогда $\delta_{j_s} = e_s \cdot p_s^{d_s}$. Пусть ξ_s есть одно из чисел ряда $1,2,\ldots,p-1$, принадлежащего экспоненте δ_{j_s}. По теореме 1 число $\eta_s = \xi_s^a$ принадлежит экспоненте $p_s^{d_s}$. По теореме 2 произведение $g=\eta_0\eta_1\ldots\eta_k$ принадлежит экспоненте $p_0^{d_0}p_1^{d_1}\ldots p_k^{d_k} = \tau$ по модулю p.

Так как все целые из (6.3) делят τ, то целые $1,2,\ldots,p-1$ есть решения (теорема 1) сравнения $x^\tau \equiv 1 \pmod p$, то $p-1 \leqslant \tau$ по теореме 2 в 4.4. Но (замечание к теореме 1 в 6.1) $\tau \mid (p-1)$. Поэтому $\tau = p-1$ и g есть примитивный корень.

Теорема 4. Пусть g есть примитивный корень по модулю p. Существует такое t, что u, определяемое равенством $(g+pt)^{p-1} = 1+pu$, не делится на p. Соответствующее $g+pt$ есть примитивный корень по модулю p^α при любом $\alpha \geqslant 2$.

Доказательство. Так как $g^{\varphi(p)} = g^{p-1} \equiv 1 \pmod p$, то $g^{p-1} = 1+pT_0$ при некотором T_0. По формуле бинома Ньютона $(g+pt)^{p-1} =$

$$g^{p-1} + \binom{p-1}{1} g^{p-2}(pt)^1 + \binom{p-1}{2} g^{p-3}(pt)^2 + \ldots + (pt)^{p-1}.$$

Возьмем два первых слагаемых в правой части равенства. Все другие слагаемые делятся на p^2. Для них

$$g^{p-1} = 1+pT_0,\quad \binom{p-1}{1} g^{p-2}(pt) = (p-1)g^{p-2}(pt) = g^{p-2}p^2t - g^{p-2}pt,$$

$(g+pt)^{p-1} = 1+pT_0 + g^{p-2}p^2t - g^{p-2}pt + pT_1$ при некотором T_1. Тогда

$(g+pt)^{p-1} = 1 + p(T_0 - g^{p-2}t + pT) = 1 + pu$ при некотором T,

$$(g+pt)^{p-1} = 1+pu \text{ при } u = T_0 - g^{p-2}t + pT. \tag{6.4}$$

Если t пробегает полную систему вычетов по модулю p, то u тоже пробегает полную систему вычетов по модулю p. Следовательно, можно выбрать такое целое t, что u не делится на p. Возведем (6.4) в степень p и получим:

$$(g+pt)^{p(p-1)} = (1+pu)^p =$$
$$1 + \binom{p}{1} pu + \binom{p}{2} (pu)^2 + \ldots + (pu)^p =$$
$$1 + p^2(u + \ldots) = 1 + p^2u_2 \text{ при } u_2 = u + \ldots$$

u_2 не делится на p, иначе $(u+\ldots) \vdots p$, $u \vdots p$, что невозможно. Возводя (6.4) в степени $p, p^2, p^3, \ldots$, получаем

$$(g+pt)^{p(p-1)} = (1+pu)^p = 1 + p^2u_2,$$
$$(g+pt)^{p^2(p-1)} = (1+p^2u_2)^p = 1 + p^3u_3, \tag{6.5}$$
$$\ldots$$

где u_2, u_3 не делится на p.

Пусть $g+pt$ принадлежат экспоненте δ по модулю p^α. Тогда

$$(g+pt)^\delta \equiv 1 \pmod{p^\alpha}, \tag{6.6}$$

откуда $(g+pt)^\delta - 1 \equiv 0 \pmod{p^\alpha}$, $((g+pt)^\delta - 1) \vdots p^\alpha$, $((g+pt)^\delta - 1) \vdots p$,

$(g+pt)^{\delta}-1\equiv 0(\mathrm{mod}\, p)$, $(g+pt)^{\delta}\equiv 1\ (\mathrm{mod}\, p)$.

Так как g есть примитивный корень по модулю p, то $g+pt$ есть тоже примитивный корень по модулю p. В самом деле, если $g\in \exp\varphi(p)=p-1\ (\mathrm{mod}\, p)$, то $\varphi(p)=p-1$ есть минимально возможное целое, которого $g^{\varphi(p)}=g^{p-1}\equiv 1(\mathrm{mod}\, p)$, $(g+pt)^{p-1}\equiv 1\ (\mathrm{mod}\, p)$. Если $(g+pt)^{\varepsilon}\equiv 1(\mathrm{mod}\, p)$ при некотором $\varepsilon<\varphi(p)$, то $g^{\varepsilon}\equiv 1\ (\mathrm{mod}\, p)$, что противоречит минимальности $\varphi(p)=p-1$. Поэтому $(g+pt)\in \exp(p-1)\ (\mathrm{mod}\, p)$. Так как $(g+pt)^{\delta}\equiv 1\ (\mathrm{mod}\, p)$, то $\delta\vdots(p-1)$ и $\delta\geqslant p-1$. Так как

$$(g+pt)\in \exp\delta\ (\mathrm{mod}\, p^c) \text{ и } (g+pt)^{\varphi(p^c)}\equiv 1(\mathrm{mod}\, p^c),$$

то $\delta\,|\,\varphi(p^c)=p^{c-1}(p-1)$ по 2.4. Так как $\delta\geqslant p-1$, то $\delta=p^{r-1}(p-1)$ при некотором $r\in\{1,2,\ldots,c\}$. Поэтому $r\leqslant c$.

Далее, $(g+pt)^{\delta}=(g+pt)^{p^{r-1}(p-1)}\equiv 1(\mathrm{mod}\, p^{\alpha})$ (по (6.6)). Тогда $1+p^{r}u_r\equiv 1(\mathrm{mod}\, p^{\alpha})$ по (6.5), $p^{r}u_r\equiv 0(\mathrm{mod}\, p^{\alpha})$, $(p^{r}u_r)\vdots p^{\alpha}$. Так как u_r не делится на p, то $p^{r}\vdots p^{\alpha}$, $r\geqslant\alpha$. Так как $r\leqslant\alpha$, то $r=\alpha$. Следовательно, $\delta=p^{\alpha-1}(p-1)=\varphi(p^{\alpha})$ есть экспонента к которой принадлежит $g+pt$ по модулю p^{α} и это $g+pt$ есть примитивный корень по модулю p^{α} при всяком $\alpha\geqslant 2$.

Теорема 5. Пусть $\alpha\geqslant 1$ и g_1 есть примитивный корень по модулю p^{α}. Нечетное целое среди чисел g_1, g_1+p^{α} есть примитивный корень по модулю $2p^{\alpha}$.

Доказательство. Пусть x есть нечетное целое. Тогда сравнение $x^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha})$ для нечетного целого x^{γ} влечет $x^{\gamma}=1+T_0\cdot p^{\alpha}$ при некотором T_0. Число T_0 четно, ибо $x^{\gamma},1,p^{\alpha}$ нечетны. Тогда $x^{\gamma}=1+\frac{T_0}{2}(2p^{\alpha})$, $x^{\gamma}\equiv 1\ (\mathrm{mod}\, 2p^{\alpha})$. Если $x^{\gamma}\equiv 1(\mathrm{mod}\, 2p^{\alpha})$, то $x^{\gamma}=1+T_0\cdot 2p^{\alpha}$, $x^{\gamma}=1+T_1p^{\alpha}$ при $T_1=2T_0$, $x^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha})$. Итак,

$$x^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha}) \longleftrightarrow x^{\gamma}\equiv 1(\mathrm{mod}\,(2p^{\alpha})). \qquad (6.7)$$

Если g_1 четно, то g_1+p^{α} нечетно. Если g_1 нечетно, то g_1+p^{α} четно. Итак, одно из целых g_1, g_1+p^{α} нечетно.

Далее, $\forall\gamma\geqslant 1$

1) $g_1^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha}), (g_1+p^{\alpha})^{\gamma}=g_1^{\gamma}+\binom{\gamma}{1}g_1^{\gamma-1}p^{\alpha}+\ldots+(p^{\alpha})^{\gamma}=$

$1+Tp^{\alpha}\equiv 1\ (\mathrm{mod}\, p^{\alpha})$ при некотором T, ибо $g_1^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha})$ и $Tp^{\alpha}\vdots p^{\alpha}$.

2) $(g_1+p^{\alpha})^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha}), g_1^{\gamma}+Tp^{\alpha}\equiv 1(\mathrm{mod}\, p^{\alpha}), g_1^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha})$.

Итак, $\forall\gamma\geqslant 1$ $(g_1^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha}) \longleftrightarrow (g_1+p^{\alpha})^{\gamma}\equiv 1(\mathrm{mod}\, p^{\alpha}))$. (6.8)

Тогда $g_1\in\exp\gamma\ (\mathrm{mod}\, p^{\alpha}) \longleftrightarrow (g_1+p^{\alpha})\in\exp\gamma\ (\mathrm{mod}\, p^{\alpha})$. (6.9)

Пусть $a=p_0^{\alpha_0}\ldots p_k^{\alpha_k}$ есть каноническая факторизация a. По 2.4

$$(p_0^{\alpha_0}\ldots p_k^{\alpha_k})=p_0^{\alpha_0}\ldots p_k^{\alpha_k}\left(1-\frac{1}{p_0}\right)\left(1-\frac{1}{p_0}\right)\ldots\left(1-\frac{1}{p_k}\right)=$$

$$(p_0^{\alpha_0}-p_0^{\alpha_0-1})(p_2^{\alpha_2}-p_2^{\alpha_2-1})\ldots(p_k^{\alpha_k}-p_k^{\alpha_k-1}).$$

Тогда $\varphi(p^{\alpha})=p^{\alpha}-p^{\alpha-1}$,

$\varphi(2p^{\alpha})=(2^1-2^0)(p^{\alpha}-p^{\alpha-1})=p^{\alpha}-p^{\alpha-1}=\varphi(p^{\alpha})$ и

$$\varphi(p^{\alpha})=\varphi(2p^{\alpha}). \qquad (6.10)$$

Если g_1 нечетно, то

$g_1^{\varphi(p^{\alpha})}\equiv 1\ (\mathrm{mod}\, p^{\alpha})$ по условию $\longleftrightarrow$

$g_1^{\varphi(p^{\alpha})}\equiv 1\ (\mathrm{mod}\, 2p^{\alpha})$ по (6.7) $\longleftrightarrow$

$g_1^{\varphi(2p^{\alpha})}\equiv 1\ (\mathrm{mod}\, 2p^{\alpha})$ по (6.10) $\longleftrightarrow$

g_1 есть примитивный корень по модулю $2p^{\alpha}$ по (6.9).

Если g_1+p^{α} нечетно, то

$g_1^{\varphi(p^{\alpha})}\equiv 1\ (\mathrm{mod}\, p^{\alpha})$ по условию $\longleftrightarrow$

$(g_1+p^{\alpha})^{\varphi(p^{\alpha})}\equiv 1\ (\mathrm{mod}\, p^{\alpha})$ по (6.8) и g_1+p^{α} есть примитивный корень по модулю $2p^{\alpha}$ по аналогии со случаем, когда g_1 нечетно.

6.3. Вычисление примитивных корней по модулям p^{α} и $2p^{\alpha}$

Пусть $p\geqslant 3$ есть нечетное простое число, целое $\alpha\geqslant 1$ и $m\in\{p^{\alpha},\ 2p^{\alpha}\}$.

Теорема. Пусть $q_1,\ldots,q_k$ есть все различные простые делители целого $c=\varphi(m)$ и целое g таково, что $(g,m)=1$. Тогда

g есть примитивный корень по модулю m $\longleftrightarrow$

g не удовлетворяет ни одному из условий:

$$g^{\frac{c}{q_1}}\equiv 1(\mathrm{mod}\, m),\ g^{\frac{c}{q_2}}\equiv 1(\mathrm{mod}\, m),\ \ldots,\ g^{\frac{c}{q_k}}\equiv 1(\mathrm{mod}\, m). \qquad (6.11)$$

Доказательство. Пусть $m\in\{p^{\alpha},2p^{\alpha}\}$, $c\geqslant 1$, $q_1,\ldots,q_k$ есть все различные простые делители числа $c=\varphi(m)$.

Необходимость. $\rightarrow$. Пусть $(g,m)=1$ и g есть примитивный корень по модулю m. Тогда $g^{\varphi(m)}\equiv 1(\mathrm{mod}\, m)$ и $\forall t$, в том числе для

$t=c/q_i$, $i=1,2,\ldots,k$, $g^t\not\equiv 1(\mathrm{mod}\, m)$, то есть

$$g^{\frac{c}{q_1}}\not\equiv 1(\mathrm{mod}\, m),\ g^{\frac{c}{q_2}}\not\equiv 1(\mathrm{mod}\, m),\ \ldots,\ g^{\frac{c}{q_k}}\not\equiv 1(\mathrm{mod}\, m), \qquad (6.12)$$

то есть g не удовлетворяет ни одному из условий (6.11).

Достаточность. $\leftarrow$. Пусть g не удовлетворяет ни одному из условий (6.11), то есть g удовлетворяет всякому из (6.12). Покажем, что g есть примитивный корень по модулю m. Допустим противное: g не есть примитивный корень по модулю m. Тогда существует минимальная экспонента $\delta<\varphi(m)$, который принадлежит g по модулю m, то есть $g^\delta\equiv 1(\mathrm{mod}\, m)$ и $c=\varphi(m)\vdots\delta$. Пусть q есть простой делитель c/δ, например, $q=q_i$. Тогда $\frac{c}{\delta}=q_i u$ при некотором u, $\frac{c}{q_i}=\delta u$, $q^{\frac{c}{q_i}}=q^{\delta u}\equiv 1(\mathrm{mod}\, m)$. Противоречие с (6.12). Поэтому g есть примитивный корень по модулю m.

Пример. 1. Найти примитивный корень по простому модулю $m=41$. Имеем $\varphi(41)=40=2^3\cdot 8$. Простые делители $\varphi(41)$ есть $40/5=8$, $40/2=20$. Число g (которое не должно делиться на 41) есть примитивный корень $\leftrightarrow$ g не удовлетворяет ни одному из сравнений

$$g^8\equiv 1\ (\mathrm{mod}\ 41),\ g^{20}\equiv 1\ (\mathrm{mod}\ 41). \qquad (6.13)$$

Проверяя целые $2,3,4,\ldots$, найдем (по модулю 41):

$2^8\equiv 10$, $3^8\equiv 1$, $4^8\equiv 18$, $5^8\equiv 18$, $6^8\equiv 10$,

$2^{20}\equiv 1$, $2^{20}\equiv 1$, $5^{20}\equiv 1$, $6^{20}\equiv 40$.

Целые 2,3,4,5 не есть примитивные корни по модулю 41, ибо каждое из них удовлетворяет одному из сравнений (6.13). Целое 6 есть примитивный корень, ибо 6 не удовлетворяет ни одному из сравнений (6.13).

2. Найти примитивный корень по простому модулю $m=3362=2\cdot 41^2$. Мы нашли, примитивный корень по модулю 41^2 есть 6. Можно найти примитивный корень как в примере 1. Но мы сделаем это проще следующим образом. Мы нашли, что примитивный корень по модулю 41^2 есть 6. По теореме 6 в 6.2, нечетное среди 6 и $6+41^2$ есть примитивный корень по модулю $2\cdot 41^2$.

6.4. Индексы по модулям p^α и $2p^\alpha$

Пусть p есть нечетное простое число, $\alpha\geqslant 1$, $m\in\{p^\alpha,2p^\alpha\}$, $c=\varphi(m)$, g есть примитивный корень по модулю m.

Теорема 1. Если γ пробегает наименьшие неотрицательные вычеты $\gamma=0,1,2,\ldots,c-1$ по модулю $c=\varphi(m)$, то g^γ пробегает приведенную систему вычетов по модулю m.

Доказательство. По теореме в 6.3 g^γ пробегает c различных чисел взаимно простых с m.

Следствие. Индексы существуют для любого числа, взаимно простого с модулем m.

Замечание. По теореме 1 всякое целое a с $(a,g)=1$ имеет среди целых $\gamma=0,1,\ldots,c-1$ единственный индекс γ'. Все другие индексы γ таковы, что $\gamma\equiv\gamma'(\mathrm{mod}\, m)$. Все целые числа с данным индексом образуют класс вычетов целых чисел по модулю m.

Теорема 2. Если g есть примитивный корень по модулю m, $(a,g)=1$, $(b,g)=1,\ldots,(l,g)=1$, то

$$\mathrm{ind}\,(ab\ldots l)\equiv \mathrm{ind}\,a+\mathrm{ind}\,b+\ldots+\mathrm{ind}\,l\ (\mathrm{mod}\, m).$$

Доказательство следует из 6.1.2.

6.5. Индексы и вычеты

Пусть простое число p нечетно, $\alpha\geqslant 1$, $m\in\{p^\alpha,2p^\alpha\}$, $c=\varphi(m)$, g есть примитивный корень по модулю m.

Теорема 1. Пусть $(n,c)=d$. Тогда

1. Сравнение

$$x^n\equiv a\ (\mathrm{mod}\ m) \qquad (6.14)$$

разрешимо (a есть вычет степени n по модулю m) $\leftrightarrow$ $\mathrm{ind}\,a$ кратен n. В случае разрешимости сравнение имеет d решений.

2. Приведенная система вычетов по модулю m имеет c/d вычетов степени n.

Доказательство. 1.Сравнение (6.14) эквивалентно сравнению

$$n\,\mathrm{ind}_g x\equiv \mathrm{ind}_g a\ (\mathrm{mod}\, c). \qquad (6.15)$$

Сравнение (6.15) разрешимо $\leftrightarrow$ $\mathrm{ind}\,a$ кратен d (параграф 4.2). Если сравнение (6.15) разрешимо, то можно найти d несравнимых по модулю c значений для $\mathrm{ind}\,x$ и d соответствующих несравнимых по модулю m значений для x.

2. Среди чисел $0,1,\ldots,c-1$, являющихся наименьшими индексами вычетов приведенной системы по модулю m, имеется c/d кратных d.

Пример. 1. Для сравнения

$$x^8\equiv 23\ (\mathrm{mod}\ 41), \qquad (6.16)$$

имеем $(8,40)=8$ и $\mathrm{ind}\,23=36$ не делится на 8. Следовательно,

сравнение (6.16) неразрешимо.

2. Для сравнения

$$x^{12} \equiv 37 \pmod{41}, \qquad (6.17)$$

имеем (12,40)=4 и $\operatorname{ind} 37 = 32$ делится на 8. Следовательно, сравнение (6.16) разрешимо и имеет 4 решения. Сравнение (6.17) эквивалентно $12 \operatorname{ind} x \equiv 32 (\operatorname{mod} 40)$ и потому $\operatorname{ind} x \equiv 8 (\operatorname{mod} 10)$, откуда для $\operatorname{ind} x$ находим четыре несравнимых по модулю 40 значений: $\operatorname{ind} x = 6,16,26,36$. Соответствующие четыре решения сравнения (6.17) есть $x \equiv 39, 18, 2, 23 \pmod{41}$.

3. Целые числа

$$1,4,10,16,18,23,25,31,37,40, \qquad (6.18)$$

индексы которых кратны 4, есть все биквадратичные вычеты (а также все вычеты всякой степени $n=12,28,36,\ldots$, где $(n,40) = 4$) среди наименьших положительных вычетов по модулю 41. Число элементов в ряду (6.18) равно 10=40/4).

Теорема 2. Число a есть вычет степени n по модулю m $\longleftrightarrow$

$$a^{c/d} \equiv 1 \pmod{m}. \qquad (6.19)$$

Доказательство. Условие $\operatorname{ind} a \equiv 0 (\operatorname{mod} d)$ эквивалентно сравнению $\frac{c}{d} \operatorname{ind} a \equiv 0 (\operatorname{mod} c)$, которое эквивалентно условию (6.19).

Пример. Невозможность сравнения $g^{c/d} \equiv 1 (\operatorname{mod} m)$ эквивалентна условию, что g есть невычет степени q по модулю m. В частности, невозможность сравнения $g^{c/2} \equiv 1 (\operatorname{mod} m)$ эквивалентна условию, что g есть квадратичный невычет по модулю m.

Теорема 3. 1. Экспонента δ, которой a принадлежит по модулю m, определяется равенством $(\operatorname{ind} a, c)=c/\delta$. В частности, принадлежность a множеству примитивных корней по модулю m определяется равенством $(\operatorname{ind} a, c)=1$.

2. Число чисел, принадлежащих экспоненте δ приведенной системы вычетов по модулю m, равно $\varphi(\delta)$. В частности, число примитивных корней равно $\varphi(c)$.

Доказательство. 1. δ есть наименьшее делитель c с условием $a^{\delta} \equiv 1 (\operatorname{mod} m)$, что эквивалентно $\delta \operatorname{ind} a \equiv 0 (\operatorname{mod} c)$, или $\operatorname{ind} a \equiv 0 (\operatorname{mod} c/\delta)$. Поэтому δ есть наименьший делитель c, при котором c/δ делит $\operatorname{ind} a$. Тогда c/δ есть наибольший делитель c, делящий $\operatorname{ind} a$. Это влечет $c/\delta=(\operatorname{ind} a, c)$. Поэтому

первая часть теоремы 1 справедлива.

2. Среди целых $0,1,\ldots,c-1$, которые есть наименьшие индексы вычетов приведенной системы по модулю m, кратные c/δ есть числа вида $(c/\delta)y$, где $y=0,1,\ldots,\delta-1$. Условие $((c/\delta)y, c)=c/d$ эквивалентно условию $(y,\delta)=1$, которое удовлетворяется $\varphi(\delta)$ значениями y. Поэтому вторая часть теоремы 1 тоже справедлива.

6.6. Индексы по модулю 2^{α}

1. Пусть $\alpha=1$. Тогда $2^{\alpha}=2$. Имеем $\varphi(2)=1$. Примитивный корень по модулю 2 равен, например, $1 \equiv -1 (\operatorname{mod} 2)$. Целое $1^0 = (-1)^0 = 1$ образует приведенную систему вычетов по модулю 2.

2. Пусть $\alpha=2$. Тогда $2^{\alpha}=4$. Имеем $\varphi(4)=2$. Примитивный корень по модулю 4 равен, например, $3 \equiv -1 (\operatorname{mod} 4)$. Целые $(-1)^0 = 1$, $(-1)^1 \equiv 3 (\operatorname{mod} 4)$ образуют приведенную систему вычетов по модулю 4.

3. Пусть $\alpha \geqslant 3$. Тогда $2^{\alpha} \geqslant 8$. Имеем $\varphi(2^{\alpha}) = 2^{\alpha} - 2^{\alpha-1} = 2^{\alpha-1}(2-1)=2^{\alpha-1}$. Примитивных корней в этом случае нет, ибо экспонента, которой принадлежит по модулю 2^{α} нечетное целое x, не более $\varphi(2^{\alpha})/2 = 2^{\alpha-1}/2=2^{\alpha-2}$. В самом деле, всякое нечетное целое x можно представить в виде $x=\pm(1+4t_0)$, $t_0=0, \pm 1, \pm 2, \ldots$ Тогда для $x^{2^{\alpha-2}}$, $\alpha=3,4,5,\ldots$,

$x^{2^{3-2}} = x^2 = (\pm(1+4t_0))^2 = 1+8t_0+16t_0^2 = 1+8t_1 =$

$1+2^3 t_1 \equiv 1 \pmod{2^3}$ при некотором $t_1=t_0+8t_0^2$;

$x^{2^{4-2}} = (x^{2^2})=x^{2^2}=(x^2)^2 = (1+8t_1)^2 = 1+16t_1+64t_1^2 =$
$1+2^4 t_2 \equiv 1 \pmod{2^4}$ при некотором t_2;

$x^{2^{5-2}} = (x^{2^3})=(x^{2^2})^2=(x^4)^2=(1+8t_1)^2=1+16t_1+64t_1^2 =$
$1+2^4 t_2 \equiv 1 \pmod{2^4}$ при некотором t_3;

$x^{2^{6-2}} = (x^{2^4})=(x^{2^3})^2=(x^4)^2=(1+2^4 t_2)^2=1+2^5 t_2+64t_2^2 =$
$1+2^5 t_3 \equiv 1 \pmod{2^5}$ при некотором t_3;

...

$x^{2^{\alpha-2}} = 1+2^{\alpha} t_{\alpha-2} \equiv 1 (\operatorname{mod} 2^{\alpha})$ при некотором $t_{\alpha-2}$.

При этом существуют числа, например пять, принадлежащие экспоненте $2^{\alpha-2}$ ($\alpha \geqslant 3$). В самом деле, $5^{\varphi(m)} \equiv 1 (\operatorname{mod} m)$. В этом случае $m=2^{\alpha}$, $\varphi(m)=2^{\alpha-1}$. По теореме 2 в 6.1 экспонента, которой целые принадлежат по модулю 2^{α}, есть делители $\varphi(2^{\alpha}) = 2^{\alpha-1}$. Показано, что $5^{2^{\alpha-2}} \equiv 1 (\operatorname{mod} 2^{\alpha})$. Для всех меньших де-

лителей $2^{\alpha-2}$:

$5^1 = 1 + 2^2$, $5^1 \not\equiv 1 \pmod{2^\alpha}$,

$5^{2^1} = 1 + 2^3 + 2^4$, $5^2 \not\equiv 1 \pmod{2^\alpha}$,

$5^{2^2} = 1 + 2^4 + 2^5 u_2$ при некотором u_2, $5^{2^2} \not\equiv 1 \pmod{2^\alpha}$,

...

$5^{2^{\alpha-3}} = 1 + 2^{\alpha-1} + 2^\alpha u_{\alpha-3}$ при некотором $u_{\alpha-3}$, $5^{2^{\alpha-3}} \not\equiv 1 \pmod{2^\alpha}$.

Числа

$$5^0,\ 5^1,\ \ldots,\ 5^{2^{\alpha-2}-1},$$

$$-5^0,\ -5^1,\ \ldots,\ -5^{2^{\alpha-2}-1}$$

образуют приведенную систему вычетов по модулю 2^α. В самом деле, число этих целых есть $2 \cdot 2^{2^{\alpha-2}}$. Целые первого ряда попарно несравнимы по модулю 2^α по теореме 1 из 6.1. Аналогично для целых второго ряда. Числа первого ряда несравнимы с целыми второго ряда, ибо первые сравнимы с 1, а вторые с -1.

Эти рассуждения можно оформить в виде следующей теоремы.

Теорема 1. Пусть

$c=1$, $c_0=1$, если $\alpha=0$ или $\alpha=1$;

$c=2$, $c_0=2^{\alpha-2}$, если $\alpha \geqslant 2$.

(Итак, всегда $c \cdot c_0=\varphi(2^\alpha)=2^{\alpha-1}$). Пусть γ и γ_0 пробегают независимо друг от друга наименьшие неотрицательные вычеты $\gamma=0, 1,\ldots,c-1$, $\gamma_0=0,1,\ldots,c_0-1$ по модулю c и по модулю c_0 соответственно. Тогда $(-1)^\gamma 5^{\gamma_0}$ пробегает приведенную систему вычетов по модулю 2^α.

Теорема 2. Сравнение

$$(-1)^\gamma 5^{\gamma_0} \equiv (-1)^{\gamma'} 5^{\gamma_0'} \pmod{2^\alpha} \tag{6.20}$$

имеет место, если и только если

$$\gamma \equiv \gamma' (\mathrm{mod}\, c),\quad \gamma_0 \equiv \gamma_0' (\mathrm{mod}\, c_0).$$

Доказательство. Теорема очевидна при $\alpha=0$. Пусть $\alpha \geqslant 1$. Пусть наименьшие неотрицательные вычеты по модулю c и c_0 для целых γ,γ_0 есть r,r_0 и для целых γ',γ_0' есть r',r_0' соответственно.

По теореме 2 из 6.1 (-1 принадлежит экспоненте c и 5 принадлежит экспоненте c_0) сравнение (6.20) имеет место, если и только если $(-1)^r 5^{r_0} \equiv (-1)^{r'} 5^{r_0'} \pmod{2^\alpha}$, то есть (по теореме 1) $r=r'$, $r'=r_0'$.

Определение. Если $a \equiv (-1)^\gamma 5^{\gamma_0} \pmod{2^\alpha}$, то пара γ,γ_0 называется *системой индексов числа a по модулю 2^α*.

Замечание. По теореме 1 всякое a, которое взаимно просто с 2^α (то есть нечетное a), имеет единственную систему индексов γ',γ_0' среди $cc_0=\varphi(2^\alpha)$ пар значений γ,γ' из теоремы 1. Зная систему γ,γ', можно получить все системы индексов целого a. В соответствии с теоремой 2 они есть все пары γ,γ_0, которые образованы из неотрицательных чисел классов вычетов $\gamma \equiv \gamma' \pmod{c}$, $\gamma_0 \equiv \gamma_0' (\mathrm{mod}\, c_0)$.

Числа с данной системой индексов γ,γ_0 образуют класс вычетов по модулю 2^α.

Теорема 3. Индексы произведения сравнимы по модулям c и c_0 с суммами индексов сомножителей.

Доказательство. Пусть $\gamma(a),\gamma_0(a)$; ...; $\gamma(l),\gamma_0(l)$ есть системы индексов чисел $a,\ldots,l$. Имеем

$$a \ldots l \equiv (-1)^{\gamma(a)+\ldots+\gamma(l)} 5^{\gamma_0(a)+\ldots+\gamma_0(l)}.$$

Поэтому $\gamma(a)+\ldots+\gamma(l)$, $\gamma_0(a)+\ldots+\gamma_0(l)$ есть индексы произведения $a \cdot \ldots \cdot l$.

6.7. Индексы по любому составному модулю

Пусть $m = 2^\alpha p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}$ есть каноническая факторизация числа m. Пусть c,c_0 есть числа, указанные в теореме 1 из 6.6; $c_s = \varphi(p_s^{\alpha_s})$; g_s есть наименьший положительный примитивный корень по модулю $p_s^{\alpha_s}$, $s=1,2,\ldots,k$.

Определение. Если

$$\begin{aligned} &a \equiv (-1)^\gamma 5^{\gamma_0} \pmod{2^\alpha}, \\ &a \equiv g_1^{\gamma_1} \pmod{p_1^{\alpha_1}}, \ldots, a \equiv g_k^{\gamma_k} \pmod{p_k^{\alpha_k}}, \end{aligned} \tag{6.21}$$

то система $\gamma,\gamma_1,\ldots,\gamma_k$ называется *системой индексов числа a по модулю m*.

Теорема 1. Числа a с данной системой индексов $\gamma,\gamma_1,\ldots,\gamma_k$ образуют класс вычетов по модулю m.

Доказательство. Определение системы индексов числа a по модулю m и (6.21) влекут, что γ,γ_0 есть система индексов числа a по модулю 2^α и числа $\gamma_1,\ldots,\gamma_k$ есть индексы a по модулям $p_1^{\alpha_1},\ldots,p_k^{\alpha_k}$ соответственно. Поэтому (определения из 6.6 и 6.4) всякое a, которое взаимно просто с m (тогда a взаимно

просто с каждым из $p_1^{\alpha_1},\ldots,p_k^{\alpha_k}$) имеет единственную систему индексов $\gamma',\gamma_0',\gamma_1',\ldots,\gamma_k'$ среди $cc_0c_1\ldots c_k=\varphi(m)$ систем $\gamma,\gamma_0,\gamma_1,\ldots,\gamma_k$, которые получаются, если $\gamma,\gamma_0,\gamma_1,\ldots,\gamma_k$ пробегают независимо друг от друга наименьшие неотрицательные вычеты по модулю $c,c_0,c_1,\ldots,c_k$. Все системы индексов числа a есть все системы $\gamma,\gamma_0,\gamma_1,\ldots,\gamma_k$, составленные из неотрицательных чисел классов

$$\gamma\equiv\gamma_1'(\mathrm{mod}\,c),\gamma_0\equiv\gamma_0'(\mathrm{mod}\,c_0),$$
$$\gamma_1\equiv\gamma_1'(\mathrm{mod}\,c_1),\ldots,\gamma_k\equiv\gamma_k'(\mathrm{mod}\,c_k).$$

Числа a с данной системой индексов $\gamma,\gamma_0,\gamma_1,\ldots,\gamma_k$ есть решения системы (6.21). Следовательно, (по 4.3) такие числа a образуют класс вычетов по модулю m.

Теорема 2. Индексы произведения сравнимы по модулям c, $c_0,c_1,\ldots,c_k$ с индексами сомножителей.

Доказательство. Индексы $\gamma,\gamma_0,\gamma_1,\ldots,\gamma_k$ числа a по модулю m есть индексы по модулям $2^\alpha,p_1^{\alpha_1},\ldots,p_k^{\alpha_k}$ соответственно.

Теорема 3. Примитивные корни по модулю m существуют лишь при $m=2,4,p^\alpha,2p^\alpha$, где простое p нечетно и целое $\alpha\geqslant1$.

Доказательство. Пусть $\tau=\varphi(2^\alpha)$ при $\alpha\leqslant2$ и $\tau=\varphi(2^\alpha)/2$ при $\alpha>2$. Пусть h есть наименьшее общее кратное чисел $\tau,c_1,\ldots,c_k$. При всяком a, взаимно простом с m, сравнение $a^h\equiv1(\mathrm{mod}\,m)$ верно по всем модулям $2^\alpha,p_1^{\alpha_1},p_2^{\alpha_2},\ldots,p_k^{\alpha_k}$. Тогда это сравнение верно по модулю m. Поэтому a не может быть примитивным корнем по модулю m, если $h<\varphi(m)$. Но последнее неравенство возможно лишь при $\alpha>2$, $k>1$, $\alpha=2$, $k=1$. Поэтому примитивные корни возможны лишь при $m=2,4,p^\alpha,2p^\alpha$. Существование примитивных корней в этих случаях было доказано ранее в 6.6, 6.2). Теорема доказана.

Замечание. 1. $\mathbb{Z}_n^*$ имеет генератор (по умножению), то есть $\mathbb{Z}_n^*$ есть циклическая группа, если и только если $n=2,4,p^k,2p^k$, где простое $p\geqslant3$ и целое $k\geqslant1$.

2. Пусть α из $\mathbb{Z}_n^*$ есть генератор для $\mathbb{Z}_n^*$. Тогда 1) $\alpha^i(\mathrm{mod}\,n)$ есть генератор для $\mathbb{Z}_n^*$, если и только если $\mathrm{нод}(i,\varphi(n))=1$; 2) число генераторов для $\mathbb{Z}_n^*$ равно $\varphi(\varphi(n))$.

3. α из $\mathbb{Z}_n^*$ есть генератор для $\mathbb{Z}_n^*$, если и только если $\alpha^{\varphi(n)/p}\not\equiv1\ (\mathrm{mod}\,n)$ для каждого простого делителя p значения функции Эйлера $\varphi(n)$.

7. ГРУППА, КОЛЬЦО, ПОЛЕ

Рассмотрим определенные в введении понятия группы, кольца поля несколько подробнее.

7.1. Группа

Определение. *Группа* есть множество G на котором определены бинарная операция $x*y$, унарная операция x^{-1}, отмеченный элемент e (единица), удовлетворяющие следующим аксиомам: $\forall x,y,z\in G$

1. $(x*y)*z = x*(y*z)$.
2. $x^{-1}*x = x*x^{-1} = e$.
3. $x*e = e*x = x$.

Группа G коммутативна (или абелева), если дополнительно

4. $x*y = y*x$.

Замечание. 1. Группа обозначается через $(G,\{*,^{-1},e\})$ или просто буквой G.

2. Иногда знак умножения $*$ заменяется точкой, иногда опускается совсем и тогда говорят о мультипликативной группе. Если вместо умножения используется знак сложения, то говорят об аддитивной группе. Иногда в мультипликативной группе единица e заменяется на 1, а в аддитивной группе на 0.

Определение. Группа *конечна*, если число ее элементов конечно. *Порядок конечной группы* есть число ее элементов.

Степень $a^n = a\cdot a\cdot\ldots\cdot a$ (n раз). По определению полагаем $a^0=e$, $a^{-n} = (a^{-1})^n$. Для аддитивной группы $n\cdot a = a+a+\ldots+a$.

Пример. 1. Множество целых чисел $\mathbb{Z}$ со сложением есть аддитивная абелева группа $(\mathbb{Z},\{+,-,e=0\})$.

2. Множество рациональных чисел $\mathbb{Q}$ без нуля с умножением есть мультипликативная абелева группа $(\mathbb{Q}-\{0\},\{\cdot,^{-1},e=1\})$.

3. Множество $\mathbb{Z}_n$ с операцией сложения по модулю n есть аддитивная абелева группа порядка n. Множество $\mathbb{Z}_n$ с операцией умножения по модулю n не есть группа, ибо не все элементы имеют обратный элемент. Множество $\mathbb{Z}_n^*$ с умножением по модулю n есть мультипликативная абелева группа порядка $\varphi(n)$ с единицей $e=1$.

Определение. Непустое множество $H\subseteq G$ есть *подгруппа* группы G, если H есть группа по отношению к операциям, определенным в G. Если $H\neq G$, то H есть *собственная подгруппа* в G.

Далее до конца параграфа рассматриваются мультипликатив-

ные абелевы группы.

Замечание. Множество $\langle a\rangle = \{a^n : n \in \mathbb{Z}\}$ всех целых степеней элемента a группы G образует абелеву группу, порожденную элементом a. Множество $\langle a\rangle$ есть подгруппа группы G.

Определение. Группа G называется *циклической*, если существует элемент $a \in G$, для которого $G = \langle a\rangle$. Элемент a называется *генератором* группы G.

Определение. *Порядок* $\mathrm{ord}(a)$ *элемента* a *группы* есть порядок циклической подгруппы $\langle a\rangle$, порожденной элементом a. Если подгруппа $\langle a\rangle$ бесконечна, то $\mathrm{ord}(a) = \infty$.

Замечание. 1. Порядок элемента a группы G есть наименьшее положительное целое число t, для которого $a^t = e$, если такое t существует. Если такого числа t не существует, то порядок $\mathrm{ord}(a) = \infty$.

2. Пусть G есть группа и пусть $a \in G$ есть элемент конечного порядка t. Тогда мощность $|\langle a\rangle| = t$.

Утверждение (Лагранж). Если G есть конечная группа и H есть подгруппа в G, то $|H|$ делит $|G|$. Если $a \in G$, то $\mathrm{ord}(a)$ делит $|G|$.

Утверждение. Каждая подгруппа циклической группы G циклична. Если G есть циклическая группа порядка n, то для каждого положительного делителя d для n группа G содержит в точности одну подгруппу порядка d.

Утверждение. Пусть G есть группа.

1. Если порядок элемента a из G есть t, то порядок элемента a^k равен $t/\mathrm{нод}(t,k)$.

2. Если G есть циклическая группа порядка n и число $d|n$, то G имеет в точности $\varphi(d)$ элементов порядка d. В частности, G имеет $\varphi(n)$ генераторов.

Пример. Рассмотрим мультипликативную группу $\mathbb{Z}_{19}^* = \{1,2,\ldots,18\}$ порядка 18. Группа $\mathbb{Z}_{19}^*$ циклична и ее генератор $\alpha = 2$. В следующей таблице приведены подгруппы группы $\mathbb{Z}_{19}^*$ и их генераторы.

Подгруппа	Генераторы	Порядок
{1}	1	1
{1,18}	18	2
{1,7,11}	7,11	3
{1,7,8,11,12,18}	8,12	6
{1,4,5,6,7,9,11,16,17}	4,5,6,9,16,17	9
{1,2,3,...,18}	2,3,10,13,14,15	18

7.2. Кольцо

Определение. *Кольцо* $(R,\{+,\cdot\})$ есть множество R, на котором определены две операции (функции):

сложение $x+y: R\times R \to R$ и умножение $x\cdot y: R\times R \to R$,

удовлетворяющие следующим аксиомам. $\forall x,y,z \in F$

1. $(x+y)+z = x+(y+z)$. 2. $x+y = y+x$.
3. $\exists 0 \in R\ x+0 = x$. 4. $\forall x \in R\ \exists (-x) \in R\ x+(-x)=0$.
5. $(xy)z = x(yz)$. 6. $(x+y)z = xz+yz$, $x(y+z) = xy+xz$.

Кольцо *коммутативно*, если умножение коммутативно:

7. $xy = yx$.

Замечание. Кольцо R конечно, если множество R содержит конечное число элементов. Кольцо есть коммутативная группа по сложению.

Пример. 1. Множество целых чисел $\mathbb{Z}$ со сложением и умножением есть коммутативное кольцо.

2. Множество $\mathbb{Z}_n$ со сложением и умножением по модулю n есть коммутативное кольцо.

Определение. Элемент e из R есть *единица* кольца R, если для всякого элемента x из R $x\cdot e = e\cdot x = x$.

Замечание. Не все кольца имеют единицу. Если единица существует, то иногда ее обозначают через 1.

Определение. Элемент a кольца R есть *обратимый элемент*, если существует (обратный) элемент $a^{-1} \in R$, для которого $a\cdot a^{-1} = a^{-1}\cdot a = e$.

Утверждение. Множество обратимых элементов кольца R образует мультипликативную группу.

Замечание. Группа обратимых элементов кольца $\mathbb{Z}_n$ есть $\mathbb{Z}_n^*$.

7.3. Поле

Определение. *Поле* $(F,\{+,\cdot\})$ есть множество F, на котором определены две операции (функции):

сложение $x+y: F\times F \to F$ и умножение $x\cdot y: F\times F \to F$,

удовлетворяющие следующим аксиомам. $\forall x,y,z \in F$

1. $(x+y)+z = x+(y+z)$. 2. $x+y = y+x$.
3. $\exists 0 \in F\ x+0 = x$. 4. $\forall x \in F\ \exists (-x) \in F\ x+(-x)=0$.
5. $(xy)z = x(yz)$. 6. $xy = yx$.
7. $\exists e \in F\ \forall x \in F\ x\cdot e = x$. 8. $\forall x \in F-\{0\}\ \exists x^{-1} \in F\ x\cdot x^{-1}=e$.

9. $(x+y)z = xz+yz$.

Замечание. Поле есть коммутативная группа по сложению. Поле без нуля есть коммутативная циклическая группа по умножению.

Определение. *Характеристика поля* есть наименьшее положительное число m, для которого $\sum_{i=1}^{m} e = 0$, если такое число существует. *Характеристика поля* есть 0, если $\sum_{i=1}^{m} e \neq 0$ для всякого $m \geqslant 1$.

Пример. Следующие числовые множества со сложением и умножением образуют поле характеристики ноль: множество рациональных чисел $\mathbb{Q}$, множество вещественных чисел $\mathbb{R}$, множество комплексных чисел $\mathbb{C}$.

Утверждение. Множество $\mathbb{Z}_n$ со сложением и умножением по модулю n есть поле, если и только если n есть простое число. Если число n просто, то $\mathbb{Z}_n$ имеет характеристику n.

Утверждение. Ненулевая характеристика всякого конечного поля есть простое число.

Определение. Подмножество H поля F есть *подполе* поля F, если H есть поле по отношению к операциям, определенным в F. В этом случае поле F есть *расширение* поля H.

7.4. Полиномиальные кольца

Определение. Если R есть коммутативное кольцо, то *полином* переменной x над кольцом R есть выражение вида $f(x) = a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0$, где каждое $a_i \in R$ и $n > 0$. Элемент a_i есть *коэффициент* при x^i в $f(x)$. Наибольшее целое m, для которого $a_m \neq 0$, есть *степень* $f(x)$, обозначаемая $\deg f(x)$; a_m есть *старший коэффициент* $f(x)$. Если $f(x) = a_0$ (*константный полином*) и $a_0 \neq 0$, то $f(x)$ имеет степень 0. Если все коэффициенты в $f(x)$ равны нулю, то $f(x)$ есть *нулевой полином* и его степень полагают равной $-\infty$. Полином $f(x)$ *нормирован*, если его старший коэффициент равен 1.

Определение. Если R есть коммутативное кольцо, то *полиномиальное кольцо* $R[x]$ есть кольцо всех полиномов переменной x с коэффициентами из R. Сложение и умножение полиномов определяется обычным образом. Сложение и умножение коэффициентов выполняется в кольце R.

Пример. Пусть $f(x)=x^3+x+1$ и $g(x)=x^2+x$ есть полиномы из кольца $\mathbb{Z}_2[x]$. Тогда $f(x)+g(x) = x^3+x^3+1$ и $f(x)\cdot g(x) = x^5+x^4+x^3+x$.

Пусть везде в дальнейшем F есть произвольное поле. Свойства полиномиального кольца $F[x]$ имеют много общего со свойствами кольца целых чисел.

Определение. Пусть $f(x) \in F[x]$ есть полином степени не меньше 1. Тогда $f(x)$ *неприводим над* F, если $f(x)$ не может быть представлен как произведение двух полиномов положительной степени из $F[x]$.

Утверждение. (Деление полиномов с остатком). Если $g(x),h(x) \in F[x]$ и $h(x) \neq 0$, то существуют единственные полиномы $q(x),r(x) \in F[x]$, для которых

$$g(x) = q(x)h(x) + r(x), \text{ где } \deg(r(x)) < \deg(h(x)).$$

Полином $q(x)$ есть *частное*, полином $r(x)$ есть *остаток* от деления $g(x)$ на $h(x)$.

Замечание. Иногда остаток от деления обозначается через $g(x) \pmod{h(x)}$, а частное через $g(x) \operatorname{div} h(x)$.

Пример. (Деление полиномов). Пусть полиномы $g(x) = x^6+x^5+x^3+x^2+x+1$ и $h(x) = x^4+x^3+1$ лежат в $\mathbb{Z}_2[x]$. Тогда частное $g(x) \operatorname{div} h(x) = x^2$ и остаток $g(x) \pmod{h(x)} = x^3+x+1$.

Определение. Если $g(x),h(x) \in F[x]$, то $h(x)$ *делит* $g(x)$ (обозначение: $h(x) \mid g(x)$), если остаток от деления $g(x) \pmod{h(x)} = 0$.

Определение. Если $g(x),h(x),f(x) \in F[x]$, то $g(x)$ *сравнимо* с $h(x)$ по модулю $f(x)$ (обозначение: $g(x) \equiv h(x) \pmod{f(x)}$), если $f(x)$ делит $g(x)-h(x)$.

Замечание. $g(x) \equiv h(x) \pmod{f(x)}$, если $g(x)$ и $h(x)$ при делении на $f(x)$ дают один и тот же остаток.

Утверждение. (Свойства сравнений). Для всех $g(x)$, $g_1(x)$, $h(x)$, $h_1(x)$, $s(x)$ из $F[x]$ справедливо следующее.

1. $g(x) \equiv g(x) \pmod{f(x)}$.
2. Если $g(x) \equiv h(x) \pmod{f(x)}$, то $h(x) \equiv g(x) \pmod{f(x)}$.
3. Если $g(x) \equiv h(x) \pmod{f(x)}$ и $h(x) \equiv s(x) \pmod{f(x)}$, то $g(x) \equiv s(x) \pmod{f(x)}$.
4. Если $g(x) \equiv g_1(x) \pmod{f(x)}$ и $h(x) \equiv h_1(x) \pmod{f(x)}$, то
$g(x)+h(x) \equiv g_1(x)+h_1(x) \pmod{f(x)}$,
$g(x)\cdot h(x) \equiv g_1(x)\cdot h_1(x) \pmod{f(x)}$.

Определение. Пусть полином $f(x) \in F[x]$. *Класс эквивалентности* полинома $g(x) \in F[x]$ есть множество всех полиномов из $F[x]$, сравнимых с $g(x)$ по модулю $f(x)$.

Замечание. Отношение сравнения по модулю $f(x)$ есть отношение эквивалентности. Оно разбивает $F[x]$ на классы эквивалентности. Если $g(x) \in F[x]$, то деление с остатком $g(x)$ на

$f(x)$ дает единственные полиномы $q(x), r(x) \in F[x]$, для которых $g(x) = q(x)f(x) + r(x)$, где $\deg(r(x)) < \deg(f(x))$. Поэтому каждый полином $g(x)$ сравним по модулю $f(x)$ с единственным полиномом степени меньшей, чем степень $f(x)$. Полином $r(x)$ можно взять в качестве представителя класса эквивалентности полиномов $g(x)$, дающих при делении на $f(x)$ остаток $r(x)$.

Определение. Пусть $F[x]/(f(x))$ есть множество всех полиномов из $F[x]$ степени меньше $n = \deg(f(x))$. Сложение и умножение в $F[x]/(f(x))$ выполняется по модулю $f(x)$.

Утверждение. $F[x]/(f(x))$ есть коммутативное кольцо.

Утверждение. Если полином $f(x)$ неприводим над F, то кольцо $F[x]/(f(x))$ есть а поле.

7.5. Векторное пространство

Определение. *Линейное векторное пространство над полем* F есть множество V элементов любой природы (векторов) $\mathbf{x},\mathbf{y},\mathbf{z}, \ldots$, на котором определены две (линейные) операции: сложение $\mathbf{x} + \mathbf{y}: V \times V \to V$ и (скалярное) умножение $a \cdot \mathbf{x}: F \times V \to V$ вектора $\mathbf{x}$ на элемент (скаляр) a из F (при этом $a\mathbf{x} = \mathbf{x}a$), которые $\forall \mathbf{x},\mathbf{y},\mathbf{z} \in V$, $\forall a,b,c \in F$ удовлетворяют следующим аксиомам.

1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
2. $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.
3. $\exists \mathbf{0} \in V \ \mathbf{x} + \mathbf{0} = \mathbf{x}$.
4. $\exists (-\mathbf{x}) \in V \ \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.
5. $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$.
6. $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$.
7. $(a \cdot b)\mathbf{x} = a(b\mathbf{x})$
8. $e \cdot \mathbf{x} = \mathbf{x}$, $e \in F$.

Вектор $\mathbf{0}$ из V есть нулевой вектор. Вектор $-\mathbf{x}$ из V противоположен (обратен по сложению, или аддитивно обратен) к вектору $\mathbf{x}$.

Замечание. Множество V есть аддидивная абелева группа.

Пример. 1. Множество всех векторов из $\mathbb{R}^3$ с операциями суммы двух векторов и умножением вектора на вещественное число есть линейное векторное пространство.

2. Множество векторов $V = \mathbb{R}^n = \{\mathbf{x} = (x_1,x_2,\ldots,x_n): x_1, x_2,\ldots,x_n \in \mathbb{R}\}$ с операциями

$$\mathbf{x}+\mathbf{y} = (x_1+y_1,\ldots, x_n+y_n),\ a\mathbf{x}=(ax_1,\ldots,ax_n),\ a\in\mathbb{R},$$
$$-\mathbf{x} = -1\cdot\mathbf{x} = (-x_1, -x_2,\ldots,-x_n),\ \mathbf{0} = (0,0,\ldots,0)$$

образует линейное векторное пространство.

3. Множество всех вещественных матриц одинакового размера с операциями сложения матриц и умножения матрицы на вещественное число образует линейное векторное пространство.

4. Множество всех вещественных полиномов степени не более n с операциями сложения полиномов и умножения полинома на вещественное число образует линейное векторное пространство.

5. Пусть $F = \mathbb{Z}_p = \{0,1,\ldots,p-1\}$; $a + b = \mathrm{rest}(a+b,p)$, $a \cdot b = \mathrm{rest}(a\cdot b,p)$, $\forall a,b \in F$;

$$V = \mathbb{Z}_p^n = \{\mathbf{x} = (x_1,x_2,\ldots,x_n):\ x_i \in F,\ i = 1,2,\ldots,n\};$$

$$\mathbf{x} + \mathbf{y} = (x_1+y_1,x_2+y_2,\ldots,x_n+y_n),\ \forall \mathbf{x},\mathbf{y} \in V;$$
$$a\mathbf{x} = (ax_1,ax_2,\ldots,ax_n),\ a \in F,\ \mathbf{x} \in V.$$

Множество всех векторов из V с операциями суммы $\mathbf{x} + \mathbf{y}$ и произведения $a\mathbf{x}$ вектора $\mathbf{x}$ на элемент a из F есть линейное векторное пространство.

Определение. Пусть V есть векторное пространство над полем F и множество $U \subseteq V$. Множество U есть *подпространство пространства* V, если U по отношению к линейным операциям, определенным в V, есть линейное векторное пространство.

Определение. Пусть $S = \{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ есть конечное подмножество векторного пространства V над полем F.

1. *Линейная комбинация* векторов из S есть выражение вида $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 +\ldots+ a_n\mathbf{v}_n$, где каждое $a_i \in F$.

2. *Оболочка* для множества S (обозначение $<S>$), есть множество всех линейных комбинаций векторов из S. Заметим, что оболочка для S есть подпространство пространства V.

3. Если U есть подпространство в V, то U *натянуто* на S (S стягивает U), если $<S> = U$.

4. Множество векторов S *линейно зависимо* над F, если в F существуют скаляры $a_1,a_2,\ldots,a_n$, не все нули, для которых $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 +\ldots+ a_n\mathbf{v}_n = 0$. Если таких скаляров не существует, то множество векторов S *линейно независимо* над F.

5. Если векторное пространство V натянуто на линейно независимую систему векторов S (или система S стягивает пространство V), то система S называется *базисом* для V.

Утверждение. Пусть V есть векторное пространство.

1. Если V имеет конечное стягивающее множество векторов, то V имеет базис.

2. Если V имеет базис, то все базисы имеют одинаковое число векторов.

Определение. Если векторное пространство V имеет базис, то число векторов в базисе называется *размерностью* V (обозначение: $\dim V$).

Пример. Если F есть поле, то n–кратное декартово произведение $V = F \times F \times \ldots \times F$, n раз, есть векторное пространство над F размерности n. *Стандартный базис* для V есть $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, где $\mathbf{e}_i$ есть вектор с единицей e в i–ой координате и нулями в остальных.

Определение. Пусть поле E есть расширение поля F. Тогда E есть линейное векторное пространство над полем F, где сложение векторов и скалярное умножение есть опрецаии сложения и умножения в поле E.

Размерность векторного пространства E над полем F называется *степенью* E над F (обозначение: $[E:F]$). Если эта степень конечна, то E есть *конечное расширение* поля F.

Утверждение. Пусть F, E, L есть некоторые поля. Если L есть конечное расширение E и E есть конечное расширение F, то L есть конечное расширение F и $[L:F] = [L:E] \cdot [E:F]$.

7.6. Конечные поля

7.6.1. Основные свойства полей

Определение. Поле F *конечно*, если число элементов в F конечно. *Порядок* F есть число элементов в F.

Утверждение. (Существование и единственность конечных полей). 1. Если F есть конечное поле, то F содержит p^m элементов для некоторого простого числа p и целого $m \geq 1$.

2. Для каждой степени p^m существует единственное (с точностью до изоморфизма) конечное поле порядка p^m. Это поле обозначается через $\mathbb{F}_{p^m}$, или иногда через $GF(p^m)$.

Замечание. Если p есть простое число, то $\mathbb{Z}_p$ есть поле, и каждое поле $\mathbb{F}_p$ порядка p изоморфно $\mathbb{Z}_p$. Поэтому конечное поле $\mathbb{F}_p$ можно (изоморфно) отождествлять с $\mathbb{Z}_p$.

Утверждение. Если $\mathbb{F}_q$ есть конечное поле порядка $q = p^m$, где число p просто, то характеристика $\mathbb{F}_q$ есть p.

Замечание. Поле $\mathbb{F}_q$ содержит (изоморфное $\mathbb{Z}_p$) подполе $\mathbb{F}_p$. Поэтому $\mathbb{F}_q$ есть расширение поля $\mathbb{F}_p$ степени m.

Утверждение. (Подполя конечного поля). Пусть $\mathbb{F}_q$ есть конечное поле порядка $q = p^m$. Тогда каждое подполе в $\mathbb{F}_q$ имеет порядок p^n при некотором n, которое есть делитель m. Обратно, если n есть положительный делитель m, то существует только одно подполе в $\mathbb{F}_q$ порядка p^n. Элемент $a \in \mathbb{F}_q$ лежит в

подполе $\mathbb{F}_{p^n}$, если и только если $a^{p^n} = a$.

Утверждение. Ненулевые элементы из $\mathbb{F}_q$ образуют мультипликативную группу, обозначаемую $\mathbb{F}_q^*$.

Утверждение. $\mathbb{F}_q^*$ есть циклическая группа порядка $q-1$. Поэтому $a^q = a$ для всех $a \in \mathbb{F}_q$.

Определение. *Примитивный элемент* в $\mathbb{F}_q$ есть генератор циклической группы $\mathbb{F}_q^*$.

Утверждение. Если $a, b \in \mathbb{F}_q$, а конечное поле имеет характеристику p, то $(a+b)^{p^t} = a^{p^t} + b^{p^t}$ для всех $t \geq 0$.

Определение. Пусть ненулевые полиномы $g(x), h(x) \in \mathbb{Z}_p[x]$. *Наибольший общий делитель* для $g(x)$ и $h(x)$ (обозначение: нод($g(x), h(x)$)) есть нормированный полином наибольшей степени из $\mathbb{Z}_p[x]$, делящий $g(x)$ и $h(x)$. По определению нод(0,0) = 0.

Утверждение. Каждый ненулевой полином $f(x) \in \mathbb{Z}_p[x]$ допускает факторизацию

$$f(x) = a f_1(x)^{e_1} f_2(x)^{e_2} \ldots f_k(x)^{e_k},$$

где $f_i(x)$ есть различные нормированные неприводимые полиномы из $\mathbb{Z}_p[x]$; e_i есть положительные целые числа и $a \in \mathbb{Z}_p$. Факторизация единственна с точностью до порядка сомножителей.

Утверждение. Пусть $f(x) \in \mathbb{Z}_p[x]$ есть неприводимый полином степени m. Тогда $\mathbb{Z}_p[x]/(f(x))$ есть конечное поле порядка p^m. Сложение и умножение полиномов выполняются по модулю $f(x)$.

Утверждение. Для всякого целого $m \geq 1$ существует нормированный неприводимый полином степени m над $\mathbb{Z}_p$.

Замечание. Элементы конечного поля $\mathbb{F}_q$, где $q=p^m$ и p есть простое число, допускают (изоморфное) *стандартное полиномиальное представление*, то есть элементы конечного поля $\mathbb{F}_q$ допускают представление полиномами из $\mathbb{Z}_p[x]$ степени меньше m. Если $g(x)$, $h(x) \in \mathbb{F}_{p^m}$, то сложение $g(x)$ и $h(x)$ выполняется как обычное сложение полиномов из $\mathbb{Z}_p[x]$. Чтобы получить произведение $g(x) \cdot h(x)$, следует сначала перемножить полиномы $g(x)$ и $h(x)$ обычным образом, а затем взять остаток от деления результата перемножения на $f(x)$. Если m=1, то $\mathbb{F}_q$ есть $\mathbb{Z}_p$ и арифметические операции выполняются по модулю p.

Утверждение. (Число нормированных неприводимых полиномов). Пусть p есть простое число и целое $m \geq 1$.

1. Число $N_p(m)$ нормированных неприводимых полиномов степени m в $\mathbb{Z}_p[x]$ задается следующей формулой.

$$N_p(m) = \frac{1}{m} \sum_{d \mid m} \mu(d) p^{m/d},$$

где $\mu(d)$ есть функция Мебиуса и суммирование берется по всем положительным делителям d числа m.

2. Вероятность случайного нормированного полинома степени m в $\mathbb{Z}_p[x]$ быть неприводимым над $\mathbb{Z}_p$ оценивается неравенством

$$\frac{1}{2m} \leqslant \frac{N_p(m)}{p^m} \approx \frac{1}{m}, \text{ откуда } N_p(m) \approx \frac{p^m}{m}.$$

Определение. Неприводимый полином $f(x) \in \mathbb{Z}_p[x]$ степени m есть *примитивный полином*, если x есть генератор для мультипликативной группы $\mathbb{F}^*_{p^m}$ всех ненулевых элементов конечного поля $\mathbb{F}_{p^m} = \mathbb{Z}_p[x]/(f(x))$.

Утверждение. Для каждого $m \geqslant 1$ существует нормированный примитивный полином степени m над $\mathbb{Z}_p$. Существует в точности $P_p(m) = \varphi(p^m-1)/m$ примитивных полиномов степени m над $\mathbb{Z}_p$, где φ есть функция Эйлера. Вероятность случайного нормированного неприводимого полинома степени m в $\mathbb{Z}_p[x]$ быть примитивным над $\mathbb{Z}_p$ есть число

$$P_p(m)/N_p(m) \approx \varphi(p^m-1)/p^m \approx 1/(6 \ln\ln p^m).$$

Утверждение. Неприводимый полином $f(x) \in \mathbb{Z}_p[x]$ степени m есть примитивный полином, если и только если $f(x)$ делит x^k-1 для $k = p^m-1$, и $f(x)$ не делит x^k-1 для всякого меньшего положительного целого k.

Утверждение. Пусть p есть простое число и пусть все различные простые делители числа p^m-1 есть $r_1, r_2, \ldots, r_t$. Тогда неприводимый полином $f(x) \in \mathbb{Z}_p[x]$ примитивен, если и только если для каждого i, $1 \leqslant i \leqslant t$,

$$x^{(p^m-1)/r_i} \not\equiv 1 \pmod{f(x)},$$

то есть x есть элемент порядка p^m-1 в поле $\mathbb{Z}_p[x]/(f(x))$.

Пример. (Конечное поле $\mathbb{F}_{2^4}$ порядка 16). Можно проверить, что полином $f(x) = x^4+x+1$ неприводим над $\mathbb{Z}_2$. Поэтому конечное поле $\mathbb{F}_{2^4}$ допускает представление в виде полиномов над $\mathbb{F}_2$ степени меньше чем 4. То есть

$$\mathbb{F}_{2^4} = \{a_3x^3+a_2x^2+a_1x+a_0 : a_i \in \{0,1\}\}.$$

Будем для удобства полином $a_3x^3+a_2x^2+a_1x+a_0$ представлять вектором $(a_3a_2a_1a_0)$ длины 4 и тогда

$$\mathbb{F}_{2^4} = \{(a_3a_2a_1a_0) : a_i \in \{0,1\}\}.$$

1. Сложение полиномов есть поразрядное сложение по модулю 2 их представляющих векторов: (1011) + (1001) = (0010).

2. Чтобы получить произведение (1101)·(1001), надо перемножить их представляющие полиномы и затем взять остаток от деления произведения на $f(x)$:

$$(x^3+x^2+1)\cdot(x^3+1) = x^6+x^5+x^2+1 \equiv x^3+x^2+x+1 \pmod{f(x)}.$$

Тогда (1101)·(1001) = (1111).

3. Мультипликативная единица в $\mathbb{F}_{2^4}$ есть (0001).

4. Обратный для (1011) есть элемент (0101), ибо $(x^3+x+1)\cdot(x^2+1) = x^5+x^2+x+1 = 1 \pmod{f(x)}$, то есть (1011)·(0101) = (0001).

$f(x)$ есть примитивный полином, ибо элемент x=(0010) есть генератор для $\mathbb{F}_{2^4}$, то есть все ненулевые элементы в $\mathbb{F}^*_{2^4}$ можно получить как степени элемента x. Вычисления приведены в следующей таблице.

i	$x^i \pmod{x^4+x+1}$	вектор
0	1	(0001)
1	x	(0010)
2	x^2	(0100)
3	x^3	(1000)
4	$x+1$	(0011)
5	x^2+x	(0110)
6	x^3+x^2	(1100)
7	x^3+x+1	(1011)
8	x^2+1	(0101)
9	x^3+x	(1010)
10	x^2+x+1	(0111)
11	x^3+x^2+x	(1110)
12	x^3+x^2+x+1	(1111)
13	x^3+x^2+1	(1101)
14	x^3+1	(1001)

7.6.2. Алгоритм Евклида для полиномов из $\mathbb{Z}_p[x]$

ВХОД. Ненулевые полиномы $f(x), h(x) \in \mathbb{Z}_p[x]$.
ВЫХОД. Наибольший общий делитель для $f(x)$ и $h(x)$.

1. Пока $h(x) \neq 0$ выполнять следующее.
   $r(x) := f(x) \pmod{h(x)}$, $f(x) := h(x)$, $h(x) := r(x)$.
2. Если $p>2$, $a \neq 1$ есть старший коэффициент $f(x)$, то $f(x) := f(x)/a \pmod{p}$.

3. Вернуть $f(x)$.

7.6.3. Расширенный алгоритм Евклида для полиномов из $\mathbb{Z}_p[x]$

ВХОД. Ненулевые полиномы $f(x),h(x) \in \mathbb{Z}_p[x]$.

ВЫХОД. $d(x)$ = нод$(f(x),h(x))$ и два полинома $u(x),v(x) \in \mathbb{Z}_p[x]$, для которых $d(x) = u(x)f(x) + v(x)h(x)$.

1. Если $h = 0$ то $d:=f$, $u:=1$, $v:=0$, вернуть (d,u,v).
2. $u_2:=1$, $u_1:=0$, $v_2:=0$, $v_1:=1$.
3. Пока $h \neq 0$ выполнять следующее.
   3.1. $q:=\lfloor f/h \rfloor$, $r:=f-hq$, $u:=u_2-qu_1$, $v:=v_2-qv_1$.
   3.2. $f:=h$, $h:=r$, $u_2:=u_1$, $u_1:=u$, $v_2:=v_1$, $v_1:=v$.
4. $d:=f$, $u:=u_2$, $v:=v_2$.
5. Если $p>2$, $a\neq 1$ есть старший коэффициент $f(x)$, то $f:=f\cdot a^{-1} \pmod p$, $u:=u\cdot a^{-1} \pmod p$, $v:=v\cdot a^{-1} \pmod p$.
6. Вернуть (d,u,v)

Пример 1. (Расширенный алгоритм Евклида для полиномов из $\mathbb{Z}_2[x]$).

Пусть $f(x)=x^{10}+x^9+x^8+x^6+x^5+x^4+1$, $h(x)=x^9+x^6+x^6+x^3+x^2+1$ есть полиномы из $\mathbb{Z}_2[x]$. Найти $d(x)$=нод$(f(x), h(x))$ и полиномы $u(x),v(x) \in \mathbb{Z}_2[x]$, для которых

$d(x) = u(x)f(x) + v(x)h(x)$.

Решение.

Исходное присваивание.

$u_2(x) := 1$, $u_1(x) := 0$, $v_2(x) := 0$, $v_1(x) := 1$.

Итерация 1.

$q(x) := \lfloor f(x)/h(x) \rfloor = x+1$,
$r(x) := f(x)-h(x)q(x) = x^8+x^7+x^6+x^2+x$,
$u(x) := u_2(x)-q(x)u_1(x)=1$, $v(x) := v_2(x)-q(x)v_1(x)=x+1$,
$f(x) := h(x)=x^9+x^6+x^5+x^3+x^2+1$, $h(x) :=r(x)=x^8+x^7+x^6+x^2+1$,
$u_2(x) := u_1(x)=0$, $u_1(x) := u(x)=1$,
$v_2(x) := v_1(x)=1$, $v_1(x) := v(x)=x+1$.

Итерация 2.

$q(x) := \lfloor f(x)/h(x) \rfloor = x+1$,
$r(x) := f(x)-h(x)q(x) = x^5+x^2+x+1$,
$u(x) := x+1$, $v(x) := x^2$,
$f(x) := h(x) = x^8+x^7+x^6+x^2+1$, $h(x) := r(x) = x^5+x^2+x+1$,
$u_2(x) := u_1(x) = 1$, $u_1(x) := u(x) = x+1$,
$v_2(x) := v_1(x) = x+1$, $v_1(x) := v(x) = x^2$.

Итерация 3.

$q(x) := \lfloor f(x)/h(x) \rfloor = x^3+x^2+x+1$,
$r(x) := f(x)-h(x)q(x) = x^3+x+1$,
$u(x) := x^4$, $v(x) := x^5+x^4+x^3+x^2+x+1$,
$f(x) := h(x) = x^5+x^2+x+1$, $h(x) := r(x) = x^3+x+1$,
$u_2(x) := u_1(x) = x+1$, $u_1(x) := u(x) = x^4$,
$v_2(x) := v_1(x) = x^2$, $v_1(x) := v(x) = x^5+x^4+x^3+x^2+x+1$.

Итерация 4.

$q(x) := \lfloor f(x)/h(x) \rfloor = x^2+1$, $r(x) := f(x)-h(x)q(x) = 0$,
$u(x) := x^6+x^4+x+1$, $v(x) := x^7+x^6+x^2+x+1$,
$f(x) := h(x) = x^3+x+1$, $h(x) := r(x) = 0$,
$u_2(x) := u_1(x) = x^4$, $u_1(x) := u(x) = x^6+x^4+x+1$,
$v_2(x) := v_1(x)=x^5+x^4+x^3+x^2+x+1$, $v_1(x):=v(x)=x^7+x^6+x^2+x+1$.

Ответ. $d(x)$ = нод$(f(x),h(x)) = x^3+x+1$,
$u(x)=x^4)$, $v(x)=x^5+x^4+x^3+x^2+x+1$.

Пример 2. (Расширенный алгоритм Евклида для полиномов из $\mathbb{Z}_5[x]$).

Пусть $f(x)=4x^8+3x^7+4x^6+2x^5+4x^4+2x^3+4x^2+3x+4$, $h(x)=3x^7+x^6+4x^4+3x^3+x^2+4x+4$ есть полиномы из $\mathbb{Z}_5[x]$. Найти $d(x)$=нод$(f(x), h(x))$ и полиномы $u(x),v(x) \in \mathbb{Z}_5[x]$, для которых

$d(x) = u(x)f(x) + v(x)h(x)$.

Решение.

Исходное присваивание.

$f(x)=4x^8+3x^7+4x^6+2x^5+4x^4+2x^3+4x^2+3x+4$,
$h(x)=3x^7+x^6+4x^4+3x^3+x^2+4x+4$
$u_2(x):=1$, $u_1(x):=0$, $v_2(x):=0$, $v_1(x):=1$.

Итерация 1.

$q(x):=\lfloor f(x)/h(x) \rfloor = \lfloor (4x^8+3x^7+4x^6+2x^5+4x^4+2x^3+4x^2+3x+4)/(3x^7+x^6+4x^4+3x^3+x^2+4x+4) \rfloor = 3x$,
$r(x):=f(x)-h(x)q(x)=4x^6+4x^3+2x^2+x+4$,
$u(x):=u_2(x)-q(x)u_1(x)=1$, $v(x):=v_2(x)-q(x)v_1(x)=2x$,
$f(x):=h(x)=3x^7+x^6+4x^4+3x^3+x^2+4x+4$,
$h(x):=r(x)=4x^6+4x^3+2x^2+x+4$,
$u_2(x):=u_1(x)=0$, $u_1(x):=u(x)=1$,
$v_2(x):=v_1(x)=1$, $v_1(x):=v(x)=2x$.

Итерация 2.

$q(x):=\lfloor f(x)/h(x) \rfloor = \lfloor (3x^7+x^6+4x^4+3x^3+x^2+4x+4)/(4x^6+4x^3+2x^2+x+4) \rfloor = 2x+4$,

$r(x):=f(x)-h(x)q(x)=x^4+3x^3+x^2+2x+3$,
$u(x):=u_2(x)-q(x)u_1(x)=3x+1$, $v(x):=v_2(x)-q(x)v_1(x)=x^2+2x+1$,
$f(x):=h(x)=4x^5+4x^3+2x^2+x+4$, $h(x):=r(x)=x^4+3x^3+x^2+2x+3$,
$u_2(x):=u_1(x)=1$, $u_1(x):=u(x)=3x+1$,
$v_2(x):=v_1(x)=2x$, $v_1(x):=v(x)=x^2+2x+1$.

Итерация 3.

$q(x):=\lfloor f(x)/h(x)\rfloor=\lfloor(4x^5+4x^3+2x^2+x+4)/(x^4+3x^3+x^2+2x+3)\rfloor=$
$4x^2+3x+2$,
$r(x):=f(x)-h(x)q(x)=2x^3+2x^2+3x+3$,
$u(x):=u_2(x)-q(x)u_1(x)=3x^3+2x^2+x+4$,
$v(x):=v_2(x)-q(x)v_1(x)=x^4+4x^3+3x^2+3$,
$f(x):=h(x)=x^4+3x^3+x^2+2x+3$, $h(x):=r(x)=2x^3+2x^2+3x+3$,
$u_2(x):=u_1(x)=3x+1$, $u_1(x):=u(x)=3x^3+2x^2+x+4$,
$v_2(x):=v_1(x)=x^2+2x+1$, $v_1(x):=v(x)=x^4+4x^3+3x^2+3$.

Итерация 4.

$q(x):=\lfloor f(x)/h(x)\rfloor=\lfloor(x^4+3x^3+x^2+2x+3)/(2x^3+2x^2+3x+3)\rfloor=$
$3x+1$,
$r(x):=0$,
$u(x):=u_2(x)-q(x)u_1(x)=x^4+x^2+2$,
$v(x):=v_2(x)-q(x)v_1(x)=2x^5+2x^4+2x^3+3x^2+3x+3$,
$f(x):=h(x)=2x^3+2x^2+3x+3$, $h(x):=r(x)=0$,
$u_2(x):=u_1(x)=3x^3+2x^2+x+4$, $u_1(x):=u(x)=x^4+x^2+2$,
$v_2(x):=v_1(x)=x^4+4x^3+3x^2+3$,
$v_1(x):=v(x)=2x^5+2x^4+2x^3+3x^2+3x+3$.

Так как $h(x)=0$, то $d(x):=f(x)=2x^3+2x^2+3x+3$, $u(x):=u_2(x)=3x^3+2x^2+x+4$, $v(x):=v_2(x)=x^4+4x^3+3x^2+3$.

Так как $p=5>2$ и $d(x)$ имеет старший коэффициент $a=2$, то
$d(x) := d(x)\cdot 2^{-1}(\mathrm{mod}\ p) = d(x)\cdot 3\ (\mathrm{mod}\ p) = x^3+x^2+4x+4$,
$u(x) := u(x)\cdot 2^{-1}(\mathrm{mod}\ p) = u(x)\cdot 3\ (\mathrm{mod}\ p) = 4x^3+x^2+3x+2$,
$v(x) := v(x)\cdot 2^{-1}(\mathrm{mod}\ p) = v(x)\cdot 3\ (\mathrm{mod}\ p) = 3x^4+2x^3+4x^2+4$.

Ответ. $d(x)=x^3+x^2+4x+4$, $u(x)=4x^3+x^2+3x+2$,
$v(x)=3x^4+2x^3+4x^2+4$.

7.6.4. Мультипликативный обратный элемент в $\mathbb{F}_{p^m}$

ВХОД. Ненулевой полином $g(x) \in \mathbb{F}_{p^m}$. (Элементы поля $\mathbb{F}_{p^m}$ представляются как элементы в $\mathbb{Z}_p[x]/(f(x))$, где $f(x) \in \mathbb{Z}_p[x]$ есть неприводимый полином степени m над $\mathbb{Z}_p$.)

ВЫХОД. $g(x)^{-1} \in \mathbb{F}_{p^m}$.

1. С помощью расширенного алгоритма Евклида для полиномов найти два полинома $u(x)$ и $v(x) \in \mathbb{Z}_p[x]$, для которых $u(x)g(x) + v(x)f(x) = 1$.
2. Вернуть $u(x)$.

7.6.5. Модулярная степень в $\mathbb{F}_{p^m}$

ВХОД. $g(x) \in \mathbb{F}_{p^m}$ и целое $0 \leqslant k < p^m-1$ с бинарным представлением $k = \sum_{i=0}^{t} k_i 2^i$. (Поле $\mathbb{F}_{p^m}$ есть $\mathbb{Z}_p[x]/(f(x))$, где $f(x) \in \mathbb{Z}_p[x]$ есть неприводимый полином степень m над $\mathbb{Z}_p$.)

ВЫХОД. $g(x)^k \pmod{f(x)}$.

1. $u(x) := 1$. Если $k = 0$, то вернуть $u(x)$.
2. $G(x) := g(x)$.
3. Если $k_0 = 1$, то $u(x) := g(x)$.
4. Для i от 1 до t выполнить следующее.
   4.1. $G(x) := G(x)^2 \pmod{f(x)}$.
   4.2. Если $k_i = 1$, то $u(x) := G(x)\cdot u(x) \pmod{f(x)}$.
5. Вернуть $u(x)$.

Утверждение. Пусть p есть простое число и пусть k есть положительное целое число.

1. Произведение всех нормированных неприводимых полиномов в $\mathbb{Z}_p[x]$, степень которых делит k, равно $x^{p^k} - x$.
2. Пусть $f(x)$ есть полином степени m из $\mathbb{Z}_p[x]$. Тогда $f(x)$ неприводим над $\mathbb{Z}_p$, если и только если $\mathrm{нод}(f(x), x^{p^i}-x) = 1$ для каждого i, $1 \leqslant i \leqslant \lfloor m/2\rfloor$.

7.6.6. Тестирование полинома из $\mathbb{Z}_p[x]$ на неприводимость

ВХОД. Простое число p и нормированный полином $f(x)$ степени m из $\mathbb{Z}_p[x]$.

ВЫХОД. Ответ на вопрос: "Является ли полином $f(x)$ неприводим над $\mathbb{Z}_p$?"

1. $u(x) := x$.
2. Для i от 1 до $\lfloor m/2\rfloor$ выполнить следующее.
   2.1. $u(x) := u(x)^p \pmod{f(x)}$.
   2.2. $d(x) := \mathrm{нод}(f(x), u(x)-x)$.
   2.3. Если $d(x) \neq 1$, то вернуть "приводимый".
3. Вернуть "неприводимый".

7.6.7. Порождение случайного неприводимого полинома над $\mathbb{Z}_p$

ВХОД. Простое число p и положительное целое m.

ВЫХОД. Неприводимый полином $f(x)$ степени m в $\mathbb{Z}_p[x]$.

1. Случайно выбираем целые $a_0, a_1, \ldots a_{m-1}$ между 0 и $p-1$ с $a_0 \neq 0$. Пусть $f(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_2x^2 + a_1x + a_0$.
2. Тестируем полином $f(x)$ на неприводимость.
   Если полином $f(x)$ приводим над $\mathbb{Z}_p$, перейти к 1.
3. Вернуть $f(x)$.

7.6.8. Тестирование неприводимого полинома на примитивность

ВХОД. Простое число p; целое $m \geq 1$; различные простые делители $r_1, r_2, \ldots, r_t$ числа $p^m - 1$; нормированный неприводимый полином $f(x)$ степени m в $\mathbb{Z}_p[x]$.

ВЫХОД. Ответ на вопрос: "Примитивен ли полином $f(x)$?"

1. Для i от 1 до t выполнить следующее.

   1.1. $l(x) := x^{(p^m-1)/r_i} \pmod{f(x)}$.
   1.2. Если $l(x) = 1$, то вернуть "Непримитивный".
2. Вернуть "Примитивный").

7.6.9. Порождение случайного нормированного примитивного полинома над $\mathbb{Z}_p$

ВХОД. Простое число p; целое $m \geq 1$; различные простые делители $r_1, r_2, \ldots, r_t$ числа $p^m - 1$.

ВЫХОД. Нормированный примитивный полином $f(x)$ степени m в $\mathbb{Z}_p[x]$.

1. Генерируем случайный нормированный неприводимый полином $f(x)$ степени m в $\mathbb{Z}_p[x]$.
2. Тестируем полином $f(x)$ на примитивность.
   Если полином $f(x)$ не примитивен над $\mathbb{Z}_p$, перейти к 1.
3. Вернуть $f(x)$.

7.6.10. Вычисление порядка элемента конечной группы (метод Гаусса)

ВХОД. Мультипликативная конечная группа G порядка n, элемент $a \in G$, факторизация $n = p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k}$.

ВЫХОД. Порядок t элемента a.

1. $t := n$.
2. Для i от 1 до k выполнить следующее.

   2.1. $t := t/p_i^{e_i}$.
   2.2. $a_1 := a^t$.

   2.3. Пока $a_1 \neq 1$, выполнить: $a_1 := a_1^{p_i}$, $t := t \cdot p_i$.

Вернуть t.

7.6.11. Вычисление генератора конечной циклической группы (метод Гаусса)

ВХОД. Циклическая конечная группа G порядка n, факторизация $n = p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k}$.

ВЫХОД. Генератор a для G.

1. Выбрать случайный элемент a в G.
2. Для i от 1 до k выполнить следующее.

   2.1. $b := a^{n/p_i}$.
   2.2. Если $b=1$, то перейти к пункту 1.

Вернуть a.

8. ПРИМЕНЕНИЕ МОДУЛЯРНОЙ АРИФМЕТИКИ В КРИПТОГРАФИИ

8.1. Криптография и ее цели

У криптографии длиная история. Есть примеры ее использования еще в древнем Египте 4000 назад. В наше время ее значение в вопросах защиты информации от несанкционированного доступа очень велико. Сначала основное ее применение было связано с государственными и военными секретами. Сейчас это защита информации в повсеместно распространенных действующих компьютерных сетях, где обращается огромный объем конфиденциальной информации.

В 1978 году Rivest, Shamir и Adleman предложили схемы шифрования и цифровой электронной подписи с открытым ключом, известном как RSA. Криптографическая стойкость схемы обеспечивается трудностью (практически почти невозможностью) решения задачи факторизации больших чисел (около 120 цифр в десятиричной записи числа). Позже были предложены схемы шифрования и электронной подписи, основаные на трудности решения проблемы дискретного логарифма и проблемы квадратичного вычета. Были найдены и другие удобные схемы шифрования, основанные на трудности решения других проблем.

Определим некоторые основные понятия криптография.

Определение. Функция $f: X \to Y$ называется *односторонней функцией*, если $f(x)$ "легко" вычислима для всех $x \in X$, но

трудно разрешима обратная задача: для данного $y \in Y = \mathrm{Im}(f)$ найти такое $x \in X$, для которого $f(x) = y$.

Пример. 1. $f(x) = \mathrm{rest}(3^x, m) = 3^x \pmod m$. Пусть модуль m есть целое число размера около 120 десятичных цифр. Значение $f(x)$ вычисляется достаточно быстро, в то время как нахождение $x = f^{-1}(y) \pmod m$ значительно труднее. $3^x \pmod m$ есть одностороняя функция.

2. Пусть факторизация натурального числа $n = pq$. Легко найти $n = f(p,q) = pq$ для всяких простых чисел p и q. Но много труднее найти факторизацию n, даже если известно, что n есть произведение двух простых чисел. $f(p,q)$ есть односторонняя функция.

Определение. *Одностороняя функция с порогом* есть односторонняя функция $f: X \to Y$ с дополнительной информацией, называемой *порогом*, которая делает возможным нахождение $x = f^{-1}(y)$ $\forall y \in Y$.

Пример. Пусть односторонняя функция $n=f(p,q)=pq$ есть произведение двух больших простых натуральных чисел p и q, размер каждого около 100 десятичных цифр цифр. Если p известно, то факторизация n легко осуществима. Число p есть пороговая информация.

Определение. *Криптография* есть применение математической техники в таких вопросах как защита информации от несанкционированного доступа, целостность информации, аутентификация адресата, аутентификация исходной информации. *Криптология* есть изучение криптографии. *Криптосистема* есть система, обеспечивающая защиту информации.

Замечание. Часто термин криптосистема используется в связи с шифрованием.

Различают два типа криптографической техники: с симметричным ключом и с открытым ключом.

Криптография при обмене информацией обеспечивает следующие цели.

1. Конфиденциальность, то есть секретность, или защита информации от несанционированного доступа.
2. Целостность информации, то есть защита информации от несанкционированного изменения.
3. Аутентификация, или идентификация адресата.
4. Неотречение (невозможность отречения) адресата от ранее принятых на себя обязательств.

Определение. *Алфавит* A есть непустое конечное множество

символов. *Слово* в алфивите A есть конечная последовательность символов из A. A^* есть множество всех слов в алфавите A. *Сообщение* m есть конечная последовательность слов из A^*. *Пространство сообщений* есть множество M всех сообщений в алфавите A. *Исходное сообщение* (или *исходный текст*) есть сообщение m из M. *Пространство шифротекстов* C есть множество сообщений в некотором алфавите, отличном от алфавита A. *Шифротекст* c есть элемент из пространства C. *Функция шифрования* (или *преобразование шифрования*) есть биекция $e: M \to C$. *Функция дешифрования* (или *преобразование дешифрования*) есть биекция $d: C \to M$. При этом $e^{-1} = d$ и $d^{-1} = e$. *Шифрование исходного текста* m есть вычисление шифротекста $c=e(m)$. *Дешифрование шифротекста* c есть вычисление исходного текста $m = d(c)$. *Ключ* есть функция шифрования или функция дешифрования. *Пространство ключей* есть множество всех ключей. *Схема шифрования* (или *шифр*) есть пара (e,d), где e есть функция шифрования и d есть соответствующая функция дешифрования. *Построить шифр* значит построить объект (M,C,e,d).

Замечание. Биекция есть взаимно однозначная функция функция $f: X \to Y$, для которой $\mathrm{Im}(f) = Y$.

Определение. *Адресат* есть санкционированный отправитель или получатель информации. *Противник* перехватыватель информации с целью ее дешифрования, порчи или искажения.

Определение. *Канал* есть средство передачи информации от одного адресата к другому. (*Физически*) *безопасный канал* есть канал, недоступный противнику для воздействия. *Небезопасный канал* есть канал, подверженный воздействию противника.

Определение. Схема шифрования *криптографически стойка* (или *безопасна*), если третья сторона (противник) без знания ключа (e,d) *не* может в приемлимое время расшифровать любой шифротекст схемы. В противном случае схема шифрования *криптографически не стойка* (то есть взламываема).

Существуют шифры с симметричным ключом и шифры с открытым ключом.

Определение. Шифр (e,d) есть *шифр с симметричным ключом* (или короче – симметричный шифр) , если ключ (e,d) "легко" вычислим по одной из двух компонент ключа.

Пример. (Симметричный ключ). Пусть $\mathcal{A} = \{a,b,c,\ldots,x,y,z\}$ есть буквы английского алфавита. Пусть M и C есть множество

всех слов длины пять в алфавите $\mathcal{A}$. Ключ e есть перестановка букв алфавита $\mathcal{A}$. Шифрование производится следующим образом. Сообщение разбивается на группы по пять букв (с условленным дописыванием букв, если длина сообщения не кратно пяти) и перестановка e применяется к каждой букве. При дешифровании обратная перестановка $d = e^{-1}$ применяется к каждой букве шифротекста. Например, пусть ключ $e: \mathcal{A} \to \mathcal{A}$ задается перестановкой

$$e = \begin{pmatrix} a & b & c & d & e & f & g & h & i & j & k & l & m & n & o & p & q & r & s & t & u & v & w & x & y & z \\ d & e & f & g & h & i & j & k & l & m & n & o & p & q & r & s & t & u & v & w & x & y & z & a & b & c \end{pmatrix}.$$

Сообщение m = thisc ipher iscer tainl ynots ecure шифруется в сообщение $c = e(m)$ = wklvf lskhu lvfhu wdlqo bqrwv hfxuh и дешифруется в $m = e^{-1}(c)$ с помощью обратной подстановки e^{-1}.

Замечание. *Проблема распределения ключей* при использовании шифра с симметричным ключом есть задача (тайной) передачи ключей лицам, обменивающимся информацией.

Определение. *Блочный шифр* есть схема шифрования, при которой исходное сообщение разбивается на блоки фиксированной длины t и шифруется блоками один за другим.

Определение. *Потоковый шифр* есть схема шифрования, при которой исходное сообщение поступает и шифруется побуквенно, (посимвольно) одна буква за другой.

Пример. *Шифр Вернама* есть потоковый шифр, действующий в алфавите $\mathcal{A}=\{0,1\}$ следующим образом. Пусть $m_1 m_2 \ldots m_t$ есть бинарное сообщение и $k_1 k_2 \ldots k_t$ есть фиксированное бинарное слово. Тогда шифротекст есть бинарное слово $c_1 c_2 \ldots c_t$, где

$$c_i = m_i + k_i \pmod 2, \quad 1 \leqslant i \leqslant t.$$

При дешифровании $m_i = c_i + k_i \pmod 2$, $1 \leqslant i \leqslant t$.

Определение. Шифр (e,d) есть *схема шифрования с открытым ключом*, если функция шифрования e *публикуется* (всем известна, открыта). Соответствующая функция дешифрования d в приемлимое время не вычисляется.

Шифры с симметричным ключом и с открытым ключом имеют свои преимущества и недостатки. Отметим некоторые из них.

Преимущества шифров с симметричным ключом. 1. Симметричные шифры обрабатывают (шифруют и дешифруют) информацию достаточно быстро. Некоторые аппаратные реализации шифров шифруют информацию со скоростью сотен мегабайт в секунду.

Программная реализация может достигать скорости не-скольких мегабайт в секунду.

2. Ключи симметричных шифров сравнительно короткие.

3. Различные симметричные шифры можно соединять вместе и получать еще более быстрые шифры.

Недостатки шифров с симметричным ключом. 1. Участники связи должны тайно обменяться ключом по некоторым каналам, на которые может воздействовать противник.

2. Ключ надо часто менять и возможно даже, что для каждого сеанса связи.

Преимущества шифров с открытым ключом. 1. Только часть ключа должна храниться в секрете.

2. Секретной частью ключа участники секретной связи не обмениваются.

3. Ключ может оставаться неизменным длительное время, иногда годами.

4. Шифры с открытым ключом можно использовать для передачи заинтересованным лицам ключей симметричных шифров.

5. Шифры с открытым ключом имеют варианты электронной подписи.

Недостатки шифров с открытым ключом. 1. Шифры с открытым ключом работают значительно медленнее симметричных шифров.

2. Размеры ключей шифров с открытым ключом много длиннее ключей симметричных шифров.

3. В настоящее время нет доказательства криптографической стойкости шифров с открытым ключом (то же самое для блочных шифров).

8.1.1. Хэш-функция

Заметим, что содержимое любого файла или его части есть некоторый текст t, составленный из символов клавиатуры компьютера и представляемый в компьютере как последовательность нулей и единиц, то есть как бинарный стринг m (бинарное слово в алфавите $\{0,1\}$). Всякий бинарный стринг можно рассматривать как запись m_2 в двоичной системе счисления натурального числа m, которое, в свою очередь, можно представить записью m_h в системе счисления по любому другому основанию h.

Между множеством всех компьютерных текстов, между множеством их бинарных стрингов в компьютере, между множеством представляемых ими чисел существует взаимно однозначное со-

ответствие. Любое из этих представлений информации в компьютере мы будем обобщенно называть текстом. В последующем из контекста всегда будет ясно, о каком представлении информации в компьютере идет речь: о тексте ли t, составленным из символов клавиатуры компьютера, о бинарном ли его стринге m_2, о натуральном ли его h-ричном числе m_h.

Хэш-функция $w=h(x)$ есть достаточно быстро (по времени) вычислимая функция, отображающая всякий исходный бинарный стринг (текст) x произвольной (разумной) длины в бинарный стринг (число) $w=h(x)$, называемый хэш-значением w (или просто хэш), который выступает как компактный представитель (паспорт) входного слова x. Хэш-функция $w=h(x)$ такова, что 1) ее значение w "легко" вычисляется для всякого текста x, но практически (в разумное время) невозможно определить исходный текст $x = h^{-1}(w)$ по его хэш-значению w; 2) практически невероятно найти два различных входных текста с одним и тем же хэш-значением, то есть два различных текста x и y, для которых $h(x) = h(y)$.

С помощью хэш-функции тексту x ставится в соответствие уникальное число $w = h(x)$.

Алгоритм вычисления значений хэш-функции публикуется.

Значение хэш-функции есть большое число, выходящее за пределы величин целых чисел, допустимых в алгоритмических языках программирования. Mathcad, например, допускает целые (10-ричные) числа длины не более 18 цифр. Для работы с большими целыми числами с длиной десятиричной записи в 100 и более цифр приходится писать специальный программный процессор. Поэтому в последующих примерах значение хэш-функции задается искусственно, для примера, небольшим числом.

Предложено много примеров хэш-функций. Приведем один из них, основанный на модулярной арифметике целых чисел. Это MASH (Modular Arithmetic Secure Hash).

8.1.2. Алгоритм MASH-1

ВХОД. Стринг x битовой длины b, $0 \leqslant b < 2^{n/2}$.

ВЫХОД. n-битовый хэш для x (n приблизительно равно битовой длине модуля M).

1. Зададим модуль $M=p\cdot q$ битовой длины m, где p и q есть случайно выбранные большие секретные простые числа примерно одного размера так, чтобы факторизация M была трудно осуществима. Определим битовую длину n хэш-результата равной наи-

большему кратному 16 меньшему m (то есть $n = 16n' < m$). Пусть H_0=0 есть начальное значение. Пусть n-битовая константа A=11110...0. Пусть знак "$\vee$" означает побитовую дизъюнкцию OR, а знак "$\oplus$" побитовую исключающую дизъюнкцию XOR (сложение по модулю 2).

2. Допишем x 0-битами, если это необходимо, чтобы получить стринг битовой длины $t\cdot(n/2)$ для наименьшего $t>1$. Разобьем полученный после дописывания текст на $(n/2)$-битовые блоки $x_1,\ldots,x_t$ и добавим конечный блок x_{t+1}, содержащий $(n/2)$-битовое представление для b.

3. Расширим каждое x_i до n-битового блока y_i разбиением его на (4-битовые) отрезки и вставляя четыре 1-бита перед каждым, кроме y_{t+1}, для которого вставляемый отрезок есть 1010 (не 1111).

4. Для $1\leqslant i\leqslant t+1$, отобразим два n-битовых входа (H_{i-1},y_i) в один n-битовый выход следующим образом.

$$H_i := ((((H_{i-1}\oplus y_i)\vee A)^2 \pmod M) \vdash n) \oplus H_{i-1}.$$

Здесь знак "$\vdash$" означает перенос крайних справа n бит m-битового результата налево.

5. Хэш-значение есть n-битовый блок H_{t+1}.

Замечание. MASH-2 отличается от MASH-1 только тем, что в H_i степень e=2 в MASH-1 заменяется в MASH-2 на $e=2^8+1=257$.

8.2. Проблема факторизации целых чисел

Перечислим несколько общих проблем, на трудности решения которых основывается криптографическая стойкость криптосистем с открытым ключом.

Криптографическая стойкость (к взлому) многих криптосистем основывается на трудности решения проблемы факторизации целых чисел. В их числе криптосистема RSA, схема RSA цифровой подписи, криптосистема Рабина, другие криптосистемы.

Определение. *Каноническая факторизация целого положительного числа* n есть представление n в виде $n = p_1^{e_1}p_2^{e_2}\ldots p_k^{e_k}$, где все p_i есть попарно различные простые числа и каждое $e_i\geqslant 1$. *Нетривиальная факторизация целого положительного* n есть его представление в виде $n = ab$, где $1 < a,b < n$. Числа a и b есть нетривиальные факторы в n. Факторы 1 и n тривиальны.

Предположение. Не существует полиномиального алгоритма решения проблемы факторизации целых чисел.

8.2.1. *Rho-алгоритм Полларда факторизации целых чисел*

ВХОД. Составное целое n, которое не есть степень простого числа.

ВЫХОД. Нетривиальный фактор d в n.

1. $a := 2$, $b := 2$.
2. Для i=1,2,... выполнить следующее.
   2.1. $a := a^2+1 \pmod n$, $b := b^2+1 \pmod n$, $b := b^2+1 \pmod n$.
   2.2. d := нод$(a-b,\ n)$.
   2.3. Если $1 < d < n$, то вернуть d.
   2.4. Если $d = n$, то алгоритм заканчивает работу и вопрос о нетривиальных факторах в n остается открытым.

Замечание. Rho–алгоритм Полларда для факторизации числа n требует $O(n^{1/4})$ модулярных умножений.

Пример. Пусть n = 455459. Результаты вычислений по *rho*-алгоритму Полларда приведены в следующей таблице.

a	b	d
5	26	1
26	2871	1
677	179685	1
2871	155260	1
44380	416250	1
179685	43670	1
121634	164403	1
155260	247944	1
44567	68343	743

Следовательно, 743 и 455459/743 = 613 есть два нетривиальных делителя числа 455459.

8.2.2. (p–1)*-алгоритм Полларда факторизации целых чисел*

Определение Пусть B есть положительное целое число. Целое число n есть B-гладкое число, или гладкое относительно границы B, если все простые делители $n \leqslant B$.

(p–1)–алгоритм Полларда факторизации есть алгоритм, для эффективного вычисления всякого простого фактора p составного целого n, для которого p–1 есть гладкое число

относительно некоторой сравнительно малой границы B.

ВХОД. Составное целое n, которое не есть степень простого числа.

ВЫХОД. Нетривиальный фактор d в n.

1. Выбрать гладкую границу B, меньшую чем n.
2. Выбрать случайное целое a, $2 \leqslant a \leqslant n$–1, и вычислить d = нод(a,n). Если $d \geqslant 2$, то вернуть d.
3. Для каждого простого $q \leqslant B$ выполнить следующее.
   3.1. $l := \lfloor (\ln n)/(\ln q) \rfloor$.
   3.2. $a := a^{q^l} \pmod n$.
4. d := нод$(a$–1, $n)$.
5. Если d=1 или d=n, то вопрос о нетривиальном факторе остается открытым (взять большее B). Иначе вернуть d.

Пример. Пусть n = 19048567. (p–1)–алгоритм Полларда для нахождения нетривиального фактора для n)

1. Выбрать границу гладкости B = 19.
2. Выбрать целое a=3 и вычислить нод$(3,n)$ = 1.
3. В следующей таблице приведены значения q,l,a после каждой итерации пункта 3 (p–1)–алгоритма Полларда.

q	l	a
2	24	2293244
3	15	13555889
5	10	16937223
7	8	15214586
11	6	9685355
13	6	13271154
17	5	11406961
19	5	554506

4. d := нод(554506–1, n) = 5281.
5. Два нетривиальных фактора в n есть p=5281 и $q := n/p$ = 3607 (эти факторы есть простые числа).

Замечание. p–1 = 5280 = $2^5 \cdot 3 \cdot 5 \cdot 11$, q–1 = 3606 = $2 \cdot 3 \cdot 601$. Число p–1 есть 19–гладкое число (ибо 2,3,5,11 $\leqslant$ 19 и B=19 угадано верно), в то время как q–1 не есть 19–гладкое число.

8.2.3. *Алгоритм квадрат–решета факторизации целых чисел*

ВХОД. Составное целое n, которое не есть степень простого числа.

ВЫХОД. Нетривиальный фактор d в n.

1. Выбрать факторный базис $S = \{p_1, p_2, \ldots, p_t\}$, где $p_1 = -1$ и p_j $(j \geqslant 2)$ есть $(j-1)$-ое простое p, для которого n есть квадратичный вычет по модулю p.

2. $m = \lfloor \sqrt{n} \rfloor$.

3. (Выбор $t+1$ пар (a_i, b_i). Значения x выбираются в порядке $0, \pm 1, \pm 2, \ldots$).

$i := 1$.

Пока $i \leqslant t+1$, выполнить следующее.

3.1. $b := q(x) = (x+m)^2 - n$.

Факторизовать b, взять p_t в S и проверить, является ли b p_t-гладким. Если нет, выбрать новое x и повторить шаг 3.1.

3.2. Если b является p_t-гладким, например,

$b = \Pi_{j=1}^{t} p_j^{e_{ij}}$, то

$a_i := x+m$, $b_i := b$, $v_i := (v_{i1}, v_{i2}, \ldots, v_{it})$,

где $v_{ij} = e_{ij} \pmod 2$ для $1 \leqslant j \leqslant t$.

3.3. $i := i+1$.

4. Методами линейной алгебры над $\mathbb{Z}_2$ найти непустое подмножество $T \subseteq \{1, 2, \ldots, t+1\}$, для которого $\sum_{i \in T} v_i = 0$.

5. $x := \Pi_{i \in T} a_i \pmod n$.

6. Для каждого j, $1 \leqslant j \leqslant t$,

$l_j := (\sum_{i \in T} e_{ij})/2$.

7. $y := \Pi_{j=1}^{t} p_j^{l_j} \pmod n$.

8. Если $x \equiv \pm y \pmod n$, то найти другое непустое подмножество $T \subseteq \{1, 2, \ldots, t+1\}$, для которого $\sum_{i \in T} v_i = 0$ и перейти к пункту 5.

(В случае (маловероятного) несуществования такого подмножества T, заменить некоторые из пар (a_i, b_i) новыми парами (пункт 3) и перейти к пункту 4.)

9. $d := \text{нод}(x-y, n)$ и вернуть d.

Пример. Пусть $n = 24961$.

1. Выберем факторный базис $S = \{-1, 2, 3, 5, 13, 23\}$ размера $t = 6$. (Числа 7,11,17,19 опущены в S, ибо символ Лежандра $\left(\frac{n}{p}\right) = -1$ для этих простых чисел.)

2. $m := \lfloor \sqrt{24961} \rfloor = 157$.

3. В следующей таблице приведены данные, собранные для первых $t+1$ значений x, для которых $q(x)$ есть 23-гладкое.

i	x	$q(x)$	факторизация $q(x)$	a_i	v_i
1	0	−312	$-2^3 \cdot 3 \cdot 13$	157	(1,1,1,0,1,0)
2	1	3	3	158	(0,0,1,0,0,0)
3	−1	−625	-5^4	156	(1,0,0,0,0,0)
4	2	320	$2^6 \cdot 5$	159	(0,0,0,1,0,0)
5	−2	−936	$-2^3 \cdot 3^2 \cdot 13$	155	(1,1,0,0,1,0)
6	4	960	$2^6 \cdot 3 \cdot 5$	161	(0,0,1,1,0,0)
7	−6	−2160	$-2^4 \cdot 3^3 \cdot 5$	151	(1,0,1,1,0,0)

4. $v_1+v_2+v_5=0$. (В обозначениях алгоритма $T = \{1,2,5\}$).

5. $x := a_1 a_2 a_5 \pmod n = 936$.

6. $l_1=1$, $l_2=3$, $l_3=2$, $l_4=0$, $l_5=1$, $l_6=0$.

7. $y := -2^3 \cdot 3^2 \cdot 13 \pmod n = 24025$.

8. Так как $936 \equiv -24025 \pmod n$, то надо найти другую линейную зависимость.

9. $v_3+v_6+v_7 = 0$; $T = \{3,6,7\}$.

10. $x := a_3 a_6 a_7 \pmod n = 23405$.

11. $l_1=1$, $l_2=5$, $l_3=2$, $l_4=3$, $l_5=0$, $l_6=0$.

12. $y := -2^5 \cdot 3^2 \cdot 5^3 \pmod n = 13922$.

13. $23405 \not\equiv \pm 13922 \pmod n$. Поэтому вычисляем $\text{нод}(x-y, n) = \text{нод}(9483, 24961) = 109$. Следовательно, два нетривиальных фактора в 24961 есть 109 и 229.

8.3. Проблема RSA

Трудность решения RSA проблемы лежит в основе криптографической стойкости криптосистемы RSA и схемы цифровой подписи RSA.

Определение. RSA *проблема* состоит в следующем. Дано 1) положительное целое число n, которое есть произведение двух различных нечетных простых чисел p и q примерно одного размера, 2) положительное целое e, для которого $\text{нод}(e, (p-1)\cdot(q-1)) = 1$, 3) целое c. Найти целое число m, для которого $m^e \equiv c \pmod n$.

Другими словами, RSA проблема есть нахождение $m = \sqrt[e]{c} \pmod n$. При выше указанных условиях для параметров n и e для каждого целого $c \in \{0,1,\ldots,n-1\}$ существует в точности одно $m \in \{0,1,\ldots,n-1\}$, для которого $m^e \equiv c \pmod n$. То есть функция $f(x) = m^e$ типа $\mathbb{Z}_n \to \mathbb{Z}_n$ есть подстановка элементов множества $\mathbb{Z}_n$.

Замечание. Если факторы в n известны, то RSA проблема легко решается. То есть RSA проблема полиномиально сводима к проблеме факторизации целых чисел. Поэтому RSA проблема не менее трудна, чем проблема факторизации целых чисел.

Предположение. RSA проблема и проблема факторизации целых чисел вычислительно эквивалентны.

Предположение. Не существует полиномиального алгоритма решения RSA проблемы.

8.4. Проблема квадратичного вычета

Криптографическая стойкость некоторых криптосистем (например, криптосистемы Голдвассера–Микали) основывается на трудности решения проблемы квадратичного вычета.

Определение. *Проблема квадратичного вычета* состоит в вычислении $\sqrt{a} \pmod n$ по данному нечетному составному положительному целому числу n и данному числу a из $\mathbb{Z}_n^*$.

Замечание. Если n есть простое число p, то проблема квадратичного вычета легко решается с помощью символа Лежандра $\left(\frac{a}{p}\right)$. Если известна факторизация составного числа n, то проблема квадратичного вычета легко решается с помощью символа Якоби $\left(\frac{a}{n}\right)$. Можно показать что проблема квадратичного вычета полиномиально сводится к проблеме факторизации целых чисел. Поэтому проблема квадратичного вычета не менее трудна, чем проблема факторизации целых чисел.

Предположение. Проблема квадратичного вычета и проблема факторизации целых чисел вычислительно эквивалентны.

Предположение. Не существует полиномиального алгоритма решения проблемы квадратичного вычета.

8.4.1. Алгоритм вычисления квадратного корня по простому модулю p

ВХОД. Нечетное простое p и целое $a \in [1, p-1]$ по модулю p.

ВЫХОД. Два квадратных корня из a по модулю p, если a есть квадратичный вычет по модулю p.

1. Вычислить символ Лежандра $\left(\frac{a}{p}\right)$. Если $\left(\frac{a}{p}\right) = -1$ то

вернуть: $\sqrt{a} \pmod p$ не существует и закончить работу.

2. Выбрать случайное целое $b \in [1, p-1]$, для которого $\left(\frac{b}{p}\right) = -1$. ($b$ есть квадратичный невычет по модулю p.)
3. Повторным делением на 2 представить $p-1$ в виде $2^s t$, где число t нечетно.
4. С помощью расширенного алгоритма Евклида вычислить $a^{-1} \pmod p$.
5. $c := b^t \pmod p$ и $r := a^{(t+1)/2} \pmod p$.
6. Для i от 1 до $s-1$ выполнить следующее.
   6.1. $d := (r^2 \cdot a^{-1})^{2^{s-i-1}} \pmod p$.
   6.2. Если $d \equiv -1 \pmod p$, то $r := r \cdot c \pmod p$.
   6.3. $c := c^2 \pmod p$.
7. Вернуть $r, -r$.

8.4.2. Алгоритм вычисления квадратного корня по простому модулю p, где p ≡ 3 (mod 4)

ВХОД. Нечетное простое p, где $p \equiv 3 \pmod 4$, и a есть квадратичный вычет по модулю p.

ВЫХОД. Два квадратных корня из a по модулю p.

1. $r := a^{(p+1)/4} \pmod p$.
2. Вернуть $r, -r$.

8.4.3. Алгоритм вычисления квадратного корня по простому модулю p, где p ≡ 5 (mod 8)

ВХОД. Нечетное простое p, где $p \equiv 5 \pmod 8$, и a есть квадратичный вычет по модулю p.

ВЫХОД. Два квадратных корня из a по модулю p.

1. $d := a^{(p-1)/4} \pmod p$.
2. Если $d = 1$, то $r := a^{(p+3)/8} \pmod p$.
3. Если $d = p-1$, то $r := 2a(4a)^{(p-5)/8} \pmod p$.
4. Вернуть $r, -r$.

8.4.4. Алгоритм вычисления квадратного корня по простому модулю p при большом s

Этот алгоритм предпочтительнее прежнего алгоритма, если $p-1 = 2^s t$ с большим s.

ВХОД. Нечетное простое p и a есть квадратичный вычет

по модулю p.

ВЫХОД. Два квадратных корня из a по модулю p.

1. Выбрать случайное $b \in \mathbb{Z}_p$, для которого b^2-4a есть квадратичный невычет по модулю p, то есть $\left(\frac{b^2-4a}{p}\right)=-1$.
2. Пусть $f(x) = x^2-bx+a \in \mathbb{Z}_p[x]$.
3. $r := x^{(p+1)/2} \pmod f$. (r будет целым числом).
4. Вернуть $r, -r$.

8.4.5. Вычисление квадратного корня по модулю n, если p и q есть простые факторы в n

ВХОД. Целое n, его простые факторы p и q, a есть квадратичный вычет по модулю n.

ВЫХОД. Четыре квадратных корня из a по модулю n.

1. Найти два квадратных корня r и $-r$ из a по модулю p.
2. Найти два квадратных корня s и $-s$ из a по модулю q.
3. С помощью расширенного алгоритма Евклида найти целые c и d, для которых $cp + dq = 1$.
4. $x := (rdq+scp) \pmod n$ и $y := (rdq-scp) \pmod n$.
5. Вернуть $\pm x \pmod n$, $\pm y \pmod n$.

8.5. Проблема дискретного логарифма

Криптографическая стойкость криптосистемы Диффи–Хеллмана, криптосистемы ЭльГамаля, схемы цифровой подписи ЭльГамаля и их модификаций основывается на трулности решения проблемы дискретного логарифма.

Определение. Пусть G есть конечная циклическая группа порядка n. Пусть α есть генератор для G и пусть $\beta \in G$. *Дискретный логарифм β по основанию α* (обозначение: $\log_\alpha \beta$) есть единственное целое x из $[0,n-1]$, для которого $\beta = \alpha^x$.

Пример. Пусть p=97. Тогда $\mathbb{Z}^*_{97}$ есть циклическая группа порядка n=96. Генератор для $\mathbb{Z}^*_{97}$ есть α=5. Так как $5^{32} \equiv 35 \pmod{97}$, то $\log_5 35 \pmod{97} = 32$ в $\mathbb{Z}^*_{97}$.

Определение. *Проблема дискретного логарифма* есть задача нахождения целого $x \in [0,p-2]$, для которого $\alpha^x \equiv \beta \pmod p$, где p есть простое число, α есть генератор для $\mathbb{Z}^*_p$, элемент $\beta \in \mathbb{Z}^*_p$.

Определение. *Обобщенная проблема дискретного логарифма* есть нахождение целого $x \in [0,n-1]$, для которого $\alpha^x \equiv \beta \pmod p$, где G есть конечная циклическая группа G порядка n, α есть генератор для G, элемент $\beta \in G$.

Замечание. Наиболее очевидный алгоритм решения обобщенной проблемы дискретного логарифма есть последовательное вычисление $\alpha^0,\alpha^1,\alpha^2,\ldots$, пока не получим β. При больших n этот метод не эффективен и практически не осуществим.

Предположение. Не существует полиномиального алгоритма решения проблемы дискретного алгоритма.

8.5.1. Алгоритм "малый шаг – большой шаг" вычисления дискретного логарифма

ВХОД. Генератор α мультипликативной циклической группы G порядка n и элемент $\beta \in G$.

ВЫХОД. Дискретный логарифм $x = \log_\alpha \beta$.

1. $m := \lceil \sqrt{n} \rceil$.
2. Построить таблицу $(j, \alpha^j \pmod n)$ для $0 \leqslant j < m$. Сортировать эту таблицу по второй компоненте.
3. Вычислить α^{-m} и $\gamma := \beta$.
4. Для i от 0 до $m-1$ выполнить следующее.
   4.1. Проверить, является ли γ второй компонентой некоторой пары в сортированной таблице.
   4.2. Если $\gamma = \alpha^j$, то вернуть $x = im+j$.
   4.3. $\gamma := \gamma \cdot \alpha^{-m}$.

Пример. (Алгоритм "малый шаг – большой шаг" вычисления дискретного логарифма). Пусть p=113. Элемент α=3 есть генератор для $\mathbb{Z}^*_{113}$ порядка n=112. Пусть β=57. Найти $\log_\alpha \beta \pmod p$. *Решение.*

1. $m := \lceil \sqrt{n} \rceil = \lceil \sqrt{112} \rceil = 11$.
2. Строим таблицу пар $(j, \alpha^j \pmod p)$ для $0 \leqslant j < m$=11.

j	0	1	2	3	4	5	6	7	8	9	10
$\alpha^j \pmod p = 3^j \pmod{113}$	1	3	9	27	81	17	51	40	7	21	63

Сортируем таблицу по второй компоненте.

j	0	1	8	2	5	9	3	7	6	10	4
$\alpha^j \pmod p = 3^j \pmod{113}$	1	3	7	9	17	21	27	40	51	63	81

3. $\alpha^{-1} \pmod p = 3^{-1} \pmod{113} = 38$,
$\alpha^{-m} \pmod p = (\alpha^{-1})^m \pmod p = 38^{11} \pmod{113} = 58$.

4. $\gamma := \beta\alpha^{-mi} \pmod{p=113} = \beta(\alpha^{-m})^i \pmod{p=113}$, $i = 0,1,2,\ldots$ Вычислять пока не получится значение из второй строки сортированной таблицы. Результаты вычислений приведены в следующей таблице.

i	0	1	2	3	4	5	6	7	8	9
$\gamma := \beta\alpha^{-mi} \pmod p$ $=57\cdot 58^i \pmod{113}$	57	29	100	37	112	55	26	39	2	3

Так как $\gamma = \beta\alpha^{-9m} = \beta\alpha^{-9m} = 3 = \alpha^1$, то $\beta = \alpha^{100}$, откуда $\log_\alpha\beta = \log_3 57 \pmod{113} = 100$.

8.5.2. Rho алгоритм Полларда вычисления дискретного логарифма

Пусть G есть мультипликативная циклическая подгруппа простого порядка n циклической группы $\mathbb{Z}_p^*$ при простом p. В соответсвии с каким-либо легко проверяемым свойством поделим группу G на три подмножества S_1, S_2, S_3 примерно одного размера, взяв, например, $1 \in S_2$. Определим последовательность элементов группы $x_0, x_1, x_2, \ldots$ и две последовательности чисел: $a_0, a_1, a_2, \ldots$ и $b_0, b_1, b_2, \ldots$, полагая

$$x_0=1, \quad x_{i+1} = f(x_i) = \begin{cases} \beta\cdot x_i \pmod p, & \text{если } x_i\in S_1, \\ x_i^2 \pmod p, & \text{если } x_i\in S_2, \\ \alpha\cdot x_i \pmod p, & \text{если } x_i\in S_3, \end{cases} \tag{8.1}$$

$i = 0,1,2,\ldots$

$$a_{i+1} = \begin{cases} a_i, & \text{если } x_i\in S_1, \\ 2a_i \pmod n, & \text{если } x_i\in S_2, \\ a_i+1 \pmod n, & \text{если } x_i\in S_3, \end{cases} \tag{8.2}$$

$$b_{i+1} = \begin{cases} b_i+1 \pmod n, & \text{если } x_i\in S_1, \\ 2b_i \pmod n, & \text{если } x_i\in S_2, \\ b_i \pmod n, & \text{если } x_i\in S_3, \end{cases} \tag{8.3}$$

Заметим, что $b_i \equiv b_{2i} \pmod n$ с очень малой вероятностью.

ВХОД. Генератор α мультипликативной циклической группы G простого порядка n, и элемент $\beta \in G$.
ВЫХОД. Дискретный логарифм $x = \log_\alpha\beta$.
1. $x_0 := 1$, $a_0 := 0$, $b_0 := 0$.
2. Для $i = 1,2,\ldots$ выполнить следующее.
   2.1. Используя ранее найденные x_{i-1}, a_{i-1}, b_{i-1} и $x_{2i-2}, a_{2i-2}b_{2i-2}$, с помощью равенств (8.1), (8.2), (8.3) вычислить x_i, a_i, b_i и x_{2i}, a_{2i}, b_{2i}.
   2.2. Если $x_i = x_{2i}$, то выполнить следующее. $r := b_i - b_{2i} \pmod n$. Если $r=0$, то алгоритм заканчивает работу безрезультатно. Иначе $x := r^{-1}(a_{2i}-a_i) \pmod n$ и вернуть x.

Замечание. Если алгоритм заканчивает работу безрезультатно, то процедуру следует повторить, выбирая случайные числа a_0, b_0 из $[1,n-1]$ и стартуя с $x_0 = \alpha^{a_0}\beta^{b_0}$.

Пример. (rho алгоритм Полларда вычисления дискретного логарифма в подгруппе группы $\mathbb{Z}_{383}^*$). Пусть элемент $\alpha=2$ есть генератор подгруппы G в группе $\mathbb{Z}_{383}^*$ порядка $n=191$ и пусть элемент $\beta=228$. Поделим элементы в G на три подмножества (взяв $1 \in S_1$) в соответствии с правилом:

$$x \in \begin{cases} S_1, & \text{если } x \equiv 1 \pmod 3, \\ S_2, & \text{если } x \equiv 0 \pmod 3, \\ S_3, & \text{если } x \equiv 2 \pmod 3. \end{cases}$$

В табл.8.1 приведены значения x_i, a_i, b_i, x_{2i}, a_{2i}, b_{2i} в конце каждой итерации пункта 2 алгоритма. Заметим, что $x_{14} = x_{28} = 144$. Наконец, вычисляем $r = b_{14}-b_{28} \pmod{191} = 125$, $r^{-1} = 125^{-1} \pmod{191} = 136$, $r^{-1}(a_{28}-a_{14}) \pmod{191} = 110$ и $\log_2 228 = 110$.

Таблица 8.1

i	x_i	a_i	b_i	x_{2i}	a_{2i}	b_{2i}
1	228	0	1	279	0	2
2	279	0	2	184	1	4
3	92	0	4	14	1	6
4	184	1	4	256	2	7
5	205	1	5	304	3	8

6	14	1	6	121	6	18
7	28	2	6	144	12	38
8	256	2	7	235	48	152
9	152	2	8	72	48	154
10	304	3	8	14	96	118
11	372	3	9	256	97	119
12	121	6	18	304	98	120
13	12	6	19	121	5	51
14	144	12	38	144	10	104

8.5.3. Алгоритм Полига-Хеллмана вычисления дискретного логарифма

ВХОД. Генератор α мультипликативной циклической группы G порядка n и элемент $\beta \in G$.

ВЫХОД. Дискретный логарифм $x = \log_\alpha \beta$.

1. Найти факторизацию $n = p_1^{e_1} p_2^{e_2} \ldots p_r^{e_r}$, где все $e_i \geq 1$.
2. Для i от 1 до r выполнить следующее. (Вычисление

   $x_i = l_0 + l_1 p_i + l_2 p_i^2 + \ldots + l_{e_i-1} p_i^{e_i-1}$,

   где $x_i = x \pmod{p_i^{e_i}}$).

   2.1. $q := p_i$, $e := e_i$.
   2.2. $\gamma := 1$, $l_{-1} := 0$.
   2.3. $\alpha 1 := \alpha^{n/q}$.
   2.4. Для j от 0 до $e-1$ выполнить следующее.
   (Вычисление l_j).

   $\gamma := \gamma \alpha^{l_{j-1} q^{j-1}}$, $\beta 1 := (\beta \cdot \gamma^{-1})^{n/q^{j+1}}$.
   $l_j := \log_{\alpha 1} \beta 1$.
   2.5. $x_i := l_0 + l_1 q + l_2 q^2 + \ldots + l_{e-1} q^{e-1}$.
3. Используя алгоритм Гаусса, решить систему сравнений

   $x \equiv x_i \pmod{p_i^{e_i}}$, $1 \leq i \leq r$,
   и взять число $x \in [0, n-1]$.
4. Вернуть x.

Пример. (Алгоритм Полига-Хеллмана вычисления дискретного логарифма в $\mathbb{Z}^*_{251}$). p=251; α=71 есть генератор для группы $\mathbb{Z}^*_{251}$ порядка n=250. Пусть β=210. Тогда дискретный логарифм $x = \log_\alpha \beta = \log_{71} 210 \pmod{251}$ вычисляется следующим образом.

1. Факторизация $n = 250 = 2 \cdot 5^3$.
2. (a) Вычисление $x_1 = x \pmod 2$.
   $\alpha 1 := \alpha^{n/2} \pmod p = 250$.

   $\beta 1 := \beta^{n/2} \pmod p = 250$.
   $x_1 := \log_{250} 250 = 1$.
   (b) Вычисление $x_2 = x \pmod{5^3} = l_0 + l_1 5 + l_2 5^2$).
   1) $\alpha 1 := \alpha^{n/5} \pmod p = 20$.
   2) $\gamma := 1$, $\beta 1 = (\beta \cdot \gamma^{-1})^{n/5} \pmod p = 149$.
   Используя исчерпывающий поиск (он предпочтительнее алгоритма "малый шаг – большой шаг", если группа имеет малый порядок (здесь порядок $\alpha 1$ равен 5), находим
   $l_0 := \log_{20} 149 = 2$.
   3) $\gamma := \gamma \alpha^2 \pmod p = 21$,
   $\beta 1 := (\beta \cdot \gamma^{-1})^{n/25} \pmod p = 113$.
   Используя исчерпывающий поиск, находим
   $l_1 := \log_{20} 113 = 4$.
   4) $\gamma := \gamma \alpha^{4 \cdot 5} \pmod p = 115$,
   $\beta 1 = (\beta \cdot \gamma^{-1})^{(p-1)/125} \pmod p = 149$.
   Используя исчерпывающий поиск, находим
   $l_2 := \log_{20} 149 = 2$.
   $x_2 := 2 + 4 \cdot 5 + 2 \cdot 5^2 = 72$.
3. Решаем систему сравнений $x \equiv 1 \pmod 2$, $x \equiv 72 \pmod{125}$ и получаем $x = \log_{71} 210 \pmod{251} = 197$.

8.6. Проблема подмножества суммы

Дано множество положительных целых чисел $M = \{a_1, a_2, \ldots, a_n\}$ и положительное целое s. Определить, существует ли подмножество элементов из M, сумма которых есть s.

Проблема подмножества суммы: по множеству положительных целых чисел $M = \{a_1, a_2, \ldots, a_n\}$ и положительному целому s определить, существует ли подмножество элементов из M, сумма которых есть s. Другими словами, определить, существуют ли $x_i \in \{0,1\}$, $1 \leq i \leq n$, для которых $\sum_{i=1}^{n} a_i x_i = s$.

Предположение. Не существует полиномиального алгоритма решения проблемы подмножества суммы.

8.6.1. Наивный (переборный) алгоритм решения проблемы суммы

ВХОД. Множество положительных целых чисел $M = \{a_1, a_2, \ldots, a_n\}$ и положительное целое s.

ВЫХОД. Найти такие $x_i \in \{0,1\}$, $1 \leq i \leq n$, если они

существуют, для которых $\sum_{i=1}^{n} a_i x_i = s$.

1. Для каждого вектора $(x_1,\ldots x_n)$ из $\mathbb{Z}_2^n$ выполнить следующее.

   1.1. $l := \sum_{i=1}^{n} a_i x_i$.

   1.2. Если $l=s$, то вернуть $(x_1,\ldots x_n)$.

2. Вернуть "Решения нет".

Замечание. Алгоритм требует $O(2^n)$ шагов работы и потому не эффективен.

8.6.2. Алгоритм "встреча посередине" решения проблемы подмножество суммы

ВХОД. Множество положительных целых чисел $M = \{a_1,a_2,\ldots,a_n\}$ и положительное целое s.

ВЫХОД. Найти такие $x_i \in \{0,1\}$, $1 \leqslant i \leqslant n$, если они существуют, для которых $\sum_{i=1}^{n} a_i x_i = s$.

1. $t := \lfloor n/2 \rfloor$.

2. Построить таблицу векторов $(\sum_{i=1}^{t} a_i x_i,\ (x_1,\ldots x_t))$ для $(x_1,\ldots,x_t) \in \mathbb{Z}_2^t$. Сортировать эту таблицу по первой компоненте.

3. Для каждого $(x_{t+1},x_{t+2},\ldots,x_n) \in \mathbb{Z}_2^{n-t}$ выполнить следующее.

   3.1. $l := s - \sum_{i=t+1}^{n} a_i x_i$ и проверить, является ли l первой компонентой некоторого вектора в таблице.

   3.2. Если $l = \sum_{i=1}^{t} a_i x_i$, то вернуть $(x_1,\ldots x_t)$.

4. Вернуть "Решения нет".

Замечание. Алгоритм требует $O(n2^{n/2})$ шагов работы и потому не эффективен.

8.7. Факторизация полиномов над конечным полем

Пусть полином $f(x) \in \mathbb{F}_q[x]$, $q = p^m$. Ставится задача нахождения факторизации $f(x) = f_1(x)^{e_1} f_2(x)^{e_2} \ldots f_t(x)^{e_t}$, где каждое $f_i(x)$ есть неприводимый полином в $\mathbb{F}_q[x]$ и каждое $e_i \geqslant 1$. Число e_i называется *кратностью* фактора $f_i(x)$.

Предположение. Не существует полиномиального алгоритма факторизации полиномов над конечными полями.

8.7.1. Бесквадратная факторизация

Пусть $f(x)$ есть нормированный полином.

Определение. Полином $f(x)$ из $\mathbb{F}_q[x]$ называется *бесквадратным*, если он не имеет кратных факторов, то есть не существует полинома $g(x)$ степени $\deg(g(x))$ такого, что $g(x)^2$ делит $f(x)$. Бесквадратная факторизация $f(x)$ есть представление

$$f(x) = \prod_{i=1}^{k} f_i(x)^i,$$

где каждое $f_i(x)$ есть бесквадратный полином и нод$(f_i(x), f_j(x)) = 1$ для $i \neq j$. (Некоторые из $f_i(x)$ в бесквадратной факторизации $f_i(x)$ могут быть 1).

Пусть $f(x) = \sum_{i=0}^{n} a_i x^i$ есть полином степени $n \geqslant 1$. (Формальная) производная для $f(x)$ есть полином $f'(x) = \sum_{i=0}^{n-1} a_{i+1}(i+1)x^i$. Если $f'(x)=0$ и так как p есть характеристика $\mathbb{F}_q$, то в каждом члене $a_i x^i$ в $f(x)$, для которого $a_i \neq 0$, показатель степени для x кратен p. Поэтому $f(x)$ имеет вид $f(x) = a(x)^p$, где $a(x) = \sum_{i=0}^{n/p} a_{ip}^{q/p} x^i$. Проблема нахождения бесквадратной факторизации для $f(x)$ сводится к нахождению таковой для $a(x)$. Если $a'(x) = 0$, то повторяем этот процесс, пока $a'(x) \neq 0$. Поэтому можно сразу предположить, что $f'(x) \neq 0$.

Находим $g(x) =$ нод$(f(x),f'(x))$. Всякий неприводимый фактор кратности k в $f(x)$ имеет кратность $k-1$ в $f'(x)$, если нод$(k,p) = 1$, а иначе имеет кратность k в $f'(x)$. Если $g(x) = 1$, то $f(x)$ не имеет кратных факторов. Если $g(x)$ имеет положительную степень, то $g(x)$ есть нетривиальный фактор в $f(x)$ и $f(x)/g(x)$ не имеет кратных факторов. Возможно, что $g(x)$ имеет кратные факторы, возможно, что $g'(x) = 0$. Тогда применяем к $g(x)$ выше описанный процесс. Шаги этих вычислений можно свести в следующий алгоритм.

8.7.2. Алгоритм бесквадратной факторизации

SQUARE-FREE$(f(x))$

ВХОД. Нормированный полином $f(x) \in \mathbb{F}_q[x]$, $q = p^m$ степени $\geqslant 1$, где $\mathbb{F}_q$ имеет характеристику p.

ВЫХОД. Бесквадратная факторизация для $f(x)$.

1. $i := 1$, $F := 1$, вычислить $f'(x)$.

2. Если $f'(x) = 0$, то $f(x) := f(x)^{1/p}$ и $F := (\text{SQUARE-FREE}(f(x)))^p$. Иначе (то есть $f'(x) \neq 0$) выполнить следующее.

2.1. $g(x) := \text{нод}(f(x), f'(x))$, $h(x) := f(x)/g(x)$.
2.2. Пока $h(x) \neq 1$, выполнять следующее.
$h1(x) := \text{нод}(h(x), g(x))$, $l(x) := h(x)/h1(x)$.
$F := F \cdot l(x)^i$, $i := i+1$, $h(x) := h1(x)$,
$g(x) := g(x)/h1(x)$.
2.3. Если $g(x) \neq 1$, то $g(x) := g(x)^{1/p}$,
$F := F \cdot (\text{SQUARE-FREE}(g(x)))^p$.
3. Вернуть F.

Замечание. Если бесквантворная факторизация

$$f(x) = \prod_{i=1}^{k} f_i(x)^i$$

найдена, то следует факторизовать бесквадратные полиномы $f_1(x)$, $f_2(x), \ldots, f_n(x)$ и получить полную (каноническую) факторизацию $f(x)$.

8.7.3. Q-матричный алгоритм Берленкампа

Пусть $f(x) = \prod_{i=1}^{t} f_i(x)$ есть нормированный полином в $\mathbb{F}_q[x]$ степени n, имеющий различные неприводимые факторы $f_i(x)$, $1 \leqslant i \leqslant t$. Множество полиномов

$$\mathcal{B} = \{b(x) \in \mathbb{F}_q[x] / (f(x) : b(x)^p \equiv b(x) \pmod{f(x)})\}$$

есть векторное пространство размерности t над $\mathbb{F}_q$. $\mathcal{B}$ состоит в точности из тех векторов в нуль-пространстве матрицы $Q-I_n$, где Q есть $n \times n$-матрица с элементами q_{ij}, определяемыми системой линейных уравнений

$$x^{iq} \pmod{f(x)} = \sum_{j=0}^{n-1} q_{ij} x^j, \quad 0 \leqslant i \leqslant n-1,$$

и где I_n есть $n \times n$ единичная матрица. Базис $B = \{v_1(x), v_2(x), \ldots, v_t(x)\}$ для B можно найти методами линейной алгебры. Для каждой пары различных факторов $f_i(x)$ и $f_j(x)$ в $f(x)$ существует $v_k(x) \in B$ и $\alpha \in \mathbb{F}_q$, для которых $f_i(x)$ делит $v_k(x)-\alpha$, но $f_j(x)$ не делит $v_k(x)-\alpha$. Эти два фактора могут быть разложены вычислением $\text{нод}(f(x), v_k(x)-\alpha)$. В алгоритме Берленкампа вектор $w = (w_0, w_1, \ldots, w_{n-1})$ определяется полиномом $w(x) = \sum_{i=0}^{n-1} w_i x^i$.

8.7.4. Q-матричный алгоритм Берленкампа факторизации полиномов над конечным полем

ВХОД. Бесквадратный нормированный полином $f(x)$ степени n в $\mathbb{F}_q[x]$.

ВЫХОД. Факторизация $f(x)$ по нормированным неприводимым полиномам.
1. Для каждого i, $0 \leqslant i \leqslant n-1$, найти полином
$$x^{iq} \pmod{f(x)} = \sum_{j=0}^{n-1} q_{ij} x^j.$$
Каждое $q_{ij} \in \mathbb{F}_q$.
2. Составить $n \times n$ матрицу Q с элементами q_{ij}.
3. Найти базис $v_1, v_2, \ldots, v_t$ для нуль-пространства матрицы $Q-I_n$, где I_n есть $n \times n$ единичная матрица. Число неприводимых факторов в $f(x)$ есть в точности t.
4. $F := \{f(x)\}$. F есть множество факторов в $f(x)$, найденных до сих пор; их число равно $f(x)$.
5. Для i от 1 to t выполнить следующее.
5.1. Для каждого полинома $h(x) \in F$, для которого deg $h(x) > 1$, выполнить следующее. Вычислить $\text{нод}(h(x), v_i(x)-\alpha)$ для каждого $\alpha \in \mathbb{F}_q$. Заменить $h(x)$ в F всеми теми полученными при вычислении нод полиномами, у которых степени $\geqslant 1$.
6. Вернуть множество F (неприводимых) факторов для $f(x)$.

Замечание. Метод эффективен при малых q. Найдены гораздо более эффективные вероятностные алгоритмы факторизации полиномов над конечными полями даже при довольно больших q.

8.8. Криптосистема RSA

Криптосистема RSA (ее предложили R.Rivest, A.Shamir, L.Adleman) есть одна из широко используемых криптосистем с открытым ключом. Ее криптографическая стойкость основана на трудной практической осуществимости проблемы факторизации больших целых чисел.

Пусть адресат A посылает информацию адресату B. Адресат A представляет исходный текст t в виде натурального числа m, шифрует сообщение m, получает шифротекст c, отправляет c адресату B. Адресат B получает шифротекст c, дешифрует c, получает число m и исходный текст t.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен выполнить следующее.

1. Выбрать два различных случайных простых числа p и q примерно одного размера.
2. Найти $n = pq$ и функцию Эйлера $\varphi = \varphi(n) = (p-1)(q-1)$.

3. Взять случайное число e, $1<e<\varphi$, такое, что нод(e,φ)=1.

4. Найти такое целое $a \in (1,\varphi)$, что $ea \equiv 1 \pmod{\varphi}$. Для этого с помощью расширенного алгоритма Евклида найти такие целые a,x, что $ea + \varphi x = 1$. Тогда $ea \equiv 1 \pmod{\varphi}$. Пусть произвольное $k \in \mathbb{Z}$. Сложив $ea \equiv 1 \pmod{\varphi}$ и $ek\varphi \equiv 0 \pmod{\varphi}$, получим $e(a+k\varphi) \equiv 1 \pmod{\varphi}$. Если $a \notin (1,\varphi)$, то найти такое целое k, что $a+k\varphi \in (1,\varphi)$, и в качестве a взять $a+k\varphi$.

5. Открытый ключ адресата есть пара чисел (n,e). Секретный ключ адресата есть число a.

Шифрование. Адресат A шифрует свой текст t и отправляет его адресату B. B дешифрует сообщение от A и получает исходный текст t. Адресат A должен выполнить следующее.

1. Получить открытый ключ (n,e) адресата B.

2. С помощью какого-либо метода M, который публикуется, представить свое письмо t как сообщение в виде натурального числа m из сегмента $[0,n-1]$.

3. Вычислить шифротекст $c = m^e \pmod{n}$.

4. Отправить свой шифротекст c адресату B.

Дешифрование. Чтобы извлечь текст t из шифротекста c, адресат B должен выполнить следующее.

1. Взять свой секретный ключ a и вычислить сообщение $m = c^a \pmod{n}$.

2. Вычислить текст t адресата A с помощью метода M.

Доказательство. Так как $ed \equiv 1 \pmod{\varphi}$, то существует целое k, для которого $ed = 1 + k\varphi$. Далее, если нод$(m,p) = 1$, то по теореме Ферма $m^{p-1} \equiv 1 \pmod{p}$. Возвывая обе части этого сравнения в степень $k(q-1)$ и умножая обе части на m, получим $m^{1+k(p-1)(q-1)} \equiv 1 \pmod{p}$. С другой стороны, если нод$(m,p) = p$, то это сравнение верно, такк как каждая его часть сравнима с 0 по модулю p. Следовательно, во всех случаях $m^{ed} \equiv m \pmod{p}$. Из тех же соображений $m^{ed} \equiv m \pmod{q}$. Наконец, так как p и q различные простые числа, то $m^{ed} \equiv m \pmod{n}$ и потому $c^d \equiv (m^e)^d \equiv m \pmod{n}$.

Пример. Адресат A пишет письмо t=NAB адресату B.

Вычисление ключей. Адресат B выполняет следующее.

1. Выбирает два разных простых числа p=499, q=631.

2. Вычисляет $n=pq$=314869 и функцию Эйлера $\varphi=(p-1)(q-1)$ = 313740.

3. Выбирает случайное число e=305183$\in(1,\varphi)$ с нод(e,φ)=1.

4. С помощью расширенного алгоритма Евклида находит такое

a = 181967 $\in (1,\varphi)$, что $ea \equiv 1 \pmod{\varphi}$.

5. Открытый ключ адресата B есть пара чисел (n=314869, e=305183). Секретный ключ адресата B есть число a = 181967.

Шифрование. Адресат A выполняет следующее.

1. Получает открытый ключ (n=314869,e=305183) адресата B.

2. Представляет свой текст t=NAB в виде натурального числа m из $[0,n-1]$ с помощью какого-либо метода, например, с помощью 27-ричной системы счисления следующим образом. Нумеруются буквы алфавита:

пробел	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	P	Q	R	S	T	U	V	W	X	Y	Z				
	16	17	18	19	20	21	22	23	24	25	26				

Текст NAB представляется в виде числа $m = 14 \cdot 27^2 + 1 \cdot 27 + 2 = 10235$.

3. Шифрует свое сообщение m=10235 числом $c = m^e \pmod{n} = 10235^{305183} \pmod{314869} = 301085$.

4. Посылает свой шифротекст c адресату B.

Дешифрование. Чтобы дешифровать шифротекст c от A, адресат B выполняет следующее.

1. Находит (с помощью своего секретного ключа a) число
$m = c^a \pmod{n} = 301085^{181967} \pmod{314869} = 10235$.

2. Представляет число m в 27-ричной системе счисления:
$m = (14\,1\,2)_{27}$ и получает исходный текст NAB.

Замечание. 1. Текст t в компьютере представляется бинарным массивом, который рассматривается как бинарная запись некоторого числа m. Предложенный выше способ представления текста числом носит иллюстративный характер и выбран из желания оперировать небольшими числами.

2. На практике для криптографической стойкости модуль n задается двоичным числом с 1024 и более двоичными разрядами.

8.9. Электронная цифровая подпись RSA с извлечением сообщения

Криптографическая стойкость цифровой подписи RSA основывается на трудности проблемы факторизации целых чисел. Так как шифрование есть взаимно однозначное преобразование, то цифровая подпись может быть получена обращением шифрования и

дешифрования.

Для схемы подписи RSA требуется взаимнооднозначная функция $R: \mathbb{N} \to \mathbb{N}$, для которой множество $R(\mathbb{N})$ значений R имеет характеристическое свойство, которое операция шифрования не сохраняет. В этом параграфе для этой цели возьмем, например, функцию $R(m) = m||m$, где $m||n$ есть конкатенация (соединение) 10-ричных записей m и n. Например, 12734||7590 = 127347590, 2354||2354 = 23542354.

Адресат A подписывает свое сообщение m. Всякий адресат B может проверить подпись A и извлечь из нее сообщение m.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать два различных случайных простых числа p и q приблизительно одного размера.
2. Найти числа $n = pq$ и $\varphi = (p-1)(q-1)$.
3. Найти такое целое число e, $1<e<\varphi$, что нод$(e,\varphi)=1$.
4. С помощью расширенного алгоритма Евклида найти такое целое a, $1<a<\varphi$, для которого $ea \equiv 1 \pmod{\varphi}$.
5. Открытый ключ адресата есть пара (n,e). Секретный ключ адресата есть a.

Вычисление подписи. Адресат A подписывает свой текст t. Любой адресат B может проверить подпись A и извлечь из нее текст t. Адресат A должен выполнить следующее.

1. Каким-либо методом M (который публикуется) представить свой текст t в виде целого числа m, $1 < m < n-1$.
2. Найти число $w=R(m)$ с помощью открытой функции $R: [0, n-1] \to M_R$, где M_R есть некоторое числовое множество, например, $R(m) = m||m$, где $a||b$ есть результат приписывания слова b к слову a. Тогда $M_R = \{w=m||m: m \in [0,n-1]\}$.
3. Найти число $s = w^a \pmod{n}$.
4. Отправить подписанный шифротекст s адресату B.

Проверка подписи и вычисление сообщения. Чтобы проверить подпись s адресата A и извлечь из нее сообщение m, адресат B должен выполнить следующее.

1. Получить открытый ключ (n,e) адресата A.
2. Найти число $w = s^e \pmod{n}$.
3. Проверить, что $w \in M_R$. Если нет, отвергнуть подпись s.
4. Найти число $m = R^{-1}(w)$.

5. С помощью метода M найти отправленный текст t.

Доказательство. Если s есть подпись для сообщения m, то $s \equiv (w)^d \pmod{n}$, где $w = R(m)$. Так как $ed \equiv 1 \pmod{\varphi}$, то $s^e \equiv (w)^{ed} \equiv w \pmod{n}$. Наконец, $R^{-1}(w) = R^{-1}(R(m)) = m$.

Пример. Адресат A подписывает свой текст t. Любой адресат B может проверить подпись A.

Вычисление ключей. Адресат A выполняет следующее.

1. Выбирает разные простые числа p=1019, q=2347.
2. Находит $n=pq$ = 2391593 и функцию Эйлера $\varphi = (p-1)(q-1) = 1018 \cdot 2346 = 2388228$.
3. Выбирает случайное число e=35, $1<e<\varphi$, с нод$(e,\varphi)=1$.
4. С помощью расширенного алгоритма Евклида находит то единственное целое a=1569407$\in(1,\varphi)$, которое удовлетворяет сравнению $ea \equiv 1 \pmod{\varphi}$, это сравнение $35d \equiv 1 \pmod{2388228}$.
5. Открытый ключ для A есть пара (n=2391593, e=35). Секретный ключ для A есть число a=1569407.

Вычисление подписи. Адресат A подписывает свой текст t = ABX и выполняет следующее.

1. Представляет свой текст t=ABX числом каким-либо методом M, например, в 27-ричной системе счисления числом $m = 1 \cdot 27^2 + 2 \cdot 27 + 24 = 807$.
2. Вычисляет $w = R(m) = R(807) = 807||807 = 807807$.
3. Вычисляет подпись $s = w^a \pmod{n} = 807807^{1569407} \pmod{2391593} = 794011$.
4. Отправляет подписанный шифротекст s адресату B.

Проверка подписи и извлечение сообщения. Адресат B получает от A подписанный шифротекст s и делает следующее.

1. Получает открытый ключ (n=2391593, e=35) адресата A.
2. С помощью открытого ключа (n,e) адресата A вычисляет: $w = s^e \pmod{n} = 794011^{35} \pmod{2391593} = 807807$.
3. Так как $w = 807807 = 807||807$ и $w \in R(m)$, то B принимает подпись A.
4. Вычисляет $m = R^{-1}(w) = 807$.
5. Представляет число $m=(807)_{10}$ в 27-ричной системе счисления $m = (1\ 2\ 24)_{27}$ и получает исходный текст t = ABX.

Замечание. 1. Есть другие более сложные функции $R(m)$ с

меньшей длиной записи значения $R(m)$, чем для здесь предложенной функции $R(m)$.

2. На практике для криптографической стойкости цифровой подписи RSA модуль n задается двоичным числом с 1024 и более двоичными разрядами.

8.9.1. Электронная цифровая подпись RSA с использованием хэш-функции

Допустима цифровая подпись RSA, основанная на использовании криптографической хэш–функции $w=h(x)$, сопоставляющей любому тексту x уникальное целое число $w=h(x)$.

Вычисление ключей. Пусть по-прежнему: p=1019, q=2347; пара (n=2391593, e=35) есть открытый ключ для A и число a=1569407 есть сек-ретный ключ для A.

Вычисление подписи. Адресат A подписывает свой текст t произвольной длины. Любой адресат B может проверить подпись A под его текстом t. Адресат A должен выполнить следующее.

1. Вычислить значение хэш-функции $h = h(t)$. Пусть для примера текст t=DXN, $m=4\cdot 27^2+24\cdot 27+14=3578$, $h=h(m)=m=3578$.

2. Вычислить $s = h^a \pmod n = 3578^{1569407} \pmod{2391593} = 2146200$. Число s есть подпись A под его текстом t.

Проверка подписи. Чтобы проверить подпись s адресата A, адресат B должен выполнить следующее.

1. Получить открытый ключ (n,e) адресата A.

2. Вычислить значение хэш-функции $h = h(t)$. Если текст t не изменялся, то h=3578.

3. Вычислить $h1 = s^e \pmod n = 2146200^{35} \pmod{2391593} = 3578$.

4. Принять подпись, если $h = h1$, и отвергнуть в противном случае. Так как $h = h1 = 3578$, то подпись принимается.

8.10. Криптосистема ЭльГамаля

Криптографическая стойкость криптосистемы Эль-Гамаля обеспечивается трудной практической осуществимостью проблемы нахождения дискретного логарифма в мультипликативной группе $\mathbb{Z}_p^*$ при больших простых числах p.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать случайное простое число p и найти генератор α мультипликативной группы $\mathbb{Z}_p^*$ целых чисел по модулю p, используя в главе 7 алгоритмы задач 10 или 21 (алгоритм Гаусса).

2. Выбрать случайное число $a\in[1,p-2]$ и найти $y=\alpha^a \pmod p$.

3. Открытый ключ адресата есть тройка чисел (p,α,y). Секретный ключ адресата есть число a.

Шифрование. Адресат A шифрует свой текст t и отправляет его адресату B. B дешифрует сообщение от A и получает исходный текст t. Адресат A должен выполнить следующее.

1. Получить открытый ключ (p,α,y) адресата B.

2. С помощью какого-либо метода M, который публикуется, представить свое письмо t как сообщение в виде натурального числа m из сегмента $[0,p-1]$.

3. Выбрать случайное число k, $1\leqslant k\leqslant p-2$.

4. Вычислить $\gamma = \alpha^k \pmod p$ и $\delta = m\cdot y^k \pmod p$.

5. Отправить свой шифротекст $c = (\gamma,\delta)$ адресату B.

Дешифрование. Чтобы получить исходный текст t по $c=(\gamma,\delta)$, адресат B должен выполнить следующее.

1. Взять свой секретный ключ a и вычислить целое число $\gamma^{p-1-a} \pmod p$.

2. Вычислить $m = (\gamma^{-a}\cdot\delta) \pmod p$, где $\gamma^{-a}=(\gamma^{-1})^a$, а число γ^{-1} есть решение сравнения $x\cdot\gamma\equiv 1 \pmod p$ и вычисляется с помощью расширенного алгоритма Евклида.

3. Вычислить исходный текст t от A с помощью метода M.

Доказательство. Дешифрованиее позволяет получить сообщение m от A, ибо $\gamma^{-a}\cdot\delta \equiv \alpha^{-ak}m\alpha^{ak} \equiv m \pmod p$.

Пример. Адресат A шифрует свой текст t=BUJ и отправляет его адресату B.

Вычисление ключей. Адресат B выполняет следующее.

1. Выбирает простое число p=2357 и находит генератор α=2 для мультипликативной группы $\mathbb{Z}_{2357}^*$.

2. Выбирает случайное число a=1751, $1\leqslant a\leqslant p-2$, и вычисляет $y = \alpha^a \pmod p = 2^{1751} \pmod{2357} = 1185$.

3. Открытый ключ адресата B есть тройка (p=2357, α=2, y = 1185). Секретный ключ адресата B есть число a=1751.

Шифрование. Адресат A шифрует свой текст t=BUJ и выполняет следующее.

1. Получает открытый ключ (p=2357, α=2, y=1185) для B.

2. Представляет свой текст t=BUJ в виде натурального числа m из $[0,p-1]$, с помощью какого-либо метода, например, с

помощью 27-ричной системы счисления числом $m = 2 \cdot 27^2 + 21 \cdot 27 + 10 = 2035$.

3. Выбирает случайное число k=1520, $1 \leqslant k \leqslant p-2$.

4. Вычисляет

$$\gamma = \alpha^k \pmod p = 2^{1520} \pmod{2357} = 1430,$$

$$\delta = m \cdot y^k \pmod p = 2035 \cdot 1185^{1520} \pmod{2357} = 697.$$

5. Посылает шифротекст $c = (\gamma$=1430, δ=697) адресату B.

Дешифрование. Чтобы дешифровать шифротекст $c = (\gamma$=1430, δ=697) от A, адресат B выполняет следующее.

1. Вычисляет

$\gamma^{p-1-a} = 1430^{605} \pmod{2357} = 872$ и получает

$m = ((\gamma^{p-1-a} \pmod p) \cdot \delta) \pmod p =$

$872 \cdot 697 \pmod{2357} = 2035.$

2. Представляет число m в 27-ричной системе счисления: $m = (2\,21\,10)_{27}$ и получает исходный текст BUJ.

Замечание. На практике для криптографической стойкости простое число p задается двоичным числом с 1024 и более двоичными разрядами.

8.11. Электронная цифровая подпись ЭльГамаля

При использовании схемы цифровой подписи ЭльГамаля по тексту письма t вычисляется значение хэш-функции $h(t)$, которое затем используется при вычислении и проверке цифровой подписи под текстом сообщения.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать случайное простое число p и найти генератор α для мультипликативной группы $\mathbb{Z}_p^*$.

2. Выбрать произвольное число a, $1 \leqslant a \leqslant p-2$.

3. Вычислить $y=\alpha^a \pmod p$.

4. Открытый ключ адресата есть тройка чисел (p,α,y). Секретный ключ адресата есть число a.

Вычисление подписи. Адресат A подписывает свой текст t (произвольной длины). Любой адресат B может проверить подпись адресата A под его текстом t. Адресат A должен выполнить следующее.

1. Вычислить значение хэш-функции $h(t)$.

2. Выбрать случайное секретное целое число k из $[1,p-2]$ такое, что нод$(k,p-1) = 1$.

3. Вычислить $k^{-1} \pmod{(p-1)}$.

4. Вычислить $r = \alpha^k \pmod p$.

6. Вычислить $s = k^{-1}(h(t)-ar) \pmod{(p-1)}$.

7. Подпись адресата A под его текстом t есть пара (r,s).

Проверка подписи. Чтобы проверить подпись (r,s) адресата A под его текстом t, адресат B должен выполнить следующее.

1. Вычислить значение хэш-функции $h(t)$.

2. Получить открытый ключ (p,α,y) адресата A.

3. Проверить, что $r \in [1,p-1]$; если нет, то отвергнуть подпись.

4. Вычислить $v_1 = y^r r^s \pmod p$.

5. Вычислить $v_2 = \alpha^{h(m)} \pmod p$.

6. Принять подпись, если $v_1=v_2$ и отвергнуть в противном случае.

Доказательство. Если адресат A подписал свое сообщение, то $s \equiv k^{-1} \cdot (h(m)-ar) \pmod{p-1}$. Умножим обе части сравнения на k и получим $ks \equiv h(m)-ar \pmod{p-1}$, откуда

$h(m) \equiv ar+ks \pmod{p-1}$,

$h(m) = ar+ks+(p-1)w$ при всяком целом w.

По теореме Ферма $\alpha^{p-1} = 1 \pmod p$. Поэтому

$v_2 = \alpha^{h(m)} = \alpha^{ar+ks+(p-1)w} = \alpha^{ar+ks} \cdot \alpha^{(p-1)w} =$

$\alpha^{ar+ks} \cdot (\alpha^{p-1})^w = \alpha^{ar+ks} \cdot 1 \pmod p =$

$(\alpha^a)^r \cdot (\alpha^k)^s \pmod p = y^r \cdot r^s \pmod p = v_1.$

Пример. Адресат A подписывает свой текст t. Любой адресат B может проверить подпись A.

Вычисление ключей. Адресат A выполняет следующее.

1. Выбирает простое число p=2357 и находит генератор α=2 мультипликативной группы $\mathbb{Z}_p^*$ целых чисел по модулю p.

2. Выбирает случайное целое a=1751 из $[1,p-2]$.

3. Вычисляет $y = \alpha^a \pmod p = 2^{1751} \pmod{2357} = 1185$.

4. Открытый ключ адресата A есть тройка (p=2357, α=2, y = 1185). Секретный ключ адресата A есть a=1751.

Вычисление подписи. Адресат A подписывает свой текст t и для этого выполняет следующее.

1. Вычисляет значение хэш-функции $h(t)$. Пусть для примера $h(t) = 1490$.

2. Выбирает случайное секретное число k=1529 из $[1,p-2]$

такое, что нод$(k,p-1) = 1$.

3. Вычисляет $k^{-1} \pmod{p-1} = 1529^{-1} \pmod{2356} = 245$.
4. Вычисляет $r = \alpha^k \pmod{p} = 2^{1529} \pmod{2357} = 1490$.
5. Вычисляет $s = k^{-1}(h(t)-ar) \pmod{(p-1)} =$
   $245 \cdot (1490-1751 \cdot 1490) \pmod{2356} = 1324$.
6. Подпись A есть пара (r=1490, s=1324).

Проверка подписи. Чтобы проверить подпись (r=1490,s=1324) адресата A под его текстом t, адресат B делает следующее.

1. Вычисляет значение хэш-функции $h(t)$. Если текст t не изменялся, то $h(t) = 1490$.
2. Получает открытый ключ (p=2357,α=2,y=1185) адресата A.
3. Проверяет, что r=1490 $\in [1,p-1] = [1,2356]$.
4. Вычисляет число
   $v_1 = y^r r^s \pmod{p} = 1185^{1490} \cdot 1490^{1324} \pmod{2357} = 1101$.
5. Вычисляет число
   $v_2 = \alpha^{h(t)} \pmod{p} = 2^{1490} \pmod{2357} = 1101$.
6. Принимает подпись, ибо $v_1=v_2$.

Замечание. Для криптографической стойкости рекомендуется брать простое число p длиной между 512 бит (лучше 768) и 1024 бит включительно.

8.12. Обобщенная криптосистема ЭльГамаля с мультипликативной группой G поля Галуа $GF(p^m)$

Числовая схема шифрования ЭльГамаля может быть обобщена для работы в любой конечной мультипликативной циклической группе G. Криптографическая стойкость схемы ЭльГамаля в группе G основана на трудности нахождения решения проблемы дискретного логарифма в G. Группа G должна удовлетворять следующим условиям.

1. *Эффективность*, то есть групповые операции в G должны вычисляться относительно просто.

2. *Криптографическая стойкость*, то есть решение проблемы дискретного логарифма в G должно быть практически неосуществимой.

Ниже следуют удовлетворяющие этим двум условиям группы, из которых первые три наиболее употребительны.

1. Мультипликативная группа $\mathbb{Z}_p^*$ целых чисел по модулю простого числа p.

2. Мультипликативная группа $\mathbb{Z}_{2^s}^*$ конечного поля $\mathbb{Z}_{2^s}$ характеристики два.

3. Группа точек эллиптической кривой над конечным полем.

4. Мультипликативная группа $\mathbb{Z}_q^*$ конечного поля $\mathbb{F}_q$, где $q = p^s$, p есть простое число, s есть положительное целое число.

5. Группа обратимых элементов $\mathbb{Z}_n^*$, где n есть составное целое число.

Адресат A шифрует свой текст t и отсылает его адресату B. B дешифрует сообщение от A и получает исходный текст t.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать подходящую (мультипликативную) циклическую группу G порядка n.
2. Найти генератор α группы G.
3. Выбрать случайное целое число a, $1 \leqslant a \leqslant n-1$.
4. Вычислить элемент $y = \alpha^a$ группы G.
5. Открытый ключ адресата есть пара (α,y) элементов группы G. Открыто также описание умножения элементов в G. Секретный ключ адресата есть число a.

Шифрование. Адресат A шифрует свой текст t и отправляет его адресату B. Адресат A должен выполнить следующее.

1. С помощью какого-либо метода M, который публикуется, представить свой текст t как элемент m группы G.
2. Получить открытый ключ (α,y) адресата B.
3. Выбрать случайное целое число k, $1 \leqslant k \leqslant n-1$.
4. Вычислить $\gamma = \alpha^k$ и $\delta = m \cdot y^k$.
5. Отправить свой шифротекст $c = (\gamma,\delta)$ адресату B.

Дешифрование. Чтобы получить исходный текст t по $c=(\gamma,\delta)$, адресат B должен выполнить следующее.

1. Взять свой секретный ключ a, вычислить γ^a и найти $\gamma^{-a} = (\gamma^a)^{-1}$.
2. Вычислить $m = (\gamma^{-a}) \cdot \delta$.
3. Вычислить исходный текст t от A с помощью метода M.

Замечание. Все адресаты могут выбрать одну и ту же циклическую группу G и ее генератор α.

Пример 1. Криптосистема ЭльГамаля с мультипликативной группой конечного поля $\mathbb{F}_{p^s}$, p=13, s=4. Пусть для удобства элемент поля $a_3x^3+a_2x^2+a_1x+a_0$ представляется p-ричным стрин-

гом $(a_3a_2a_1a_0)$.

Адресат A шифрует свой текст t=ZAM и отправляет его адресату B. B дешифрует сообщение от A и получает исходный текст t.

Вычисление ключей. Адресат B выполняет следующее.

1. Выбирает мультипликативную группу G конечного поля $(E_{13}^4,\{+,\cdot\})$, элементы которого представляются полиномами из $\mathbb{Z}_{13}[x]$ над $\mathbb{Z}_{13}$ степени меньше 4 и умножение в котором выполняется по модулю неприводимого полинома $f(x)$ = (1 0 1 0 2) = x^4+x^2+2 из $\mathbb{Z}_{13}[x]$. Группа G имеет порядок $n = p^s-1 = 13^4 - 1 = 28560$.
2. Находит генератор $\alpha = x+5 = (0\ 0\ 1\ 5)$.
3. Выбирает случайное число $a = 2 \in [1,n-1]$.
4. Вычисляет $y = \alpha^a = \alpha^2 = (x+5)^2 (\mathrm{mod}\ f(x)) = x^2+10x+12 =$ (0 1 10 12).
5. Открытый ключ для B есть пара (α=(0 0 1 5), y=(0 1 10 12) вместе с полиномом $f(x)$, который определяет умножение в G, если $f(x)$ и α не есть параметры, общие всем адресатам. Секретный ключ для B есть число a=2.

Шифрование. Адресат A шифрует свой текст t=ZAM и отправляет его адресату B. Адресат A выполняет следующее.

1. Представляет свой текст t=ZAM как элемент m группы G. Чтобы зашифровать письмо t, адресат A кодирует текст t каким-либо способом, например, в 27-ричной системе счисления 10-ричным числом $u=26\cdot 27^2+1\cdot 27+13=18994_{10}$, а затем вычисляет 13-ричное представление числа u в виде сообщения m = $(8\ 8\ 5\ 1)_{13}$, рассматриваемом как полином $8x^3+8x^2+5x+1$ из $\mathbb{Z}_{13}[x]$.
2. Получает открытый ключ (α=(0 0 1 5), y=(0 1 10 12)) адресата B.
3. Выбирает произвольное целое число k=2134, $1 \leqslant k \leqslant n-1$.
4. Вычисляет следующие элементы из G.

$\gamma = \alpha^k = (0\ 0\ 1\ 5)^{2134} = (x+5)^{2134} (\mathrm{mod}\ f(x)) =$
$8x^3+9x^2+7x+5 = (8\ 9\ 7\ 5)$, $y^k = (0\ 1\ 10\ 12)^{2134} =$
$(x^2+10x+12)^{2134} (\mathrm{mod}\ f(x)) = 10x^3+12x^2+3x+1 =$
(10 12 3 1),

$\delta = m\cdot y^k = (8\ 8\ 5\ 1)\cdot(10\ 12\ 3\ 1) = (8x^3+8x^2+5x+1)\cdot$
$(10x^3+12x^2+3x+1) (\mathrm{mod}\ f(x)) = 4x^3+6x^2+7x+3 = (4\ 6\ 7\ 3)$.

5. Отправляет шифротекст c = (γ=(8 9 7 5), δ=(4 6 7 3)) адресату B.

Дешифрование. Чтобы получить исходный текст t по c, адресат B выполняет следующее.

1. Пользуясь своим секретным ключом a, адресат B вычисляет следующие элементы группы G.

$\gamma^a = (8\ 9\ 7\ 5)^2 = (8x^3+9x^2+7x+5)^2 (\mathrm{mod}\ f(x)) =$
$10x^3+12x^2+3x+1 = (10\ 12\ 3\ 1)$,

$\gamma^{-a} = (\gamma^a)^{-1} = (10\ 12\ 3\ 1)^{-1} =$
$(10x^3+12x^2+3x+1)^{-1} (\mathrm{mod}\ f(x)) =$
$5x^3+7x^2+6x+11 = (5\ 7\ 6\ 11)$.

2. Вычисляет в группе G элемент

$m = (\gamma^{-a})\cdot\delta = (5\ 7\ 6\ 11)\cdot(4\ 6\ 7\ 3) = (5x^3+7x^2+6x+11)\cdot$
$(4x^3+6x^2+7x+3) (\mathrm{mod}\ f(x)) = 8x^3+8x^2+5x+1 = (8\ 8\ 5\ 1)$.

3. Чтобы получить текст t по элементу m, адресат B производит следующие вычисления.

$m = (8\ 8\ 5\ 1)_{13} = 8\cdot 13^3 + 8\cdot 13^2 + 5\cdot 13 + 1 = 18994_{10} =$
$(26\ 1\ 13)_{27}$, откуда текст t = ZAM.

Пример 2. Криптосистема ЭльГамаля с мультипликативной группой конечного поля $\mathbb{F}_{p^s}$, p=2, s=4. Пусть для удобства элемент поля $a_3x^3+a_2x^2+a_1x+a_0$ представляется бинарным стрингом $(a_3a_2a_1a_0)$.

Вычисление ключей. Адресат B выполняет следующее.

1. Выбирает мультипликативную группу $G=\mathbb{Z}_{2^4}^*$ конечного поля $(E_2^4,\{+,\cdot\})$, элементы которого представляются полиномами из $\mathbb{Z}_2[x]$ над $\mathbb{Z}_2$ степени меньше 4 и умножение в котором выполняется по модулю неприводимого полинома $f(x) = x^4+x+1$ из $\mathbb{Z}_2[x]$. Группа G имеет порядок n=15.
2. Находит генератор α=(0010) = $0\cdot x^3+0\cdot x^2+1\cdot x+0 = x$.
3. Выбирает случайное число $a = 7 \in [1,n-1]$.
4. Вычисляет $y = \alpha^a = \alpha^7 = x^7 (\mathrm{mod}\ f(x)) =$
$1\cdot x^3+0\cdot x^2+1\cdot x+1 = (1011)$.
5. Открытый ключ для B есть пара (α=(0010),y=(1011)) вместе с полиномом $f(x)$, который определяет умножение в G, если $f(x)$ и α не есть параметры, общие всем адресатам). Секретный ключ для B есть число a=7.

Шифрование. Чтобы зашифровать свое сообщение m=(1100), A получает открытый ключ (α=(0010),y=(1011)) адресата B, выбирает случайное целое число k=11 и вычисляет

$\gamma = \alpha^k = (0010)^{11} = x^{11} (\mathrm{mod}\ f(x)) =$
$x^3+x^2+x = (1110)$, $y^k = (1100)^{11} = (0100)$ и
$\delta = m\cdot(\alpha^a)^{11} = (x^3+x^2)(x^3+x+1) (\mathrm{mod}\ f(x)) =$

$x^2+1 = (0101)$.

A посылает шифротекст $c = (\gamma=(1110), \delta=(0101)$ адресату B.

Дешифрование. Чтобы дешифровать шифротекст c, B вычисляет
$\gamma^a = (1110)^7 = (x^3+x^2+x)^7 (\mathrm{mod}\ f(x)) = x^3 = (0100)$,
$(\gamma^a)^{-1} = (0100)^{-1} = (x^3)^{-1}\ (\mathrm{mod}\ f(x)) = x^3+x^2+1 = (1101)$,
$m = (\gamma^{-a})\cdot\delta = (1101)\cdot(0101) = (x^3+x^2+1)(x^2+1)(\mathrm{mod}\ f(x)) = x^3+x^2 = (1100)$.

8.13. Обобщенная электронная цифровая подпись ЭльГамаля с мультипликативной группой G поля Галуа $GF(p^m)$

Схема электронной цифровой подписи ЭльГамаля, основанная на числовой мультипликативной группе $\mathbb{Z}_p^*$, может быть обобщена на любую конечную мультипликативную абелеву группу G. Алгоритм подписи использует криптографическую хэш–функцию $w = h(t)$, где t есть текст, а число $w=h(t)$ из $\mathbb{Z}_n$, где n есть число элементов в G, есть значение хэш–функции h на t. Предполагается, что каждый элемент r из G может рассматриваться как текст, для которого тоже может быть вычислено значение хэш-функции $h(r)$.

Алгоритм вычисления хэш-функции публикуется.

Криптографическая стойкость подписи основана на трудной осуществимости проблемы нахождения дискретного логарифма в группе G большого порядка.

При использовании схемы цифровой подписи ЭльГамаля по тексту письма t вычисляется значение хэш-функции $h(t)$, которое затем используется при вычислении и проверке цифровой подписи под текстом письма.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать подходящую (мультипликативную) циклическую группу G порядка n.
2. Найти генератор α группы G.
3. Выбрать случайное число a, $1 \leqslant a \leqslant n-1$.
4. Вычислить элемент $y = \alpha^a$ группы G.
5. Открытый ключ адресата есть пара (α,y) элементов группы G. Открыто также описание умножения элементов в G. Секретный ключ адресата есть число a.

Вычисление подписи. Адресат A подписывает свой текст t

(произвольной длины). Любой адресат B может проверить подпись адресата A под его текстом t. Адресат A должен выполнить следующее.

1. Вычислить значение хэш-функции $h(t)$.
2. Выбрать случайное секретное целое число k из $[1, n-1]$, для которого нод$(k,n) = 1$.
3. Вычислить целое число $k^{-1}\ (\mathrm{mod}\, n)$.
4. Вычислить элемент $r = \alpha^k$ группы G.
5. Вычислить значение хэш-функции $h(r)$.
6. Вычислить число $s = k^{-1}(h(t) - ah(r))\ (\mathrm{mod}\, n)$.
7. Подпись адресата A под его письмом t есть пара (r,s).

Проверка подписи. Чтобы проверить подпись (r,s) адресата A под его текстом t, адресат B должен выполнить следующее.

1. Вычислить значение хэш-функции $h(t)$.
2. Получить открытый ключ (α,y) для A.
3. Вычислить значение хэш-функции $h(r)$.
4. Вычислить в группе G элементы
$v_1 = y^{h(r)}\cdot r^s$ и $v_2 = \alpha^{h(t)}$.
5. Принять подпись, если $v_1=v_2$ и отвергнуть в противном случае.

Пример 1. Схема электронной (цифровой) подписи ЭльГамаля с мультипликативной группой конечного поля $\mathbb{F}_{p^s}$, p=13, s=4. Пусть для удобства элемент поля $a_3x^3+a_2x^2+a_1x+a_0$ представляется p–ричным стрингом $(a_3a_2a_1a_0)$.

Адресат A подписывает свой текст t (произвольной длины). Любой адресат B может проверить подпись A.

Вычисление ключей. Адресат A выполняет следующее.

1. Выбирает мультипликативную группу $G=\mathbb{Z}_{13^4}-\{0\}$ конечного поля $(\mathbb{Z}_{13^4},\{+,\cdot\})$, элементы которого представляются полиномами из $\mathbb{Z}_{13}[x]$ над $\mathbb{Z}_{13}$ степени меньше 4 и умножение в котором выполняется по модулю неприводимого полинома $f(x) = (1\ 0\ 1\ 0\ 2) = x^4+x^2+2$ из $\mathbb{Z}_{13}[x]$. Группа G имеет порядок $n = p^s-1 = 13^4-1 = 28560$.
2. Находит генератор $\alpha = x+5 = (0\ 0\ 1\ 5)$.
3. Выбирает случайное число $a = 2 \in [1,n-1]$.
4. Вычисляет $y = \alpha^a = \alpha^2 = (x+5)^2(\mathrm{mod}\ f(x)) = x^2+10x+12 = (0\ 1\ 10\ 12)$.
5. Открытый ключ для A есть пара (α=(0 0 1 5), y=(0 1 10 12)) вместе с полиномом $f(x)$, который определяет умножение в G, если $f(x)$ и α не есть параметры, общие всем адресатам).

Секретный ключ для A есть число a=2.

Вычисление подписи. Адресат A подписывает свой текст t (произвольной длины). Адресат A выполняет следующее.

1. Вычисляет значение хэш-функции $h(t)$. Пусть для примера $h(t)$ = 13708.

2. Выбирает случайное секретное целое число k=2141, $1 \leqslant k \leqslant n-1$, такое, что нод$(k,n)$ = 1.

3. Вычисляет в группе G элемент $r = \alpha^k = (0\ 0\ 1\ 5)^{2141} = (x+5)^{2141} \pmod{f(x)} = (3\ 8\ 0\ 4)$.

4. Вычисляет целое число
$k^{-1} \pmod n = 2141^{-1} \pmod n = 16421$.

5. Вычисляет значение хэш-функции $h(r)$, например, следующим образом. По r = (3 8 0 4) вычисляет в $\mathbb{Z}_n$ 10-ричное число $(3\ 8\ 0\ 4)_{13} = 3p^3+8p^2+4 = 3\cdot 13^3 + 8\cdot 13^2 + 4 = 7947_{10}$. Пусть для примера $h(r) = 7947_{10}$.

6. Вычисляет в $\mathbb{Z}_n$ число $s = k^{-1}(h(t) - ah(r)) \pmod n = 16421\cdot(13708 - 2\cdot 7947) \pmod n = 3614$.

7. Подпись адресата A под его текстом t есть пара (r=(3 8 0 4), s=3614_{10}).

Проверка подписи. Чтобы проверить подпись (r,s) адресата A под его письмом t, адресат B выполняет следующее.

1. Вычисляет значение хэш-функции $h(t)$. Если текст t не изменялся, то $h(t) = 13708_{10}$.

2. Получает открытый ключ (α=(0 0 1 5), y=(0 1 10 12)) адресата A.

3. Вычисляет значение хэш-функции $h(r)$. Если вектор r не изменялся, то $h(r) = 7947_{10}$.

4. Вычисляет в группе G элементы
$v_1 = y^{h(r)}\cdot r^s = (0\ 1\ 10\ 12)^{7947}\cdot(3\ 8\ 0\ 4)^{3614} = (x^2+10x+12)^{7947}\cdot(3x^3+8x^2+4)^{3614} \pmod{f(x)} = (6\ 2\ 5\ 12)$,
$v_2 = \alpha^{h(m)} = (0\ 0\ 1\ 5)^{13708} = (x+5)^{13708} \pmod{f(x)} = (6\ 2\ 5\ 12)$.

5. Так как $v_1 = v_2$, то B принимает подпись адресата A.

Пример 2. Вычисление ключей. Рассмотрим конечное поле $\mathbb{F}_{2^5}$, построенное с помощью неприводимого полинома $f(x) = x^5+x^2+1$ над $\mathbb{Z}_2$. Элементы этого поля есть 32 набора из 0 и 1 длины 5 с нулем 00000. Элемент α = (00010) есть генератор мультипликативной циклической группы $G = \mathbb{F}^*_{2^5}$ поля. Порядок группы G есть n=31. Адресат A выбирает число a = 19 и вычисляет $y = \alpha^a = (00010)^{19} = (00110)$. Открытый ключ адресата A есть пара

наборов из 0 и 1 длины пять (α=(00010), y=(00110)). Секретный ключ для A есть число a = 19.

Вычисление подписи. Чтобы подписать текст m = 10110101, адресат A выбирает случайное число k = 24 и вычисляет $r = \alpha^{24}$ = (11110) и $k^{-1} \pmod{31} = 22$. Потом адресат A вычисляет $h(m)$ = 16, $h(r)$ = 7 (значения хэш-функции не связано с сообщением m и вектором r и взято в качестве примера) и $s = 22\cdot(16 - 19\cdot 7) \pmod{31} = 30$. Подпись адресата A под сообщением m есть пара (r=(11110), s=30).

Проверка подписи. Адресат B вычисляет

$h(m) = 16$, $h(r) = 7$,
$v_1 = y^{h(r)}r^s = (00110)^7\cdot(11110)^{30} = (11011)$,
$v_2 = \alpha^{h(m)} = \alpha^{16} = (11011)$.

Так как $v_1 = v_2$, то B принимает подпись адресата A.

Замечание. При вычислении подписи используются вычисления в группе G и вычисления в $\mathbb{Z}_n$. При проверке подписи используются только вычисления в группе G.

8.14. Электронная цифровая подпись DSA

Схема цифровой подписи DSA (Digital Signature Algorithm) есть вариант цифровой подписи ЭльГамаля. Схема цифровой подписи DSA основывается на трудности вычисления дискретного логарифма в $\mathbb{Z}^*_p$. Остается недоказанным, что подпись DSA криптографически стойка, даже если было бы известно, что проблема дискретного в $\mathbb{Z}^*_p$ есть трудная проблема. Схема цифровой подписи DSA требует использования хэш-функции.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать простое число q, $2^{159} < q < 2^{160}$.

2. Выбрать число t, $0 \leqslant t \leqslant 8$, и простое число p, $2^{511+64t} < p < 2^{512+64t}$ такое, что q делит $p-1$.

3. Найти генератор $\alpha \in \mathbb{Z}^*_p$ для циклической подгруппы порядка q в группе $\mathbb{Z}^*_p$. Для этого адресат должен выполнить следующее.

   3.1. Выбрать элемент $g \in \mathbb{Z}^*_p$ и найти $\alpha = g^{(p-1)/q} \pmod p$.

   3.2. Если α=1, то перейти к шагу 3.1 с другим g.

4. Выбрать произвольное число a, $1 \leqslant a \leqslant q-1$.

5. Вычислить $y = \alpha^a \pmod p$.

6. Открытый ключ адресата есть (p,q,α,y); секретный ключ адресата есть число a.

Вычисление подписи. Адресат A подписывает свой текст t (произвольной длины). Любой адресат B может проверить подпись A под текстом t с помощью открытого ключа адресата A. Адресат A выполняет следующее.

1. Вычисляет значение хэш-функции $h(t)$.
2. Выбрать произвольное секретное число k, $0<k<q$.
3. Вычислить $k^{-1} \pmod q$.
4. Вычислить $r=(\alpha^k \pmod p) \pmod q$.
5. Вычислить $s=k^{-1}(h(t)+ar) \pmod q$.
6. Подпись адресата A есть пара чисел (r,s).

Проверка подписи. Чтобы проверить подпись (r,s) адресата A под его текстом t, адресат B должен выполнить следующее.

1. Вычислить значение хэш-функции $h(t)$.
1. Взять открытый ключ (p,q,α,y) адресата A.
2. Проверить, что $0 < r < q$ и $0 < s < q$. Если нет, то отвергнуть подпись.
3. Вычислить $w = s^{-1} \pmod q$ и $h(m)$.
4. Вычислить $u_1 = w \cdot h(m) \pmod q$ и $u_2 = rw \pmod q$.
5. Вычислить $v = (\alpha^{u_1} y^{u_2} \pmod p) \pmod q$.
6. Принять подпись, если $v = r$ и отвергнуть в противном случае.

Доказательство. Если (r,s) есть подпись адресата A на сообщении m, то должно быть $h(m) \equiv -ar+ks \pmod q$. Умножим обе части этого сравнения на w и получим $w \cdot h(m)+arw \equiv k \pmod q$, что есть просто $u_1+au_2 \equiv k \pmod q$. Возвышая α в степень, равную обеим частям этого сравнения, получим

$$(\alpha^{u_1} y^{u_2} \pmod p) \pmod q \equiv (\alpha^k \pmod p) \pmod q.$$

Следовательно, $v = r$.

Пример. Адресат A подписывает свой текст t и всякий адресат B может проверить подпись A.

Вычисление ключей. Адресат A делает следующее.

1. Выбирает простое число q=27367.
2. Выбирает простое число p=656809, для которого q делит $(p-1)$. Пусть $(p-1)/q = 24$.
3. Выбирает случайное число $g = 2732 \in \mathbb{Z}_p^*$ и вычисляет $\alpha = g^{(p-1)/q} \pmod p = 2732^{24} \pmod{656809} = 68909$. Так как $\alpha \neq 1$, то α есть генератор для единственной циклической подгруппы порядка q в группе $\mathbb{Z}_p^*$. (Если α=1, то следует выбрать другое g).
4. Выбирает случайное число $a = 80 \in [1,q-1]$.
5. Вычисляет $y = \alpha^a \pmod p = 68909^{80} \pmod{656809} = 50951$.
6. Открытый ключ адресата A есть (p=656809, q=27367, α=68909, y=50951). Секретный ключ адресата A есть a=80.

Вычисление подписи. Чтобы подписать свой текст t (произвольной длины), адресат A делает следующее.

1. Вычисляет значение хэш-функции $h(t)$. Пусть для примера $h(t) = 1499$.
2. Выбирает случайное секретное число $k = 74 \in [0,q]$.
3. Вычисляет $k^{-1} \pmod q = 21080$.
4. A вычисляет $r = (\alpha^k \pmod p) \pmod q = (68909^{74} \pmod{656809}) \pmod{27367} = 145325 \pmod{27367} = 8490$.
5. A вычисляет $s = k^{-1} \cdot (h(t)+ar) \pmod q = 21080 \cdot (1499 + 80 \cdot 8490 \pmod{27367} = 14746$.
6. Подпись A под его текстом t есть пара чисел (r = 8490, s=14746).

Проверка подписи. Чтобы проверить подпись (r = 8490, s = 14746) адресата A под его текстом t, адресат B выполняет следующее.

1. Вычисляет значение хэш-функции $h(t)$. Если текст t не изменялся, то $h(t) = 1499$.
2. Берет открытый ключ адресата A: (p=656809, q=27367, α=68909, y=50951).
3. Проверяет, что $r = 8490 \in [0,q] = [0,27367]$, $s = 9 \in [0,q] = [0,27367]$. Если проверка не проходит, то подпись отвергнуть.
4. Вычисляет $w = s^{-1} \pmod q = 15699$.
5. Вычисляет
   $u_1 = w \cdot h(t) \pmod q = 15699 \cdot 1499 \pmod{27367} = 24548$,
   $u_2 = rw \pmod q = 8490 \cdot 15699 \pmod{27367} = 7220$.
6. Вычисляет $v = (\alpha^{u_1} y^{u_2} \pmod p) \pmod q = (68909^{24548} \cdot 50951^{7220} \pmod{656809}) \pmod{27367} = (280146 \cdot 334407 \pmod{656809}) \pmod{27367} = 145325 \pmod{27367} = 8490$.
6. Так как $v = 8490 = r$, то B принимает подпись A.

Для криптографической стойкости рекомендуется брать q длиной 160 бит, размер p при любом кратном 64 лежит между 512 (лучше 768) и 1024 бит включительно.

8.15. Криптосистема Рабина

Желательное свойство схемы шифрования есть наличие доказательства, что взлом схемы так же труден, как и трудность решения какой-либо другой известной трудно вычислимой проблемы, например, проблемы факторизации целых чисел или проблемы вычисления дискретного алгоритма. Есть пока недоказанное предположение, что взлом схемы RSA так же труден, как трудность факторизации целых чисел. Схема Рабина была первой схемой с доказанной трудностью взлома, равной трудности факторизации.

Вычисление ключей. Каждый адресат вычисляет свой открытый ключ и соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать два больших различных случайных простых числа p и q, примерно одного размера.
2. Вычислить $n = pq$.
3. Открытый ключ адресата есть n. Секретный ключ адресата есть пара (p,q).

Шифрование. Адресат A шифрует свой текст t и отправляет шифротекст адресату B, который адресат B дешифрует. Адресат A делает следующее.

1. Каким-либо способом представляет свой текст t как натуральное число m из сегмента $[0,n-1]$.
2. Берет открытый ключ n для B.
3. Вычисляет $c = m^2 \pmod n$.
4. Отправляет шифротекст c адресату B.

Дешифрование. Чтобы дешифровать шифротекст c от A, адресат B делает следующее.

1. Решает уравнение $c \equiv m^2 \pmod n$ и, используя свой секретный ключ (p,q), находит четыре квадратных корня m_1, m_2, m_3, m_4 из c по модулю n. (Число c имеет один или два корня из c по модулю n, если нод$(m,n) \neq 1$, что возможно с очень малой вероятностью).
2. Отосланное сообщение было одно из чисел m_1,m_2,m_3,m_4. Адресат B каким-то образом решает (позже мы покажем как),

какое из этих сообщений есть m.

Замечание. Если p и q выбраны сравнимыми с 3 по модулю 4, то алгоритм вычисления четырех квадратных корней из c по модулю n можно упростить следующим образом.

1. Применив расширенный алгоритм Евклида, найти целые числа a и b, для которых $ap + bq = 1$.
2. Вычислить $r = c^{(p+1)/4} \pmod p$.
3. Вычислить $s = c^{(q+1)/4} \pmod q$.
4. Вычислить $x = (aps + bqr) \pmod n$.
5. Вычислить $y = (aps - bqr) \pmod n$.
6. x, $-x$, y, $-$y по модулю n есть четыре квадратных корня из c по модулю n.

Пример. Адресат A посылает текст t=RAB адресату B.

Вычисление ключей. Адресат B выполняет следующее.

1. Выбирает два разных простых числа p=2131, q=2437.
2. Вычисляет $n = p \cdot q = 5193247$.
3. Открытый ключ адресата B есть число n=5193247. Секретный ключ адресата B есть пара чисел (p=2131, q=2437).

Шифрование. Адресат A выполняет следующее.

1. Представляет свой текст t=RAB в виде натурального числа m из $[0,n-1]$ с помощью какого-либо метода, например, с помощью 27-ричной системы счисления: m=18 · 27²+1 · 27+2=13151.
2. Удваивает слово из двух последних цифр справа в m и получает $m1$ = 1315151.
3. Получает открытый ключ n=5193247 адресата B.
4. Вычисляет $c = m1^2 \pmod n = 1315151^2 \pmod{5193247} = 852957$.
5. Посылает шифротекст c = 852957 адресату B.

Дешифрование. Чтобы дешифровать шифротекст c, адресат B делает следующее.

1. Решает уравнение $c \equiv m^2 \pmod n$ относительно m, то есть находит четыре квадратных корня из c по модулю n:
$m_1 = 1315151$, $m_2 = -1315151$, $m_3 = 2050346$, $m_4 = -2050346$.
2. Прибавляет к отрицательным корням модуль n. Тогда
$m_1 = 1315151$, $m_2 = 3878096$, $m_3 = 2050346$, $m_4 = 3142901$.

Так как только m_1 имеет справа два одинаковых 2-буквенных слова, то B дешифрует c как m_1 и получает $m = 13151_{10}$.

В 27-ричной записи $m = (18\ 1\ 2)_{27}$, откуда t=RAB.

Замечание. Шифрование по схеме Рабина довольно быстро,

ибо используется только одна модульная операция возведения в квадрат. Дешифрование осуществляется медленнее, чем шифрование, но по скорости сравнимо с дешифрованием по схеме RSA.

8.16. Электронная цифровая подпись Рабина с извлечением сообщения

Криптосистема цифровой подписи Рабина подобна схеме RSA. Пространство подписи M_φ есть Q_n (множество квадратичных вычетов по $\bmod n$) и подписи есть квадратные из них корни по $\bmod n$. Функция $R: M \to M_\varphi$ из пространства сообщений M в подписываемое пространство M_φ публикуется.

Вычисление ключей. Каждый адресат вычисляет свой открытый ключ и соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать два больших различных случайных простых числа p и q примерно одного размера.
2. Вычислить $n = pq$.
3. Открытый ключ адресата есть n. Секретный ключ адресата есть пара чисел (p,q).
4. Публикуемая функция R определяется следующим образом. Если число m есть подписываемое сообщение, то $R(m)$ есть ближайшее к m число, для которого существуют четыре различных корня $\sqrt{R(m)} \pmod n$. (Такое число существует). При этом $R(m) > m$, если четыре корня $\sqrt{R(m)} \pmod n$ для $m<n$ существуют, и $R(m) < m$ в противном случае.

Вычисление подписи. Адресат A представляет свой текст t в виде сообщения (числа) и подписывает его. Любой адресат B может проверить подпись адресата A и извлечь из подписи текст t. Адресат A должен сделать следующее.

1. Представить свой текст t в виде натурального числа m с помощью какого-либо метода.
2. Вычислить $w = R(m)$ и $i = R(m)-m$.
3. Вычислить $s = \sqrt{w} \pmod n$.
4. Подпись A под сообщением m есть пара чисел (s,i).

Проверка подписи. Чтобы проверить подпись (s,i) адресата A и извлечь из s текст t, адресат B должен сделать следующее.

1. Получить открытый ключ n адресата A.
2. Вычислить $w = s^2 \pmod n$.

3. Проверить, что $w \in M_R$. Если нет, то подпись отвергнуть.
4. Получить $m = R^{-1}(w)-i$ и найти текст t.

Пример. Вычисление ключей. Адресат A делает следующее.

1. Выбирает два различных случайных простых числа p=1699, q=1597 примерно одного размера.
2. Вычисляет $n = pq = 1699 \cdot 1597 = 2713303$.
3. Открытый ключ для A есть n=3185549. Секретный ключ для A есть пара чисел (p=1699, q=1597).

Вычисление подписи. Адресат A подписывает свой текст t = RAD и делает следующее.

1. Представляет свой текст t=RAD в виде натурального числа m с помощью какого-либо метода, например, с помощью 27-ричной системы счисления: $m1 = 18 \cdot 27^2+1 \cdot 27+4 = 13153$.
2. Удваивает правое слово из двух букв в $m1$, приписывает его к $m1$ справа и получает m=1315353.
3. Вычисляет $w = R(m) = R(1315353) = 1315358$ и $i=R(m)-m = 1315358 - 1315353 = 5$.
4. Вычисляет $\mathbf{s1} = \sqrt{w} \pmod n = \sqrt{1315358} \pmod{2713303} = (2264649, -2264649, 399147, -399147)$.
5. К каждой отрицательной компоненте в **s1** прибавляет модуль n и получает **s** = (2264649, 448654, 399147, 2314156).
6. Подпись адресата A есть любое число в **s** и число i, например, (s=2314156, i=5).

Проверка подписи. Чтобы проверить подпись (s=2314156, i = 5) адресата A и извлечь из s текст t, адресат B должен сделать следующее.

1. Получить открытый ключ n=3185549 адресата A.
2. Вычислить $w = s^2 \pmod n = 2314156^2 \pmod{3185549} = 1315358$, $m = w-i = 1315358-3 = 1315353$.
3. Так как m справа имеет сдвоенное двухбуквенное слово 53, то подпись принимается.
4. $m1 = 13153 = (18\ 1\ 4)_{27}$ и потому текст t=RAD.

8.17. Модифицированная цифровая подпись Рабина с извлечением сообщения

Представленная здесь техника подобна технике в цифровой подписи ISO/IEC 9796. Последняя предлагает детерминированный метод для связи сообщения с элементами пространства подписей M_S с таким расчетом, чтобы вычисление квадратного корня (или

чего-то близкого к корню) было всегда возможно.

Утверждение. Пусть p и q есть различные простые числа, каждое из которых сравнимо с 3 по модулю 4. Пусть $n = pq$.

1. Если нод$(x,n) = 1$, то $x^{(p-1)(q-1)/2} \equiv 1 \pmod n$.

2. Если $x \in Q_n$, то $x^{(n-p-q+5)/8} \pmod n$ есть $\sqrt{x} \pmod n$.

3. Если x есть целое, для которого символ Якоби $\left(\frac{x}{n}\right) = 1$, и если $d = (n-p-q+5)/8$, то $x^{2d} \pmod n = \begin{cases} x, & \text{if } x \in Q_n, \\ n-x, & \text{if } x \notin Q_n. \end{cases}$

4. Если $p \not\equiv q \pmod 8$, то $\left(\frac{2}{n}\right) = -1$ и потому умножение любого целого x на 2 или на $2^{-1} \pmod n$ обращает символ Якоби для x. (Целые вида $n = pq$, где $p \equiv q \equiv 3 \pmod 4$ и $p \not\equiv q \pmod 8$ иногда называют целыми Вильямса.)

Пространство сообщений $M = \{m \in \mathbb{Z}_n : m \leqslant \lfloor (n-6)/16 \rfloor\}$.

Подписываемое пространство $M_S = \{m \in \mathbb{Z}_n : m \equiv 6 \pmod{16}\}$.

Пространство подписей $S = \{s \in \mathbb{Z}_n : (s^2 \bmod n) \in M_S\}$.

Функция $R(m) = 16m+6 \ \forall m \in M$.

$M_R = \{m \in \mathbb{Z}_n : m \equiv 6 \pmod{16}\}$ есть множество значений для R.

Вычисление ключей. Каждый адресат вычисляет свой открытый ключ и соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать два различных случайных простых числа $p \equiv 3 \pmod 8$, $q \equiv 7 \pmod 8$.

2. Вычислить $n = pq$.

3. Вычислить $d = (n-p-q+5)/8$.

4. Открытый ключ адресата есть n. Секретный ключ адресата есть число d.

Вычисление подписи. Адресат A представляет свой текст t в виде сообщения (числа) и подписывает его. Любой адресат B может проверить подпись адресата A и извлечь из подписи текст t. Адресат A должен сделать следующее.

1. Представить свой текст t в виде натурального числа $m \leqslant \lfloor (n-6)/16 \rfloor$ с помощью какого-либо метода.

2. Вычислить $w = R(m) = 16m + 6$.

3. Вычислить символ Якоби $J = \left(\frac{w}{n}\right)$.

4. Если $J = 1$, то вычислить $s = w^d \pmod n$.

5. Если $J = -1$, то вычислить $s = (w/2)^d \pmod n$.

Если $J \neq 1$ или -1, то $J = 0$ и потому нод$(w,n) \neq 1$. Это потребует факторизации числа n. Вероятность факторизации большого n на практике ничтожно мала.

6. Подпись адресата A есть число s.

Проверка подписи. Чтобы проверить подпись s адресата A и извлечь из s текст t, B должен сделать следующее.

1. Получить открытый ключ n адресата A.

2. Вычислить $u = s^2 \pmod n$.

3. Если $u \equiv 6 \pmod 8$, то взять $w = u$.

4. Если $u \equiv 3 \pmod 8$, то взять $w = 2u$.

5. Если $u \equiv 7 \pmod 8$, то взять $w = n-u$.

6. Если $u \equiv 2 \pmod 8$, то взять $w = 2(n-u)$.

7. Если $w \in M_R$, принять подпись. Если нет, отвергнуть подпись.

8. Получить $m = R^{-1}(w) = (w-6)/16$.

Доказательство. Подпись s зависит от знака символа Якоби для $v = w$ и $v = w/2$. Только одно из w, $w/2$ имеет символ Якоби 1. Значение v таково, что $v = 3$ или $6 \pmod 8$. $s^2 \pmod n$ равно v или $n-v$ в зависимости от принадлежности или непринадлежности v к Q_n. Последнее может быть установлено, ибо $n \equiv 5 \pmod 8$.

Пример (*модифицированная схема подписи Рабина*).

Вычисление ключей. Адресат A делает следующее.

1. Выбирает два различных случайных простых числа p=1811$\equiv 3 \pmod 8$, q=1759 $\equiv 7 \pmod 8$ примерно одного размера.

2. Вычисляет $n = pq = 1811 \cdot 1759 = 3185549$.

3. Вычисляет $d = (n-p-q+5)/8 = (3185549-1811-1759+5)/8 = 397748$.

4. Открытый ключ для A есть n=3185549. Секретный ключ для A есть число d=397748.

Вычисление подписи. Адресат A подписывает свой текст t = RAB и делает следующее.

1. Представляет свой текст t=RAB в виде натурального числа m с помощью какого-либо метода, например, с помощью 27-ричной системы счисления: $m = 18 \cdot 27^2+1 \cdot 27+2 = 13151$.

2. Вычисляет $w = R(m) = 16m+6 = 16 \cdot 13151+6 = 210422$.

3. Вычисляет символ Якоби $J = \left(\frac{w}{n}\right) = \left(\frac{210422}{3185549}\right) = -1$.

4. Так как $J=-1$, то находит $s = (w/2)^d \pmod n = (210422/2)^{397748} \pmod{3185549} = 548579$.

5. Подпись адресата A под сообщением m есть s=548579.

Проверка подписи. Чтобы проверить подпись s=548579 адресата A и извлечь из s текст t, B делает следующее.

1. Получает открытый ключ n=3185549 адресата A.

2. Вычисляет $u = s^2 \pmod n = 548579^{3185549} \pmod{3185549} = 105211$.

3. Так как $u \equiv 3 \pmod 8$, то берет $w = 2u = 2 \cdot 105211 = 210422$.

4. Так как $w \equiv 6 \pmod{16}$ и $w \in M_R$, то принимает подпись.

5. Получает $m = R^{-1}(w)=(w-6)/16 = (105211-6)/16 = 13151$.

6. Получает $m = (18\ 1\ 2)_{27}$ и текст t = RAB.

Замечание. (*Образцы значений параметров для ISO/IEC 9796*)

Следующая таблица приводит образцы значений параметров в в процессе получения подписи для 150-битового сообщения и 1024-битовой подписи.

Параметр	k (бит)	d (бит)	z (байт)	r (бит)	t (байт)
Значение	1024	150	19	3	64

8.18. Криптосистема МакЭлиса

Криптографическая схема МакЭлиса основана на кодах, корректирующих ошибки. Выбирается конкретный код, для которого известен эффективный декодирующий алгоритм. Затем код маскируется общим линейным кодом. Так как проблема декодирования произвольного линейного кода является **NP**-полной (трудной), то описание исходного кода может служить секретным ключом, а описание трансформированного кода – открытым ключом.

Схема шифрования МакЕлиса (с кодами Гоппа) успешно противостоит взлому до сих пор. Это первая схема шифрования, которая использовала в процессе шифрования рандомизацию. Схема МакЭлиса довольно эффективна, но из-за большой длины открытого ключа она используется в практике не очень часто.

Натуральные числа k,n,t фиксируются как общие параметры системы и публикуются.

Вычисление ключей. Каждый адресат создает свой открытый

ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать порождающую $k \times n$-матрицу G для некоторого бинарного (n,k)-линейного кода, исправляющий t ошибок, и для которого известен эффективный декодирующий алгоритм.

2. Выбрать случайную бинарную неособенную квадратную $k \times k$-матрицу S.

3. Выбрать случайную перестановочную $n \times n$-матрицу P.

4. Вычислить $k \times n$-матрицу $G' = SGP$.

5. Открытый ключ адресата есть пара (G',t). Секретный ключ адресата есть тройка (S,G,P).

Шифрование. A шифрует свое сообщение m для B, которое B дешифрует. A делает следующее.

1. A получает открытый ключ (G',t) для B.

2. A представляет свое сообщение как бинарный стринг m длины k.

3. A выбирает случайный бинарный вектор ошибок z длины n, имеющий самое большее t единиц.

4. A вычисляет бинарный вектор $c = mG'+z$.

5. A посылает шифротекст c адресату B.

Дешифрование. Чтобы получить исходный текст m из c, адресат B делает следующее.

1. B вычисляет $c' = cP^{-1}$, где P^{-1} есть обратная матрица для матрицы P.

2. B использует декодирующий алгоритм для кода, порожденного матрицей G и из c' получает m'.

3. B вычисляет $m = m'S^{-1}$.

Доказательство. Так как

$$c' = cP^{-1} = (mG'+z)P^{-1} = (mSGP+z)P^{-1} = (mS)G+zP^{-1},$$

и zP^{-1} есть вектор с максимумом t единиц, то декодирующий алгоритм для кода, порожденного G, корректирует c' до $m' = mS$. Наконец, $mS^{-1} = m$ и дешифрование выполнено.

Специальный тип кода, исправляющего ошибки, называемый *кодом Гоппа*, может быть использован в шаге 3 вычисления ключей. Для каждого неприводимого полинома $g(x)$ степени t над $\mathbb{F}_{2^m}$ существует бинарный код Гоппа длины $n = 2^m$ и размерности $k \geq n-mt$, способный исправить всякий пакет из t или менее ошибок. Для таких кодов известны эффективные алгоритмы.

Замечание. МакЭлис предложил параметры n=1024, t=50, k

$\geqslant 524$. Более поздние исследования показали, что для криптографической стойкости шифра можно взять n=1024, t=38, $k \geqslant 644$.

Схема шифрования и дешифрования МакЭлиса осуществляется относительно быстро. Недостаток метода состоит в громоздкости открытого ключа. Менее обременительный недостаток состоит в существовании мультипликативного коэффициента для n/k. Для рекомендуемых параметров n=1024, t=38, $k \geqslant 644$ размер открытого ключа равен 2^{19} бит при коэффициенте, равным 1.6. Из этих соображений метод МакЭлиса получил небольшое распространение.

8.19. Рюкзачная схема шифрования Меркле–Хеллмана

Рюкзачная криптосистема с открытым ключом основана на проблеме подмножества суммы, которая является **NP**–полной. Задача состоит в выборе легко решаемого примера проблемы подмножества суммы, который затем маскируется примером трудно решаемой общей проблемой подмножества суммы. Исходное рюкзачное множество может служить секретным ключом, в то время как трансформированное рюкзачное множество служит в качестве открытого ключа. Рюкзачная схема шифрования Меркле–Хеллмана важна из соображений истории, так как это была первая конкретная реализация схемы шифрования с открытым ключом. Потом было предложено много модификаций этой схемы шифрования, но большинство из них, включая первую, оказались криптографически нестойкими. Примечательным исключением оказалась рюкзачная схема шифрования Хора-Ривеста.

8.19.1. Базовая рюкзачная схема шифрования Меркле–Хеллмана

Рюкзачная схема шифрования Меркле–Хеллмана маскирует некоторый легко решаемый пример проблемы подмножества суммы, называемый супервозрастающей проблемы подмножества суммы, модулярным умножением и подстановкой. Однако это не рекомендуется к использованию.

Определение. *Супервозрастающая последовательность* есть последовательность $(b_1, b_2, \ldots, b_n)$ положительных целых чисел, обладающих свойством $b_i > \sum_{j=1}^{i-1} b_j$ для каждого i, $2 \leqslant i \leqslant n$.

Алгоритм. Решение супервозрастающей проблемы подмножества суммы.

ВХОД: Супервозрастающая последовательность $(b_1, b_2, \ldots, b_n)$

и целое число s, являющееся суммой подмножества b_i.

ВЫХОД: набор $(x_1, \ldots, x_n)$, где $x_i \in \{0,1\}$ и $\sum_{i=1}^{n} x_i b_i = s$.

1. $i := n$.
2. Пока $i \geqslant 1$, выполнять следующее.
   2.1. Если $s \geqslant b_i$, то $x_i := 1$, $s := s - b_i$. Иначе $x_i := 0$.
   2.2. $i := i-1$.
3. Вернуть $(x_1, x_2, \ldots, x_n)$.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Целое число n публикуется.
2. Выполнить шаги 3 – 7.
3. Выбрать супервозрастающую последовательность $(b_1, b_2, \ldots, b_n)$ и модуль M, для которого $M > b_1 + b_2 + \ldots + b_n$.
4. Выбрать случайное целое число W, $1 \leqslant W \leqslant M-1$, для которого нод$(W, M) = 1$.
5. Выбрать случайную подстановку π целых чисел $\{1, 2, \ldots, n\}$.
6. Вычислить $a_i = W \cdot b_{\pi(i)} \pmod{M}$ для $i = 1, 2, \ldots, n$.
7. Открытый ключ адресата есть набор $(a_1, a_2, \ldots, a_n)$. Секретный ключ адресата есть набор $(\pi, M, W, (b_1, b_2, \ldots, b_n))$.

Шифрование. Адресат A шифрует свое сообщение m к B, которое адресат B дешифрует. Адресат A делает следующее.

1. A получает открытый ключ $(a_1, a_2, \ldots, a_n)$ адресата B.
2. A представляет свое сообщение m как бинарный стринг $m = m_1 m_2 \ldots m_n$ длины n.
3. A вычисляет число $c = m_1 a_1 + m_2 a_2 + \ldots + m_n a_n$.
4. A посылает свой шифротекст c к B.

Дешифрование. Чтобы получить исходный текст m из c, адресат B делает следующее.

1. B вычисляет $d = W^{-1} c \pmod{M}$.
2. Решая супервозрастающую проблему подмножества суммы, B находит числа $r_1, r_2, \ldots, r_n$, $r_i \in \{0,1\}$, для которых $d = r_1 b_1 + r_2 b_2 + \ldots + r_n b_n$.
3. B находит биты $m_i = r_{\pi(i)}$, $i = 1, 2, \ldots, n$, исходного сообщения от A.

Доказательство. Дешифрование дает исходный текст, ибо $d \equiv W^{-1} c \equiv W^{-1} \sum_{i=1}^{n} m_i a_i \equiv \sum_{i=1}^{n} m_i b_{\pi(i)} \pmod{M}$.

Так как $0 \leqslant d < M$, то $d = \sum_{i=1}^{n} m_i b_{\pi(i)} \pmod M$, и потому решение супервозрастающей проблемы подмножества суммы на шаге 2 дешифрования дает быты сообщения после применения подстановки π.

Пример. Вычисление ключей. Пусть $n=6$. Адресат B выбирает супервозрастающую последовательность $(12,17,33,74,157,316)$, $M=737$, $W=635$, подстановку π для $\{1,2,3,4,5,6\}$, полагая $\pi(1)=3$, $\pi(2)=6$, $\pi(3)=1$, $\pi(4)=2$, $\pi(5)=5$, $\pi(6)=4$. Открытый ключ для B есть набор $(319,196,250,477,200,559)$. Секретный ключ для B есть набор $(\pi, M, W, (12,17,33,74,157,316))$.

Шифрование. Чтобы зашифровать битовое сообщение $m=101101$, адресат A вычисляет число $c = 319+250+477+559 = 1605$ и посылает шифротекст c адресату B.

Дешифрование. Чтобы дешифровать шифротекст c, адресат B вычисляет число $d = W^{-1}c \pmod M = 136$ и решает супервозрастающую проблему подмножества суммы

$$136 = 12r_1 + 17r_2 + 33r_3 + 74r_4 + 157r_5 + 316r_6,$$

получая $136 = 12+17+33+74$. Поэтому $r_1=1$, $r_2=1$, $r_3=1$, $r_4=1$, $r_5=0$, $r_6=0$. Применение подстановки π дает биты сообщения адресата A: $m_1 = r_3 = 1$, $m_2 = r_6 = 0$, $m_3 = r_1 = 1$, $m_4 = r_2 = 1$, $m_5 = r_5 = 0$, $m_6 = r_4 = 1$.

8.20. Рюкзачная схема шифрования Хора–Ривеста

Схема Хора–Ривеста есть единственно известная рюкзачная схема шифрования с открытым ключом, которая не использует модулярное умножение для маскировки легкой проблемы подмножества суммы.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать конечное поле $\mathbb{F}_q$ характеристики p, где $q = p^h$, $h \leqslant p$, и для которого легко решается проблема дискретного логарифма (см. в конце параграфа замечание, пункт 2).

2. Выбрать случайный нормированный неприводимый поли-ном $f(x)$ степени h над $\mathbb{Z}_p$ Элементы из $\mathbb{F}_q$ представляются полиномами в $\mathbb{Z}_p[x]$ степени меньше h. Умножение полиномов выполняется по модулю $f(x)$.

3. Выбрать случайный примитивный элемент $g(x)$ поля $\mathbb{F}_q$.

4. Для каждого элемента основного поля $i \in \mathbb{Z}_p$, A найти

дискретный логарифм $a_i = \log_{g(x)}(x+i)$ элемента поля $(x+i)$ по основанию $g(x)$.

5. Выбрать случайную подстановку π на множестве чисел $\{0,1,2,\ldots,p-1\}$.

6. Выбрать случайное целое число d, $0 \leqslant d \leqslant p^h-2$.

7. Вычислить $c_i = (a_{\pi(i)}+d) \pmod{(p^h-1)}$, $0 \leqslant i \leqslant p-1$.

8. Открытый ключ адресата есть набор $((c_0,c_1,\ldots,c_{p-1}), p, h)$. Секретный ключ адресата есть набор $(f(x), g(x), \pi, d)$.

Шифрование. Адресат A шифрует свое сообщение m к B, которое адресат B дешифрует. Адресат A делает следующее.

1. A получает открытый ключ $((c_0,c_1,\ldots,c_{p-1}),p,h)$ адресата B.

2. A представляет свое сообщение m в виде бинарного стринга длины $\lfloor \lg \binom{p}{h} \rfloor$, где $\binom{p}{h}$ есть биномиальный коэффициент.

3. A рассматривает m как бинарное представление числа. A преобразует это число в бинарный вектор $M = (M_0, M_1, \ldots, M_{p-1})$ длины p, имеющий в точности h единиц, следующим образом.

3.1. $l := h$.

3.2. Для i от 1 до p выполнить следующее.

Если $m \geqslant \binom{p-i}{l}$, то $M_{i-1} := 1$, $m := m - \binom{p-i}{l}$, $l := l-1$. Иначе $M_{i-1} := 0$. (Заметим, что $\binom{n}{0} = 1$ для $n \geqslant 0$; $\binom{0}{l} = 0$ для $l \geqslant 1$).

4. A вычисляет $c = \sum_{i=0}^{p-1} M_i c_i \pmod{(p^h-1)}$.

5. A посылает шифротекст c адресату B.

Дешифрование. Чтобы получить исходный текст m из c, адресат B делает следующее.

1. B вычисляет $r = (c-hd) \pmod{(p^h-1)}$.

2. B вычисляет $u(x) = g(x)^r \pmod{f(x)}$.

3. B вычисляет $s(x) = u(x)+f(x)$, нормированный полином степени h над $\mathbb{Z}_p$.

4. B факторизует $s(x)$ в произведение линейных множителей над $\mathbb{Z}_p$. Пусть тогда $s(x) = \prod_{j=1}^{h}(x+t_j)$, где $t_j \in \mathbb{Z}_p$.

5. B вычисляет бинарный вектор $M = (M_0, M_1, \ldots, M_{p-1})$ следующим образом. Компоненты в M, которые равны 1, имеют индексы $\pi^{-1}(t_j)$, $1 \leqslant j \leqslant h$. Остальные компоненты равны нулю.

6. B вычисляет сообщение m по M следующим образом.

6.1. $m := 0$, $l := h$.

6.2. Для i от 1 до p выполнить следующее.

Если $M_{i-1} = 1$, то $m := m+\binom{p-i}{l}$, $l := l-1$.

Доказательство. Заметим, что

$$u(x) = g(x)^r \pmod{f(x)}$$

$$\equiv g(x)^{c-hd} \equiv g(x)^{(\sum_{i=0}^{p-1} M_i c_i)-hd} \pmod{f(x)}$$

$$\equiv g(x)^{(\sum_{i=0}^{p-1}(a_{\pi(i)}+d)-hd} \pmod{f(x)}$$

$$\equiv g(x)^{\sum_{i=0}^{p-1} M_i a_{\pi(i)}} \pmod{f(x)}$$

$$\equiv \prod_{i=0}^{p-1}[g(x)^{a_{\pi(i)}}]^{M_i} \equiv \prod_{i=0}^{p-1}(x+\pi(i))^{M_i} \pmod{f(x)}.$$

Так как $\prod_{i=0}^{p-1}(x+\pi(i))^{M_i}$ и $s(x)$ есть нормированные полиномы степени h, сравнимые по модулю $f(x)$, то $s(x) = u(x) + f(x) = \prod_{i=0}^{p-1}(x+\pi(i))^{M_i}$. Следовательно, все h корней $s(x)$ лежат в $\mathbb{Z}_p$. Применение к этим корням подстановки π^{-1} дает равные единице координаты в M.

Пример. Вычисление ключей. Адресат B делает следующее.

1. B выбирает $p = 7$ и $h = 4$.
2. B выбирает неприводимый полином

$$f(x) = x^4 + 3x^3 + 5x^2 + 6x + 2$$

степение 4 over $\mathbb{Z}_7$. Элементы конечного поля $\mathbb{F}_{7^4}$ представимы полиномами в $\mathbb{Z}_7[x]$ степение меньше 4. Произведение выполняется по модулю $f(x)$.

3. B выбирает случайный примитивный элемент $g(x) = 3x^3 + 3x^2 + 6$.
4. B вычисляет следующие дискретные логарифмы.

$$a_0 = \log_{g(x)}(x) = 1028,$$
$$a_1 = \log_{g(x)}(x+1) = 1935,$$
$$a_2 = \log_{g(x)}(x+2) = 2054,$$
$$a_3 = \log_{g(x)}(x+3) = 1008,$$
$$a_4 = \log_{g(x)}(x+4) = 379,$$
$$a_5 = \log_{g(x)}(x+5) = 1780,$$
$$a_6 = \log_{g(x)}(x+6) = 223.$$

5. B выбирает случайную подстановку π на $\{0,1,2,3,4,5,6\}$, полагая $\pi(0)=6$, $\pi(1)=4$, $\pi(2)=0$, $\pi(3)=2$, $\pi(4)=1$, $\pi(5)=5$, $\pi(6) = 3$.

6. B выбирает случайное число $d = 1702$.
7. B вычисляет

$$c_0 = (a_6 + d) \pmod{2400} = 1925,$$
$$c_1 = (a_4 + d) \pmod{2400} = 2081,$$
$$c_2 = (a_0 + d) \pmod{2400} = 330,$$
$$c_3 = (a_2 + d) \pmod{2400} = 1356,$$
$$c_4 = (a_1 + d) \pmod{2400} = 1237,$$
$$c_5 = (a_5 + d) \pmod{2400} = 1082,$$
$$c_6 = (a_3 + d) \pmod{2400} = 310.$$

8. Открытый ключ для адресата B есть набор $((c_0,c_1,c_2,c_3, c_4,c_5,c_6), p=7,h=4)$. Секретный ключ для B есть набор $(f(x), g(x),\pi,d)$.

Шифрование. Чтобы зашифровать сообщение $m = 22$ for B, адресат B делает следующее.

1. A получает открытый ключ адресата B.
2. A представляет m как бинарный стринг $m = 10110$ длины $\lfloor \lg \binom{7}{4} \rfloor = 5$. (Заметим, что $\lfloor \lg \binom{7}{4} \rfloor = 5$).
3. A преобразует m в бинарный вектор $M = (1,0,1,1,0,0,1)$ длины 7.
4. A вычисляет $c = (c_0+c_2+c_3+c_6) \pmod{2400} = 1521$.
5. A посылает шифротекст $c = 1521$ адресату B.

Дешифрование. Чтобы дешифровать шифротекст c=1521, адресат B делает следующее.

1. B вычисляет $r = (c-hd) \pmod{2400} = 1913$.
2. B вычисляет $u(x) = g(x)^{1913} \pmod{f(x)} = x^3+3x^2+2x+5$.
3. B вычисляет $s(x) = u(x)+f(x) = x^4+4x^3+x^2+x$.
4. B факторизует $s(x) = x(x+2)(x+3)(x+6)$ (так что $t_1=0$, $t_2=2$, $t_3=3$, $t_4=6$).
5. Равные единице компоненты вектора M имеют индексы $\pi^{-1}(0)=2$, $\pi^{-1}(2)=3$, $\pi^{-1}(3)=6$, $\pi^{-1}(6)=0$. Следовательно $M=(1, 0,1,1,0,0,1)$.
6. B преобразует M в число m=22, являющееся исходным текстом адресата A.

Замечание. 1. Хотя схема Хор–Ривеста описана для случая простого числа p, ее можно распространить на случай основного поля $\mathbb{F}_q$ с $q=p^m$ при некотором целом положительным m. Тогда элементы из $\mathbb{F}_{q^h}$ $(h \leqslant q)$ представляются как полиномы степени h из $\mathbb{Z}_q[x]$.

2. Чтобы проблема дискретного логарифма на шаге 1 алгоритма вычисления ключей была осуществимой, число $p^{h-1}-1$ должно факторизоваться только с малыми факторами и тогда может быть использован алгоритм Полига–Хеллмана.

3. Для криптографической стойкости рекомендуется брать параметры $p \approx 200$, $h \approx 25$. Авторы метода предложили p=197, h=24. В этом случае наибольший простой фактор для $197^{24}-1$ есть 10316017 и плотность рюкзачного множества ≈ 1.077. Авторами предлагались другие возможные параметры: $\{q=p^1=211,\ h = 24\}$, $\{q=3^5,\ h=24\}$ (основное поле $\mathbb{F}_{3^5}$), и $\{p=2^8,\ h=25\}$ (основное поле $\mathbb{F}_{2^8}$).

4. Шифрование проходит достаточно быстро. Дешифрование много медленнее. Его узкое место есть вычисление $u(x)$. Корни полинома $s(x)$ можно найти непосредственной проверкой элементов из $\mathbb{Z}_p$.

5. Недостаток схемы есть большой размер открытого ключа, именно около $(ph \cdot \lg p)$ бит. Для параметров p=197, h=24 это около 36000 бит.

6. Схема перестает быть криптографически стойкой, если известны части секретного ключа.

8.21. Вероятностное шифрование с открытым ключом

Минимальное требование криптографической стойкости схемы шифрования есть трудность несанкционированного дешифрования перехваченного шифротекста. Иногда желательны более сильные требования к криптографической стойкости схемы.

Схемы RSA, Рабина, рюкзачные схемы шифрования детерминированы в том смысле, что при фиксированном открытом ключе исходный текст m всегда шифруется в один и тот же шифротекст c. Детерминированная схема может иметь, например, следующие недостатки.

1. Схема не стойка для всех возможных исходных текстов. Например, схема RSA сообщения 0 и 1 шифруются в 0 и 1 соответственно.

2. Иногда по шифротексту о шифровке можно получить частичную информацию. Например, если в RSA шифротекст $c = m^e \pmod n$ соответствует исходному тексту m, то при нечетном n символ Якоби $\left(\frac{c}{n}\right) = \left(\frac{m^e}{n}\right) = \left(\frac{m}{n}\right)^e = \left(\frac{m}{n}\right)$ о сообщении m.

3. Два одинаковых шифротекста соответствуют двум одинаковым исходным текстам.

Вероятностное шифрование использует рандомизацию для получения очень высокой степени стойкости.

Определение. Схема шифрования с открытым ключом *полиномиально стойка*, если противник не может в полиномиальное время выбрать два текста m_1 и m_2 и затем правильно отличить шифротексты для m_1 и m_2 с вероятностью значительно больше 1/2.

Определение. Схема шифрования с открытым ключом *семантически стойка*, если для всякого сообщения, если противник может в полиномиальное время получить частичную информацию об исходном тексте по шифротексту, то он может ее вычислить также в полиномиальное время без шифротекста.

Интуитивно, схема шифрования с открытым ключом семантически стойка, если по шифротексту нельзя в полиномиальное время получить частичную информацию об исходном тексте.

Утверждение. Схема шифрования с открытым ключом семантически стойка, если и только если она полиномиально стойка.

8.22. Вероятностная схема шифрования Голдвассера–Микали

Вероятностная схема шифрования Голдвассера–Микали семантически стойка в предположении трудности решения проблемы квадратичного вычета.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать два больших различных случайных простых числа p и q примерно одного размера.

2. Вычислить $n = pq$.

3. Выбрать в $\mathbb{Z}_n$ квадратичный невычет y по модулю n, для которого символ Якоби $\left(\frac{y}{n}\right) = 1$ (то есть y есть псевдоквадрат по модулю n). Для этого сначала (случайным подбором с использованием символа Лежандра) находят квадратичный невычет a по модулю p и квадратичный невычет b по модулю q. Затем с помощью алгоритма Гаусса вычисляют такое целое y, $0 \leqslant y \leqslant n-1$, что $y \equiv a \pmod p$ и $y \equiv b \pmod q$. Так как y есть квадратичный невычет по модулю p, то y есть квадратичный невычет по модулю n. По свойствам символов Лежандра и Якоби

$\left(\frac{y}{n}\right) = \left(\frac{y}{p}\right)\left(\frac{y}{q}\right) = (-1)(-1) = 1$. То есть y есть псевдоквадрат по модулю n.

4. Открытый ключ адресата есть пара (n,y). Секретный ключ адресата есть пара (p,q).

Шифрование. Адресат A шифрует свой текст T и отправляет шифротекст адресату B, которое B дешифрует. Адресат A должен выполнить следующее.

1. Получить открытй ключ (n,y) адресата B.
2. Представить свой текст T каким-либо методом M как бинарный стринг $m = m_0 m_1 \ldots m_t$.
3. Для i от 0 до t выполнить следующее.
   3.1. Взять случайное число $x \in \mathbb{Z}_n$.
   3.2. Если m_i=1, то $c_i := yx^2 \pmod n$. Иначе
   $c_i := x^2 \pmod n$.
4. Послать шифротекст $c=(c_0,c_1,\ldots,c_t)$ адресату B.

Дешифрование. Чтобы получить исходный текст t по шифротексту c, адресат B должен сделать следующее.

1. Для i от 0 до t выполнить следующее.
   1.1. Вычислить символ Лежандра $e_i = \left(\frac{c_i}{p}\right)$.
   1.2. Если $e_i = 1$, то $m_i := 0$. Иначе $m_i := 1$.
2. Положить $m=m_0 m_1 \ldots m_t$. Методом M восстановить текст T.

Доказательство. Если бит m_i=0, то $c_i = x^2 \pmod n$ есть квадратичный вычет по модулю n. Если бит m_i=1, то так как y есть псевдоквадрат по модулю n, то $c_i = yx^2 \pmod n$ есть тоже псевдоквадрат по модулю n. Число c_i есть квадратичный вычет по модулю n, если и только если c_i есть квадратичный вычет по модулю p, то есть если символ Лежандра $\left(\frac{c_i}{p}\right) = 1$. Так как B знает p, то он может вычислить этот символ Лежандра и получить бит m_i.

Пример. Вычисление ключей. Адресат B делает следующее.

1. Выбирает простые числа p=499, q=547.
2. Вычисляет $n = pq = 314869$.
3. Находит (подбором) квадратичный невычет a=231 по $\bmod p$ и квадратичный невычет b=426 по $\bmod q$ (то есть a и b со значениями JACOBI(a,p) = JACOBI(b,q) = −1). Методом Гаусса решает систему сравнений $y \equiv a \pmod p$, $y \equiv b \pmod q$ и находит y = 135460 $\in \mathbb{Z}_n$.

3. Открытый ключ адресата B есть (n=314869, y=135460). Секретный ключ адресата B есть (p=499, q=631).

Шифрование. Адресат A шифрует свой текст T=BLM и отправляет шифротекст адресату B, которое B дешифрует. Адресат A делает следующее.

1. Получает открытй ключ (n,y) адресата B.
2. Представляет свой текст T=BLM в виде натурального числа m с помощью какого-либо метода, например, с помощью 27-ричной системы счисления $m=2\cdot 27^2+12\cdot 27+13=1795$, затем по основанию 2 находит $m=11100000011_2$ как бинарный стринг $m = m_0 m_1 \ldots m_t$ при t=10.
3. Для i от 0 до t=10 берет случайные числа x_i из $\mathbb{Z}_n$ и вычисляет соответствующие c_i:

x_0=233486, $c_0 = yx_0^2 \pmod n$=206861, ибо m_0=1,
x_1=148997, $c_1 = yx_1^2 \pmod n$=252056, ибо m_1=1,
x_2=294740, $c_2 = yx_2^2 \pmod n$=236339, ибо m_2=1,
x_3=144311, $c_3 = x_3^2 \pmod n$=229061, ибо m_3=0,
x_4=150802, $c_4 = x_4^2 \pmod n$=144548, ибо m_4=0,
x_5=178883, $c_5 = x_5^2 \pmod n$=250695, ибо m_5=0,
x_6=214657, $c_6 = x_6^2 \pmod n$=13058, ибо m_6=0,
x_7= 25787, $c_7 = x_7^2 \pmod n$=280910, ибо m_7=0,
x_8=100889, $c_8 = x_8^2 \pmod n$=135027, ибо m_8=0,
x_9=109032, $c_9 = yx_9^2 \pmod n$=31370, ибо m_9=1,
x_{10}=249721, $c_{10} = yx_{10}^2 \pmod n$=264850, ибо m_{10}=1.

4. Посылает шифротекст $c = (c_0,c_1,\ldots,c_t)$ адресату B.

Дешифрование. Чтобы получить исходный текст t по шифротексту c, адресат B должен сделать следующее.

1. Для i от 0 до t выполнить следующее.
   1.1. Вычислить символ Лежандра $e_i = \left(\frac{c_i}{p}\right)$.
   1.2. Если $e_i = 1$, то $m_i := 0$. Иначе $m_i := 1$.
   e_i: −1,−1,−1,1,1,1,1,1,1,−1,−1,
   m_i: 1, 1, 1,0,0,0,0,0,0, 1, 1.
2. Положить $m=m_0 m_1 \ldots m_t=11100000011_2$. Методом M восстановить текст T: $m = \sum_{i=0}^{t} m_{t-i} \cdot 2^i = 1795_{10} = (2\ 12\ 13)_{27}$. Текст t=BLM.

8.23. Вероятностная схема шифрования Блюма–Голдвассера

Вероятностная схема шифрования Блюма-Голдвассера есть наиболее эффективная из известных вероятностных схем шифрования, по эффективности сравнимая со схемой шифрования RSA. Она семантически стойка в предположении трудности проблемы факторизации. Схема использует Блюм-Блюм-Шуб генератор для порождения псевдослучайной битовой последовательности, которая затем покоординатно XOR-суммируется с исходным текстом. Результирующая битовая последовательность вместе с шифрованием случайного "зерна" передается получателю, который использует свою информацию, чтобы вычислить "зерно", реконструировать псевдослучайную битовую последовательность и исходный текст.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать два больших случайных различных простых числа p и q, сравнимых с 3 по модулю 4.
2. Вычисить $n = pq$.
3. Использовать расширенный алгоритм Евклида и вычислить целые a и b, для которых $ap + bq = 1$.
4. Открытый ключ адресата есть n. Секретный ключ адресата есть набор (p,q,a,b).

Шифрование. Адресат A шифрует свое сообщение m и отправляет шифротекст адресату B, которое B дешифрует. Адресат A делает следующее.

1. A получает открытый ключ n адресата B.
2. Пусть $k = \lfloor \log_2 n \rfloor$ и $h = \lfloor \log_2 k \rfloor$. A представляет свое сообщение m как стринг $m = m_1 m_2 \ldots m_t$ длины t, где каждое m_i есть бинарный стринг длины h.
3. A выбирает в качестве "зерна" x_0 случайный квадратичный вычет по модулю n, взяв, например, случайное число $r \in \mathbb{Z}_n^*$ и положив $x_0 := r^2 \pmod n$.
4. Для i от 1 до t для A выполнить следующее.

   4.1. Вычислить $x_i = x_{i-1}^2 \pmod n$.

   4.2. Пусть r_i есть h наименьших значащих бит в в бинарном представлении числа x_i.

   4.3. Вычислить $c_i = r_i \oplus m_i$.

5. A вычисляет $x_{t+1} = x_t^2 \pmod n$.
6. A посылает шифротекст $c=(c_1,\ldots,c_t,x_{t+1})$ адресату B.

Дешифрование. Чтобы получить исходный текст m по шифротексту c, адресат B делает следующее.

1. B вычисляет $d_1 = ((p+1)/4)^{t+1} \pmod{(p-1)}$.
2. B вычисляет $d_2 = ((q+1)/4)^{t+1} \pmod{(q-1)}$.
3. B вычисляет $u = x_{t+1}^{d_1} \pmod p$.
4. B вычисляет $v = x_{t+1}^{d_2} \pmod q$.
5. B вычисляет $x_0 = vap + ubq \pmod n$.
6. Для i от 1 до t B должен сделать следующее.

   6.1. $x_i := x_{i-1}^2 \pmod n$.

   6.2. Пусть r_i есть h наименьших значащих бит в x_i.

   6.3. $m_i := r_i \oplus c_i$ (ибо $r_i \oplus c_i = r_i \oplus r_i \oplus m_i = m_i$).

Доказательство. Так как x_t есть квадратичный вычет по модулю n, то x_t есть квадратичный вычет по модулю p. Поэтому $x_t^{(p-1)/2} \equiv 1 \pmod p$. Заметим, что

$x_{t+1}^{(p+1)/4} \equiv (x_t^2)^{(p+1)/4} \equiv x_t^{(p+1)/2} \equiv x_t^{(p-1)/2} x_t \equiv x_t \pmod p$. Аналогично получаем, что

$x_t^{(p+1)/4} \equiv x_{t-1} \pmod p$. Тогда

$x_{t+1}^{((p+1)/4)^2} \equiv x_{t-1} \pmod p$.

Повторяя эти рассуждения, получаем следующее.

$u \equiv x_{t+1}^{d_1} \equiv x_{t+1}^{((p+1)/4)^{t+1}} \equiv x_0 \pmod p$. Аналогично

$v \equiv x_{t+1}^{d_2} \equiv x_0 \pmod q$.

Так как $ap + bq = 1$, то $vap + ubq \equiv x_0 \pmod p$ и $vap + ubq \equiv x_0 \pmod q$. Поэтому $x_0 \equiv vap + ubq \pmod n$ и адресат B вычисляет использованное адресатом A при шифровании случайное "зерно" x_0, по которому восстанавливается исходный текст.

Пример. Вычисление ключей. Адресат B выбирает простые числа p=499, q=547, сравнимые с 3 по модулю 4, и вычисляет $n = pq = 272953$. Используя расширенный алгоритм Евклида, B вычисляет целые a=–57, b=52 такие, что $ap + bq = 1$. Открытый ключ адресата B есть $n = 272953$. Секретный ключ адресата B есть набор (p,q,a,b).

Шифрование. Параметры k=18, h=4. Адресат A представляет

свое сообщение m как стринг $m = m_1m_2m_3m_4m_5$ (t=5), где m_1 = 1001, m_2=1100, m_3=0001, m_4=0000, m_5=1100. Адресат A выбирает случайный квадратичный вычет

$x_0 = 159201\ (= 3992 \pmod n)$ и вычисляет следующее.

i	$x_i = x_{i-1}^2 \pmod n$	r_i	$c_i = r_i \oplus m_i$
1	180539	1011	0010
2	193932	1100	0000
3	245613	1101	1100
4	130286	1110	1110
5	40632	1000	0100

и $x_6 = x_5^2 \pmod n = 139680$. Адресат A посылает шифротекст $c = (c_1$=0010,c_2=0000,c_3=1100,c_4=1110,c_5=0100,x_0=139680) адресату B.

Дешифрование. Чтобы дешифровать c, адресат B вычисляет:

$d_1 = ((p+1)/4)^6 \bmod (p-1) = 463,$

$d_2 = ((q+1)/4)^6 \bmod (q-1) = 337,$

$u = x_6^{463} \pmod p = 20,$

$v = x_6^{337} \pmod p = 24,$

$x_0 = vap + ubq \pmod n = 159201.$

Адресат B использует x_0 и получает все x_i и r_i так же, как это делал адресат A при шифровании. Затем B получает m_i XOR-суммированием r_i с блоками c_i шифротекста.

Замечание. Для криптографической стойкости модуль n следует брать длиной 1024 бит. Например, если n есть 1025-битовое число, то можно взять k=1024 и h=10.

8.24. Электронная цифровая подпись Фейге–Фиата–Шамира

Схема цифровой подписи Фейге–Фиата–Шамира есть некоторая модификация более ранней схемы Фиата–Шамира цифровой подписи. Схема требует хэш-функции h: $\{0,1\}^* \to \{0,1\}^k$. Здесь $\{0,1\}^k$ есть множество всех битовых стрингов длины k и $\{0,1\}^*$ есть множество всех битовых стрингов произвольной длины.

Криптографическая стойкость схемы основана на трудности вычисления (дискретного) квадратного корня по модулю n.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать случайные различные секретные простые числа p,q и вычислить $n = p \cdot q$.
2. Выбрать положительное целое число k и случайные различные целые числа $s_1, s_2, \ldots, s_k \in \mathbb{Z}_n^*$.
3. Вычислить $v_j = s_j^{-2} \pmod n$, $1 \leqslant j \leqslant k$,
4. Открытый ключ адресата есть набор $((v_1, v_2, \ldots, v_k), n)$. Секретный ключ адресата есть k-набор $(s_1, s_2, \ldots, s_k)$.

Вычисление подписи. Адресат A подписывает свое бинарное сообщение m произвольной длины. Всякий адресат B может проверить подпись A с помощью открытого ключа адресата A. Адресат A делает следующее.

1. A выбирает случайное целое r, $1 \leqslant r \leqslant n-1$.
2. A вычисляет $u = r^2 \pmod n$.
3. A вычисляет $e = (e_1, e_2, \ldots, e_k) = h(m||u)$, где каждое $e_i \in \{0,1\}$. (Здесь $m||u$ означает конкатенацию m и u, то есть приписывание u к m справа. Например, $ab||aca = abaca$).
4. A вычисляет $s = r \cdot \prod_{j=1}^{k} s_j^{e_j} \pmod n$.
5. Подпись адресата A под текстом m есть пара (e,s).

Проверка подписи. Чтобы проверить подпись (e,s) адресата A под текстом m, адресат B делает следующее.

1. B получает открытый ключ $((v_1, \ldots, v_k), n)$ адресата B.
2. B вычисляет $w = s^2 \cdot \prod_{j=1}^{k} v_j^{e_j} \pmod n$.
3. B вычисляет $e' = h(m||w)$.
4. B принимает подпись, если и только если $e = e'$.

Доказательство. $w \equiv s^2 \cdot \prod_{j=1}^{k} v_j^{e_j} \equiv r^2 \cdot \prod_{j=1}^{k} s_j^{2e_j} \cdot \prod_{j=1}^{k} v_j^{e_j} \equiv r^2 \cdot \prod_{j=1}^{k} (s_j^2 v_j) \equiv r^2 \equiv u \pmod n$. Следовательно, $w = u$ и потому $e = e'$.

Пример. Вычисление ключей. Адресат A берет простые числа p=3571, q=4523 и вычисляет $n = pq$ = 16151633. A выбирает положительное целое число k=5 и случайные различные целые числа s_1=42, s_2=73, s_3=85, s_4=101, s_5=150 из $\mathbb{Z}_n^*$. Адресат A выполняет следующие вычисления.

j	1	2	3	4	5
s_j	42	73	85	101	150
$s_j^{-1} \pmod n$	4999315	885021	6270634	13113207	11090788
$v_j = s_j^{-2} \pmod n$	503594	4879739	7104483	1409171	6965302

Открытый ключ для A есть набор $((v_1,v_2,\ldots,v_k),n)$ = ((503594,4879739,7104483,1409171,6965302), 16151633).

Секретный ключ для A есть набор $(s_1,s_2,\ldots,s_k)$ = (42,73,85,101,150).

Вычисление подписи. Пусть h: $\{0,1\}^* \to \{0,1\}^5$ есть хэш-функция. A выбирает случайное целое r = 23181 $\in [1,n-1]$ и вычисляет $u = r^2 \pmod n$ = 4354872. A вычисляет $e = h(m||u)$ = $s_1s_2s_3s_4s_5$=10110 (значение хэш-функции взято искусственно, это произвольный стринг из 0 и 1 длины 5). A вычисляет

$$s = rs_1^{e_1}s_2^{e_2}s_3^{e_3}s_4^{e_4}s_5^{e_5} \pmod n = r\,s_1^1 s_2^0 s_3^1 s_4^1 s_5^0 \pmod n =$$

$$23181 \cdot 42 \cdot 85 \cdot 101 \pmod n = 7978909.$$

Подпись A под текстом m есть (e=10110, s=7978909).

Проверка подписи. B вычисляет $s^2 \pmod n$ = 2926875 и $v_1v_3v_4 \pmod n$ = 503594 · 7104483 · 1409171) (mod n) = 15668174. B вычисляет $w = s^2v_1v_3v_4 \pmod n$ = 4354872. Так как $w = u$, то $e' = h(m||w) = h(m||u) = e$, e'=e. Следовательно, адресат B принимает подпись адресата A.

Замечание. Для криптографической стойкости модуль n следует брать длиной 1024 бит.

Если n есть t-битовое число, то секретный ключ имеет размер kt бит. Его можно уменьшить выбором случайных чисел s_j, $1 \leqslant j \leqslant k$, битовой длины $t' < t$. Однако число t' не может быть слишком малым во избежание возможности вычисления чисел s_j. Размер открытого ключа есть $(k+1)t$ бит. Например, если t=768 и k=128, то секретный и открытый ключи требуют 98304 и 99072 бит соответственно.

8.25. Электронная цифровая подпись GQ

Схема Гилу–Куискуатера (Guillou–Quisquater, GQ) цифровой подписи требует хэш-функции h: $\{0,1\}^* \to \mathbb{Z}_n$, где n есть некоторое положительное целое число. Схема основана на трудности решения проблемы факторизации целых чисел.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать случайные различные секретные простые числа p,q и вычислить $n = pq$.
2. Выбрать целое $e \in [1,n-1]$, для которого нод$(e,(p-1)\cdot(q-1)) = 1$.
3. Выбрать целое J_A, $1 \leqslant J_A \leqslant n$, для которого нод$(J_A,n) = 1$. (Бинарное представление для J_A можно использовать для передачи информации об адресате, такой как имя, адрес, номер водительских прав и т.д.)
4. Найти такое целое $a \in \mathbb{Z}_n$, что $J_A \cdot a^e \equiv 1 \pmod n$, и сделать это следующим образом.
   4.1. Вычислить $J_A^{-1} \pmod n$.
   4.2. Вычислить
   $d_1 = e^{-1} \pmod{(p-1)}$ и $d_2 = e^{-1} \pmod{(q-1)}$.
   4.3. Вычислить
   $a_1 = (J_A^{-1})^{d_1} \pmod p$ и $a_2 = (J_A^{-1})^{d_2} \pmod q$.
   4.4. Найти такое a, что одновременно
   $a \equiv a_1 \pmod p$ и $a \equiv a_2 \pmod q$.
5. Открытый ключ адресата есть набор (n,e,J_A). Секретный ключ адресата есть число a.

Вычисление подписи. Адресат A подписывает бинарное сообщение m произвольной длины. Всякий адресат B может проверить подпись адресата A. Адресат A делает следующее.

1. A выбирает случайное целое число k и вычисляет $r = k^e \pmod n$.
2. A вычисляет $l = h(m||r)$.
3. A вычисляет $s = ka^l \pmod n$.
4. Подпись адресата A под текстом m есть пара (s,l).

Проверка подписи. Чтобы проверить подпись (s,l) адресата A под (бинарным) текстом m, адресат B делает следующее.

1. B получает открытый ключ (n,e,J_A) адресата A.
2. B вычисляет $u = s^e(J_A)^l \pmod n$ и $l' = h(m||u)$.
3. B принимает подпись A, если и только если $l = l'$.

Доказательство. Заметим, что

$$u \equiv s^e(J_A)^l \equiv (ka^l)^e(J_A)^l \equiv k^e(a^eJ_A)^l \equiv k^e \equiv r \pmod n.$$

Следовательно, $u = r$ и потому $l = l'$.

Пример. Вычисление ключей. Адресат A выбирает простые числа p=20849, q=27457 и вычисляет $n = pq$ = 572450993. A выбирает целые e=47, J_A=1091522, решает сравнение $J_Aa^e \equiv 1 \pmod n$ и получает a=214611724. Открытый ключ адресата A есть набор (n=572450993, e=47, J_A=1091522). Секретный ключ для A есть целое a=214611724.

Вычисление подписи. Чтобы подписать (бинарное) сообщение m=1101110001, адресат A выбирает случайное целое k=42134 и вычисляет $r = k^e \pmod n = 297543350$. Затем A вычисляет $l = h(m||r) = 2713833$ (значение хэш-функции взято искусственно) и

$$s = ka^l \pmod n = 42134 \cdot (214611724^{2713833}) \pmod n = 252000854.$$

Подпись адресата A под текстом m есть пара (s=252000854, l = 2713833).

Проверка подписи. B вычисляет

$s^e \pmod n = 252000854^{47} \pmod n = 398641962$,

$(J_A)^l \pmod n = 109152^{2713833} \pmod n = 110523867$,

$u = s^e(J_A)^l \pmod n = 297543350$.

Так как $u = r$, $l' = h(m||u) = h(m||r) = l$, то $l' = l$ и адресат B принимает подпись адресата A.

Замечание. Для криптографической стойкости модуль n следует брать модуль n длины $\geqslant$ 768 бит, число e длины $\geqslant$ 128 бит, значение хэш-функции от 128 до 160 бит. Тогда открытый ключ будет длины 896+u бит, где u есть число бит для представления J_A. Секретный ключ a был бы 768 бит.

8.26. Электронная цифровая подпись Шнорра

Схема Шнорра есть некоторая модификация схемы DSA без ограничений DSA на простые p и q при вычислении ключей. В схемы Шнорра, как и в схеме DSA, используется подгруппа порядка q группы $\mathbb{Z}_p^*$, где p есть некоторое большое простое число. Схема также требует хэш-функции h: $\{0,1\}^* \to \mathbb{Z}_q$. Схема основана на трудности решения проблемы дискретного логарифма.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать простые числа p и q, для которых q делит $p-1$.
2. Выбрать генератор $\alpha \in \mathbb{Z}_p^*$ для циклической подгруппы порядка q в группе $\mathbb{Z}_p^*$. Для этого адресат делает следующее.
   2.1. Выбрать элемент $g \in \mathbb{Z}_p^*$ и найти $\alpha = g^{(p-1)/q} \pmod p$.
   2.2. Если α=1, то перейти к шагу 2.1 с другим g.
3. Выбрать произвольное число a, $1 \leqslant a \leqslant q-1$.
4. Вычислить $y = \alpha^a \pmod p$.
5. Открытый ключ адресата есть (p,q,α,y). Секретный ключ

адресата есть число a.

Вычисление подписи. Адресат A подписывает бинарное сообщение m произвольной длины. Всякий адресат B может проверить подпись адресата A. Адресат A должен сделать следующее.

1. Выбрать случайное секретное целое k, $1 \leqslant k \leqslant q-1$.
2. Вычислить $r = \alpha^k \pmod p$, $e = h(m||r)$, $s = ae+k \pmod q$.
3. Подпись адресата A под текстом m есть пара (s,e).

Проверка подписи. Чтобы проверить подпись (s,e) адресата A под (бинарным) текстом m, адресат B должен сделать следующее.

1. Получить открытый ключ (p,q,α,y) адресата A.
2. Вычислить $v = \alpha^s y^{-e} \pmod p$, $e' = h(m||v)$.
3. Принять подпись A если и только если $e' = e$.

Доказательство. В подписи адресата A число $v \equiv \alpha^s y^{-e} \equiv \alpha^s \alpha^{-ae} \equiv \alpha^k \equiv r \pmod p$, откуда $h(m||v) = h(m||r)$ и $e' = e$.

Пример. Вычисление ключей. Адресат A выбирает простые числа p=129841, q=541, для которых $q \mid (p-1)$. Число $(p-1)/q$ = 240. Затем A выбирает случайное целое $g = 26346 \in \mathbb{Z}_p^*$ и вычисляет $\alpha = g^{(p-1)/q} \pmod p = 26346^{240} \pmod p = 26$. Так как $\alpha \neq 1$, то α порождает в $\mathbb{Z}_p^*$ единственную циклическую подгруппу порядка 541. Адресат A выбирает произвольное число $a = 423 \in [1,q-1]$ и вычисляет $y = \alpha^a \pmod p = 26^{423} \pmod p = 115917$. Открытый ключ адресата A есть набор (p=129841, q=541, α=26, y=115917). Секретный ключ для A есть число a = 423

Вычисление подписи. Чтобы подписать (бинарное) сообщение m=11101101, адресат A выбирает случайное число $k = 327 \in [1, q-1]$ и вычисляет $r = \alpha^k \pmod p = 26^{327} \pmod p = 49375$ и $e = h(m||r) = 155$ (значение хэш-функции взято искусственно). A вычисляет $s = ae+k \pmod q = 423 \cdot 155 + 327 \pmod{541} = 431$. Подпись адресата A под текстом m есть пара (s=431, e=155).

Проверка подписи. B вычисляет

$v = \alpha^s y^{-e} \pmod p = 26^{431} \cdot 115917^{-155} \pmod p = 49375$,

$e' = h(m||v) = 155$.

Так как $e = e'$, то B принимает подпись A.

Замечание. Для криптографической стойкости рекомендуется брать q длиной 160 бит, размер p лежит между 512 (лучше 768) и 1024 бит включительно.

8.27. Электронная цифровая подпись Ниберга–Рюппеля с извлечением сообщения

Схема Ниберга–Рюппеля есть некоторая модификация схемы DSA без ограничений DSA на простые p и q при вычислении ключей. Схема основана на трудности решения проблемы дискретного логарифма.

Вычисление ключей. Каждый адресат создает свой открытый ключ и ему соответствующий секретный ключ. Каждый адресат должен сделать следующее.

1. Выбрать простые числа p и q, для которых q делит $p-1$.
2. Выбрать генератор $\alpha \in \mathbb{Z}_p^*$ для циклической подгруппы порядка q в группе $\mathbb{Z}_p^*$. Для этого нужно сделать следующее.
   2.1. Выбрать элемент $g \in \mathbb{Z}_p^*$ и найти $\alpha = g^{(p-1)/q} \pmod p$.
   2.2. Если $\alpha=1$, то перейти к шагу 2.1 с другим g.
3. Выбрать произвольное число a, $1 \leqslant a \leqslant q-1$.
4. Вычислить $y = \alpha^a \pmod p$.
5. Открытый ключ адресата есть (p,q,α,y). Секретный ключ адресата есть число a.

Вычисление подписи. Адресат A подписывает сообщение m. Всякий адресат B может проверить подпись A и извлечь сообщение m из подписи. Адресат A должен сделать следующее.

1. Вычислить $m' = R(m)$.
2. Выбрать случайное секретное число k, $1 \leqslant k \leqslant q-1$, и вычислить $r = \alpha^{-k} \pmod p$.
3. Вычислить $e = m'r \pmod p$.
4. Вычислить $s = ae + k \pmod q$.
5. Подпись A под m есть пара (e,s).

Проверка подписи. Чтобы проверить подпись (e,s) адресата A под сообщением m, адресат B должен сделать следующее.

1. Получить открытый ключ (p,q,α,y) адресата A.
2. Проверить, что $e \in [1,p-1]$. Если нет, то отвергнуть подпись.
3. Проверить, что $s \in [0,q-1]$. Если нет, то отвергнуть подпись.
4. Вычислить $v = \alpha^s y^{-e} \pmod p$ и $m' = ve \pmod p$.
5. Проверить, что $m' \in M_R$. Если нет, то отвергнуть подпись.
6. Вычислить $m = R^{-1}(m')$.

Доказательство. Для подписи адресата A число $v \equiv \alpha^s y^{-e} \equiv \alpha^{s-ae} \equiv \alpha^k \pmod p$. Следовательно, $ve \equiv \alpha^k m' \alpha^{-k} \equiv m' \pmod p$, что и требовалось.

Пример. Вычисление ключей. Адресат A выбирает простые числа p=1256993, q=3571, где q делит $(p-1)$. Число $(p-1)/q = 352$. A выбирает случайное число $g = 42077 \in \mathbb{Z}_p^*$ и вычисляет $\alpha = g^{(p-1)/q} \pmod p = 42077^{352} \pmod p = 441238$. Так как $\alpha \neq 1$, то α порождает в $\mathbb{Z}_p^*$ единственную циклическую подгруппу порядка 3571. A выбирает случайное целое $a = 2774 \in [1,q-1]$ и вычисляет $y = \alpha^a \pmod p = 1013657$. Открытый ключ адресата A есть набор (p=1256993, q=3571, α=441238, y=1013657). Секретный ключ для A есть число $a = 2774$.

Вычисление подписи. Чтобы подписать сообщение m, адресат A вычисляет $m' = R(m) = 1147892$ (значение $R(m)$ взято искусственно). A выбирает случайное $k = 1001 \in [1,q-1]$ и вычисляет $r = \alpha^{-k} \pmod p = 441238^{-1001} \pmod p = 1188935$, $e = m'r \pmod p = 138207$, $s = ae + k \pmod q = 2774 \cdot 138207 + 1001 \pmod q = 1088$. Подпись адресата A под сообщением m есть пара (e=138207, s=1088).

Проверка подписи и извлечение сообщения. B вычисляет

$v = 441238^{1088} \cdot 1013657^{-138207} \pmod{1256993} = 504308$,

$m' = v \cdot 138207 \pmod{1256993} = 1147892$.

B проверяет, что $m' \in M_R$ и получает $m = R^{-1}(m')$).

Замечание. Для криптографической стойкости рекомендуется брать q длиной 160 бит, размер p лежит между 512 (лучше 768) и 1024 бит включительно.

9. РЕКУРРЕНТНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ В $\mathbb{R}$

9.1. Конечные разности

Пусть $\mathbb{N}=\{0,1,2,\ldots\}$ и $\mathbb{R}$ есть множества натуральных и вещественных чисел соответственно. Пусть функция $x(k): \mathbb{N} \to \mathbb{R}$ отображает $\mathbb{N}$ в $\mathbb{R}$. Функция $x(k)$ задает последовательность x_0, $x_1, x_2, \ldots, x_k, \ldots$ вещественных чисел $x_k = x(k)$, которую мы иногда тоже будем называть функцией.

Определение. Пусть функция $x_k = x(k)$.

Конечная разность 1-го порядка $\Delta^1 x_k = x_{k+1} - x_k$.

Если конечные разности порядка n уже определены, то конечные разности порядка $n+1$

$\Delta^{n+1} x_k = \Delta^1(\Delta^n x_k) = \Delta^n x_{k+1} - \Delta^n x_k$,

Замечание. 1. Конечная разность 0-го порядка $\Delta^0 x_k = x_k$.

2. Иногда вместо $\Delta^1 x_k$ пишут $\Delta^1 x(k,k+1)$, а вместо $\Delta^n x_k$ пишут $\Delta^n x(k,k+1,\ldots,k+n)$.

3. $\Delta^2 x_k = \Delta^1 x_{k+1} - \Delta^1 x_k = x_{k+2} - x_{k-1} - x_{k-1} + x_k$,

$$\Delta^2 x_k = x_{k+2} - 2x_{k-1} + x_k = \sum_{i=0}^{2} (-1)^{2-i} C_2^i x_{k+i}.$$

$$\Delta^3 x_k = x_{k+3} - 3x_{k-2} + 3x_{k-1} + x_k = \sum_{i=0}^{3} (-1)^{3-i} C_3^i x_{k+i}.$$

И так далее.

9.1.1. Свойства конечных разностей

Лемма 1. Для дискретной функции $x_k = x(k)$

$$\Delta^n x_k = \sum_{i=0}^{n} (-1)^{n-i} C_n^i x_{k+i}, \text{ где } C_n^i = \frac{n!}{i!(n-i)!}. \qquad (9.1)$$

Доказательство. Индукция по n.

Базис. n=0. $\Delta^0 x_k = x_k = (-1)^0 C_0^0 x_k = \sum_{i=0}^{0} (-1)^{0-i} C_n^i x_{k+i}$.

n=1. $\Delta^1 x_k = x_{k+1} - x_k = (-1)^1 C_1^0 x_k + (-1)^0 C_1^1 x_{k+1} = \sum_{i=0}^{1} (-1)^{n-1} C_n^i x_{k+i}$.

Предположение индукции. Предположим, что формула (9.1) справедлива для конечных разностей порядка n.

Шаг индукции. Покажем, что формула (9.1) справедлива для конечных разностей порядка n+1.

$\Delta^{n+1} x_k = \Delta^n x_{k+1} - \Delta^n x_k =$ [предположение индукции] $=$

$$\sum_{i=0}^{n} (-1)^{n-i} C_n^i x_{k+1+i} - \sum_{i=0}^{n} (-1)^{n-i} C_n^i x_{k+i} =$$

[выделим последнее слагаемое в 1-ой и 1-ое во 2-ой сумме]

$$(-1)^0 C_n^n x_{k+n+1} - (-1)^n C_n^0 x_k +$$

$$\sum_{i=0}^{n-1} (-1)^{n-i} C_n^i x_{k+1+i} - \sum_{i=1}^{n} (-1)^{n-i} C_n^i x_{k+i} =$$

$$(-1)^0 C_n^n x_{k+n+1} + (-1)^{n+1} C_n^0 x_k +$$

$$\sum_{i=1}^{n} (-1)^{n+1-i} C_n^{i-1} x_{k+i} + \sum_{i=1}^{n} (-1)^{n+1-i} C_n^i x_{k+i} =$$

[объединим две суммы под одним знаком]

$$(-1)^0 C_n^n x_{k+n+1} + \left\{ \sum_{i=1}^{n} (-1)^{n+1-i} (C_n^{i-1} + C_n^i) x_{k+i} \right\} + (-1)^{n+1} C_n^0 x_k =$$

$$(-1)^0 C_{n+1}^{n+1} x_{k+n+1} + \sum_{i=1}^{n} (-1)^{n+1-i} C_{n+1}^i x_{k+i} + (-1)^{n+1} C_{n+1}^0 x_k =$$

$$\sum_{i=0}^{n+1} (-1)^{n+1-i} C_{n+1}^i x_{k+i}$$ и формула (9.1) установлена.

Шаг индукции установлен. Лемма доказана.

Следствие. Оператор Δ^n обладает свойством линейности:

1) $\Delta^n(x_k + y_k) = \Delta^n x_k + \Delta^n y_k$,

2) $\Delta^n(cx_k) = c \cdot \Delta^n x_k \quad \forall c \in \mathbb{R}$.

Лемма 2. Для дискретной функции $x_k = x(k)$

$$x_{k+n} = \sum_{i=0}^{n} C_n^i \Delta^i x_k. \qquad (9.2)$$

Доказательство. Индукция по n.

n=0. $x_k = C_0^0 \Delta^0 x_k = \sum_{i=0}^{0} C_n^i \Delta^i x_k$.

n=1. $x_{k+1} = x_k + (x_{k+1}-x_k) = C_1^0\Delta^0 x_k + C_1^1\Delta^1 x_k = \sum_{i=0}^{1} C_n^i \Delta^i x_k$.

Предположение индукции. Допустим, что формула (9.2) справедлива для n.

Шаг индукции. Покажем формула (9.2) верна для n+1.

$x_{k+n+1} = x_{k+n}+(x_{k+n+1}-x_{k+n}) =$ [предположение индукции] =

$$\sum_{i=0}^{n} C_n^i \Delta^i x_k + \sum_{i=0}^{n} C_n^i \Delta^i x_{k+1} - \sum_{i=0}^{n} C_n^i \Delta^i x_k =$$

$$\sum_{i=0}^{n} C_n^i \Delta^i x_k + \sum_{i=0}^{n} (C_n^i \Delta^i x_{k+1} - C_n^i \Delta^i x_k) =$$

$$\sum_{i=0}^{n} C_n^i \Delta^i x_k + \sum_{i=0}^{n} C_n^i (\Delta^i x_{k+1} - \Delta^i x_k) =$$

$$\sum_{i=0}^{n} C_n^i \Delta^i x_k + \sum_{i=0}^{n} C_n^i \Delta^{i+1} x_k = \sum_{i=0}^{n} C_n^i \Delta^i x_k + \sum_{i=1}^{n+1} C_n^{i-1} \Delta^i x_k =$$

[выделим первое слагаемое в 1-й сумме и последнее в 2-й]

$$C_n^0 \Delta^0 x_k + \sum_{i=1}^{n} C_n^i \Delta^i x_k + \sum_{i=1}^{n} C_n^{i-1} \Delta^i x_k + C_n^n \Delta^{n+1} x_k =$$

$$C_n^0 \Delta^0 x_k + \sum_{i=1}^{n} (C_n^i + C_n^{i-1}) \Delta^i x_k + C_n^n \Delta^{n+1} x_k =$$

$$C_{n+1}^0 \Delta^0 x_k + \sum_{i=1}^{n} C_{n+1}^i \Delta^i x_k + C_{n+1}^{n+1} \Delta^{n+1} x_k = \sum_{i=0}^{n+1} C_{n+1}^i \Delta^i x_k.$$

Шаг индукции установлен. Лемма доказана.

9.2. Рекуррентные уравнения

Пусть $F(t_1,\ldots,t_{n+1})$: $\mathbb{R}^{n+1} \to \mathbb{R}$ есть некоторая $(n+1)$-местная функция.

Определение. Уравнение вида R

$$x(k+n) = F(k,x(k),x(k+1),\ldots,x(k+n-1)) \qquad (R)$$

с неизвестной функцией натурального аргумента $x(k)$: $\mathbb{N}\to\mathbb{R}$, $k\in\mathbb{N}=\{0,1,2,\ldots\}$, называется *рекуррентным* (возвратным).

Число n называется порядком уравнения (R).

Определение. Функция $x(k)$ есть *частное решение рекуррентного уравнения* R, если $x(k)$ удовлетворяет уравнению R для всякого натурального k.

Рекуррентное уравнение R имеет много решений. Достаточно задать произвольные начальные значения $x(0),x(1),\ldots,x(n-1)$, а затем по формуле R последовательно вычислять значения $x(n),x(n+1),x(n+2),\ldots,x(n+k),\ldots$

Всякая функция натурального аргумента $x(k)$ задает числовую последовательность X: $x_0,x_1,x_2,\ldots,x_k,\ldots$, для которой $x_k=x(k)$, k=0,1,2,... Ее можно записать как $x(0),x(1),x(2),\ldots,x(k),\ldots$

Определение. Числовая последовательность X есть *частное решение рекуррентного уравнения* R, если $x_{k+n} = F(k,x_k,x_{k+1},\ldots,x_{k+n-1})$, k=0,1,2,...

Определение. *Числовая последовательность* X (функция $x(k)$) *называется рекуррентной*, если существует рекуррентное соотношение, решением которого последовательность X (функция $x(k)$) является.

Определение. *Общее решение рекуррентного уравнения* R есть функция $G(k,C_1,\ldots,C_n)$ такая, что

1) $\forall\, c_1,\ldots,c_n \in \mathbb{R}$ функция $y(k)=G(k,c_1,\ldots,c_n)$ есть некоторое частное решение уравнения R;

2) для всякого частного решения $y(k)$ уравнения R существуют числа $c_1,\ldots,c_n \in \mathbb{R}$, для которых $y(k)=G(k,c_1,\ldots,c_n)$.

Не существует общего метода решения рекуррентных уравнений. Но некоторые его частные виды такие методы допускают.

Замечание. Рекуррентное уравнение R можно задать в виде

$$x(k+n) - F(k,x(k),x(k+1),\ldots,x(k+n-1)) = 0,$$

или в неявной форме

$$H(k,x(k),x(k+1),\ldots,x(k+n-1),x(k+n)) = 0,$$

где H есть некоторая n+2-местная функция $\mathbb{R}^{n+2} \to \mathbb{R}$.

Пусть F есть поле вещественных чисел.

Определим сумму двух функций (последовательностей) $x(k)$+ $y(k)$ и произведение $a\cdot x(k)$ на число a поля F обычным обра-

зом. Тогда множество функций (последовательностей) $x(k)$ с этими операциями образует линейное векторное пространство над полем F.

9.3. Линейные рекуррентные уравнения с переменными коэффициентами

Определение. Уравнения R_0 и R_1

$$L(x(k))=a_n(k)x(k+n)+a_{n-1}(k)x(k+n-1)+\ldots+a_0(k)x(k)=0, \qquad (R_0)$$

$$L(x(k))=a_n(k)x(k+n)+a_{n-1}(k)x(k+n-1)+\ldots+a_0(k)x(k)=f(k), \quad (R_1)$$

где $a_i(k)$: $\mathbb{N}\to\mathbb{R}$, $a_n\neq 0$, $i=0,1,\ldots,n$, $f(k)\not\equiv 0$ – известные функции, $x(k)$ – неизвестная функция, называются *линейными рекуррентными уравнениями* (ЛРУ), *однородным и неоднородным* соответственно *с переменными коэффициентами*. Коэффициент $a_n(k)$ называется *старшим*. Уравнение со старшим коэффициентом 1 называется *нормированным уравнением*.

Замечание. Разделив R_0 и R_1 на старший коэффициент $a_n(k)$, получим нормированные ЛРУ:

$$L(x(k))=x(k+n)+a_{n-1}(k)x(k+n-1)+\ldots+a_0(k)x(k) = 0, \text{ или}$$
$$x(k+n) = -a_{n-1}(k)x(k+n-1)-\ldots-a_0(k)x(k), \qquad (R_0')$$

$$L(x(k))=x(k+n)+a_{n-1}(k)x(k+n-1)+\ldots+a_0(k)x(k) = f(k) \qquad (R_1')$$

Порядок уравнения есть число n в R_0 и R_1.

Однородное ЛРУ $L(x(k)) = 0$ называется соответствующим неоднородному ЛРУ $L(x(k)) = f(k)$.

Последовательности, являющиеся решениями ЛРУ, иногда называют возвратными последовательностями.

Утверждение. Оператор $L(x)$ линеен.

Доказательство. Пусть $x(k)$ и $y(k)$ – произвольные функции и $c\in\mathbb{R}$. Тогда

$$1)\ L(x+y) = a_n\cdot(x(k+n) + y(k+n)) +$$
$$a_{n-1}\cdot(x(k+n-1) + y(k+n-1)) +\ldots+ a_0\cdot(x(k)+y(k)) =$$
$$a_n\cdot x(k+n) + a_{n-1}\cdot x(k+n-1) +\ldots+ a_0\cdot x(k) +$$
$$a_n\cdot y(k+n) + a_{n-1}\cdot y(k+n-1) +\ldots+ a_0\cdot y(k) = L(x) + L(y);$$

$$2)\ L(cx) = a_n cx(n+k) + a_{n-1}cx(k+n-1) +\ldots+ a_0 cx(k) =$$
$$c(a_n\cdot x(k+n) + a_{n-1}\cdot x(k+n-1) +\ldots+ a_0\cdot x(k)) = cL(x).$$

Теорема. Множество всех решений однородного ЛРУ образует линейное векторное пространство.

Доказательство. Пусть $x(k)$ и $y(k)$ – произвольные решения однородного ЛРУ R_0 и $c\in\mathbb{R}$. Тогда $L(x(k))=0$, $L(y(k))=0$ и

$$1)\ L(x+y) = L(x)+L(y) = 0+0 = 0.$$
$$2)\ L(cx) = cL(x) = c\cdot 0 = 0.$$

Множество M решений однородного ЛРУ замкнуто относительно линейных операций. Восемь аксиом линейного пространства выполняются для M как для подмножества линейного пространства всех функций $x(k)$. Следовательно, M – линейное пространство.

Теорема. Общее решение неоднородного ЛРУ есть сумма какого-либо его частного решения и общего решения соответствующего однородного ЛРУ.

Доказательство. Пусть $x(k)$ и $y(k)$ – пара решений неоднородного ЛРУ $L(x) = f(k)$. Тогда их разность есть решение однородного ЛРУ $L(x) = 0$, ибо

$$L(x-y) = L(x) - L(y) = f(k) - f(k) = 0.$$

Пусть $x_{оо} = G(C_1,\ldots,C_k,k)$ есть общее решение однородного ЛРУ R_0 и $x_{чн}(k)$ – какое либо частное решение неоднородного ЛРУ R_1. Покажем, что функция $x_{он} = x_{оо} + x_{чн}$ есть общее решение для R_1.

1. Функция $x_{он} = x_{оо} + x_{чн}$ есть решение ЛРУ R_1 при любых $C_1,\ldots,C_n$, ибо $L(x_{он}) = L(x_{оо} + x_{чн}) = L(x_{оо}) + L(x_{чн}) = 0 + f(k) = f(k)$.

2. Покажем, что для всякого частного решения $z(k)$ не-однородного ЛРУ R_1 найдутся числа $c_1,\ldots,c_k$, для которых $z(k) = G(k,c_1,\ldots,c_n) + x_{чн}(k)$.

Зафиксируем числа $d_1,\ldots,d_n$ и построим функцию $u(k) = G(d_1,\ldots,d_n,k) + x_{чн}(k)$. Разность $z(k) - u(k)$ есть решение однородного ЛРУ R_0 и потому найдутся числа $e_1,\ldots,e_n$, для которых функция $z(k) - u(k) = G(e_1,\ldots,e_n,k)$. Отсюда

$$z(k) = u(k)+G(e_1,\ldots,e_n,k) =$$
$$G(d_1,\ldots,d_n,k) + x_{чн}(k) + G(e_1,\ldots,e_n,k) = w(k) + x_{чн}(k),$$

где функция $w(k) = G(d_1,\ldots,d_n,k) + G(e_1,\ldots,e_n,k)$ как сумма двух решений R_0 есть снова решение R_0. Поэтому найдутся числа $c_1,\ldots,c_n$, для которых $w(k) = G(c_1,\ldots, c_n,k)$ и тогда $z(k) = G(k,c_1,\ldots,c_n) + x_{чн}(k)$.

Пусть $f_1(k),\ldots,f_n(k)$: $\mathbb{N}\to\mathbb{R}$ – некоторые функции.

Определение. Нетривиальная *система функций* $f_1(k),\ldots,$

$f_n(k)$ *линейно зависима*, если существуют числа $c_1,\dots,c_n$, не равные нулю одновременно, для которых

$c_1f_1(k) + \dots + c_nf_n(k) \equiv 0$ на $\mathbb{N}$.

Определение. *Система функций* $f_1(k),\dots,f_n(k)$ *линейно независима* (на $\mathbb{N}$), если тождество $c_1f_1(k) + \dots + c_nf_n(k) \equiv 0$ на $\mathbb{N}$ влечет $c_1 = \dots = c_n = 0$.

Определение. *Фундаментальная система решений* (ФСР) для однородного ЛРУ $L(x) = 0$ порядка n есть система из n его линейно независимых решений.

Замечание. ФСР для неоднородного ЛРУ $L(x) = f(k)$ совпадает с ФСР для соответствующего однородного уравнения.

Определение. ФСР $x_1(k),\dots,x_n(k)$ *нормирована*, если

$$x_i(j) = \delta_{ij} = \begin{cases} 1, & \text{если } i=j, \\ 0, & \text{если } i\neq j, \end{cases} \quad i=1,2,\dots,n,\ j=0,1,\dots,n-1.$$

Определение. *Определитель Казарати* (Казаратиан) системы функций $f_1,\dots,f_n$ в точках $t_k,t_{k+1},\dots,t_{k+n-1}$ (точка есть натуральное число) есть определитель

$$D = D(f_1,\dots,f_n|t_k,t_{k+1},\dots,t_{k+n-1}) = \begin{vmatrix} f_1(t_k) & f_2(t_k) & \dots & f_n(t_k) \\ f_1(t_{k+1}) & f_2(t_{k+1}) & \dots & f_n(t_{k+1}) \\ & & \dots & \\ f_1(t_{k+n-1}) & f_2(t_{k+n-1}) & \dots & f_n(t_{k+n-1}) \end{vmatrix}.$$

Теорема. Система решений $x_1(k),\dots,x_n(k)$ ЛРУ R_0 $L(x)=0$ есть ФСР $\longleftrightarrow$ Казаратиан $D = D(x_1,\dots,x_n|0,1,\dots,k+n-1)$ для этой системы решений в точках $0,1,\dots,n-1$ отличен от 0.

Доказательство. *Необходимость*. Пусть $x_1(k),\dots,x_n(k)$ есть ФСР для ЛРУ R_0. Покажем, что Казаратиан

$$D = D(x_1,\dots,x_n|0,\dots,n-1) \neq 0.$$

Допустим противное: $D=0$. Система линейных уравнений $\sum_{i=1}^{n} c_ix_i(j)=0$, $j=0,1,\dots,n-1$, имеет нетривиальное решение $c_1^0,\dots,c_n^0$, ибо ее определитель $D=0$. Тогда $\sum_{i=1}^{n} c_i^0x_i(j)=0$, $j=0,\dots,n-1$. Пусть $y(k) = \sum_{i=1}^{n} c_i^0x_i(k)$, $k=0,1,2,\dots$. Числа $y(j)=0$, $j=0,\dots,n-1$. Функция $y(k)$ есть решение R_0, ибо

$$L(y(k)) = a_n(k)y(k+n) + \sum_{j=1}^{n} a_{n-j}(k)y(k+n-j) =$$

$$a_n(k)\sum_{i=1}^{n} c_i^0x_i(k+n) + \sum_{j=1}^{n} a_{n-j}(k)\sum_{i=1}^{n} c_i^0x_i(k+n-j) =$$

$$a_n(k)\sum_{i=1}^{n} c_i^0x_i(k+n) + \sum_{i=1}^{n} c_i^0\sum_{j=1}^{n} a_{n-j}(k)x_i(k+n-j) =$$

$$\sum_{i=1}^{n} c_i^0\cdot\left(a_n(k)x_i(k+n) + \sum_{j=1}^{n} a_{n-j}(k)x_i(k+n-j)\right) =$$

$$\sum_{i=1}^{n} c_i^0\cdot L(x_i(k)) = 0.$$

Так как $y(0) = \dots = y(n-1) = 0$, то в силу вида R_0' функция $y(k)\equiv 0$ на $\mathbb{N}$. Значит, $y(k) = \sum_{i=1}^{n} c_i^0x_i(k)\equiv 0$ на $\mathbb{N}$ при нетривиальном наборе $(c_1^0,\dots,c_n^0)$ и потому система функций $x_1(k),\dots,x_1(k)$ линейно зависима на $\mathbb{N}$. Противоречие. Следовательно, Казаратиан $D \neq 0$.

Достаточность. Пусть для системы решений $x_1(k),\dots,x_n(k)$ для R_0 Казаратиан $D \neq 0$. Покажем, что система решений $x_1,\dots,x_n$ линейно независима. Допустим противное: система решений $x_1,\dots,x_n$ линейно зависима. Тогда существует нетривиальный набор чисел $c_1^0,\dots,c_n^0$, для которого $\sum_{i=1}^{n} c_i^0x_i(k) \equiv 0$ на $\mathbb{N}$. Тогда система линейных уравнений $\sum_{i=1}^{n} c_ix_i(j)=0$, $j=0,\dots,n-1$, имеет нетривиальное решение $c_1^0,\dots,c_n^0$, а потому ее определитель, который есть определитель Казарати, равен нулю. Противоречие. Следовательно, система решений $x_1,\dots,x_n$ линейно независима.

Замечание. 1. ФСР существует. В качестве начальных условий можно взять столбцы неравного нулю Казаратиана в точке 0 и с помощью R_0' достроить их до ФСР.

2. Если задать начальные данные:

$$\begin{aligned} &x_1(0)=1,\ x_1(1)=0,\ \dots,\ x_1(n-1)=0,\\ &x_2(0)=0,\ x_2(1)=1,\ \dots,\ x_2(n-1)=0,\\ &\qquad\dots\\ &x_n(0)=0,\ x_n(1)=0,\ \dots,\ x_n(n-1)=1, \end{aligned}$$

то можно построить нормированную ФСР $x_1(k),\dots,x_n(k)$, для

которой Казаратиан $D(x_1,\dots,x_n|0,1,\dots,n-1) = 1$.

Теорема. Общее решение однородного ЛРУ R_0

$x_{oo}(k) = C_1x_1(k) + \dots + C_nx_n(k)$,

где $x_1,\dots,x_n$ есть ФСР для R_0, а произвольные постоянные $C_1,\dots,C_n$ пробегают $\mathbb{R}$ независимо друг от друга.

Доказательство. x_{oo} есть решение для R_0 $\forall C_1,\dots,C_n$, ибо

$L(x_{oo}) = L(C_1x_1(k) + \dots + C_nx_n(k)) =$

$C_1L(x_1(k)) + \dots + C_nL(x_n(k)) = 0$.

Пусть $x(k)$ есть какое-либо частное решение R_0. Система уравнений $\sum_{i=1}^{n} C_ix_i(j) = x(j)$, $j=0,\dots,n-1$, имеет единственное решение $c_1,\dots,c_n$, ибо ее определитель есть не равный нулю Казаратиан. Пусть функция $y(k) = \sum_{i=1}^{n} c_ix_i(k)$, $k=0,1,\dots$ Так как $x(j) = y(j)$, $j=0,\dots,n-1$, и так как начальные условия полностью определяют функцию, то $x(k) = y(k) = \sum_{i=1}^{n} c_ix_i(k)$ $\forall k \in \mathbb{N}$. Следовательно, частное решение $x(k)$ есть линейная комбинация функций из ФСР.

Следствие. Линейное пространство решений для R_0 имеет размерность n. Общее решение неоднородного ЛРУ R_1

$x_{oн}(k) = x_{oo}(k) + x_{чн}(k)$, $k=0,1,2,\dots$.

Замечание. Если $x_0(k),\dots,x_{n-1}(k)$ есть нормированная ФСР, то есть если $x_i(j)=1$ при $i=j$, и $x_i(j)=0$ при $i\neq j$, то всякое другое решение $y(k)$ удовлетворяет равенству $y(k) = y(0)x_0(k) + y(1)x_1(k) + \dots + y(n-1)x_{n-1}(k)$, $k=0,1,\dots$ $n-1$.

Определение. Решение $x(k)$ неоднородного ЛРУ R_1 называется *главным*, если его начальные условия нулевые:

$x(i) = 0$, $i=0,\dots,n-1$.

Теорема. Пусть задано нормированное однородное ЛРУ

$$\begin{aligned} &x(k+n)+a_{n-1}(k)x(k+n-1)+\dots+a_0(k)x(k) = 0, \text{ или} \\ &a_{n-1}(k)x(k+n-1)+\dots+a_0(k)x(k) = -x(k+n). \end{aligned} \qquad (R_0')$$

Пусть $x_1(k),\dots,x_n(k)$ есть частные решения уравнения R_0' и $a_0(k) \neq 0$ $\forall k \in \mathbb{N}$. Тогда Казаратиан

$D = D(x_1,\dots,x_n|k,\dots,k+n-1) =$

$\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(x_1,\dots,x_n|0,\dots,n-1)$.

Доказательство. Так как $x_1,\dots,x_n$ есть частные решения уравнения R_0', то $\forall k=0,1,2,\dots$

$$\begin{aligned} &a_0(k)x_1(k) + \dots + a_{n-1}(k)x_1(k+n-1) = -x_1(k+n), \\ &a_0(k)x_2(k) + \dots + a_{n-1}(k)x_2(k+n-1) = -x_2(k+n), \\ &\dots \\ &a_0(k)x_n(k) + \dots + a_{n-1}(k)x_n(k+n-1) = -x_n(k+n), \end{aligned}$$

или короче

$$a_0(k)x_i(k) + \dots + a_{n-1}(k)x_i(k+n-1) = -x_i(k+n),$$
$$i = 1,2,\dots,n.$$

По правилу Крамера

$a_0(k) = \dfrac{D_n}{D} = \dfrac{D_n(x_1,\dots,x_n|k+n,k+2,\dots,k+n-1)}{D(x_1,\dots,x_n|k,\dots,k+n-1)}$, откуда

$D_n = a_0(k)\cdot D$, или

$$-\begin{vmatrix} x_1(k+n) & x_1(k+1) & \dots & x_1(k+n-1) \\ x_2(k+n) & x_2(k+1) & \dots & x_2(k+n-1) \\ & & \dots & \\ x_n(k+n) & x_n(k+1) & \dots & x_n(k+n-1) \end{vmatrix} =$$

$$a_0(k)\cdot\begin{vmatrix} x_1(k) & x_1(k+1) & \dots & x_1(k+n-1) \\ x_2(k) & x_2(k+1) & \dots & x_2(k+n-1) \\ & & \dots & \\ x_n(k) & x_n(k+1) & \dots & x_n(k+n-1) \end{vmatrix}.$$

Последовательно меняя местами столбец 1 в D_n со столбцами $2,3,\dots,n$, получим

$$(-1)^n\begin{vmatrix} x_1(k+1) & x_1(k+2) & \dots & x_1(k+n-1) & x_1(k+n) \\ x_2(k+1) & x_2(k+2) & \dots & x_2(k+n-1) & x_2(k+n) \\ & & \dots & & \\ x_n(k+1) & x_n(k+2) & \dots & x_n(k+n-1) & x_n(k+n) \end{vmatrix} =$$

$$a_0(k)\cdot\begin{vmatrix} x_1(k) & x_1(k+1) & \dots & x_1(k+n-1) \\ x_2(k) & x_2(k+1) & \dots & x_2(k+n-1) \\ & & \dots & \\ x_n(k) & x_n(k+1 & \dots & x_n(k+n-1) \end{vmatrix},$$

или

$(-1)^n D(x_1,\dots,x_n|k+1,\dots,k+n) =$

$a_0(k)\cdot D(x_1,\dots,x_n|k,\dots,k+n-1)$,

откуда, умножая равенство на $(-1)^n$, получаем

$$D(x_1,\dots,x_n|k+1,\dots,k+n) =$$
$$(-1)^n a_0(k)\cdot D(x_1,\dots,x_n|k,\dots,k+n-1).$$

Меняем k на $k-1$ и получаем

$$D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$(-1)^n a_0(k-1)\cdot D(x_1,\dots,x_n|k-1,\dots,k+n-2).$$

Снова применяем правило Крамера к системе

$$a_0(k-1)x_i(k-1) + \dots + a_{n-1}(k-1)x_i(k+n-2) = -x_i(k+n-1),$$
$$i = 1,2,\dots,n,$$

и получаем

$$(-1)^n D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$a_0(k-1)\cdot D(x_1,\dots,x_n|k-1,\dots,k+n-2),$$

откуда, умножая равенство на $(-1)^n$, получаем

$$D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$(-1)^n a_0(k-1)\cdot D(x_1,\dots,x_n|k-1,\dots,k+n-2),$$

Меняем k на $k-1$ и получаем

$$D(x_1,\dots,x_n|k-1,\dots,k+n-2) =$$
$$(-1)^n a_0(k-2)\cdot D(x_1,\dots,x_n|k-2,\dots,k+n-3).$$

Тогда

$$D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$(-1)^n(-1)^n a_0(k-1)a_0(k-2)\cdot D(x_1,\dots,x_n|k-2,\dots,k+n-3).$$

И так далее. Наконец, получим

$$D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(x_1,\dots,x_n|0,\dots,n-1).$$

Замечание. 1. Теорема верна при условии $a_0(k)\neq 0\ \forall k\in\mathbb{N}$.

2. Если система решений $x_1(k),\dots,x_n(k)$ фундаментальна, то Казаратиан D для нее в точках $k,\dots,k+n-1$ не равен нулю для всякого натурального k.

3. Если для системы решений $x_1(k),\dots,x_n(k)$ Казаратиан D не равен нулю в точках $k,\dots,k+n-1$ при некотоом $k\in\mathbb{N}$, то эта система решений фундаментальна.

4. В точках $k,\dots,k+n-1$ для n частных решений либо Казаратиан $D=0\ \forall k\in\mathbb{N}$ (система решений не есть ФСР), либо $D\neq 0\ \forall k\in\mathbb{N}$ (система решений есть ФСР).

5. Фундаментальная система решений ОЛРУ есть базис пространства его решений.

6. Если система решений $x_1(k),\dots,x_n(k)$ фундаментальна, то для любой другой фундаментальной системы решений $y_1(k),\dots,y_n(k)$ справедливо равенство

$$D(y_1,\dots,y_n|k,\dots,k+n-1) =$$
$$c\cdot D(x_1,\dots,x_n|k,\dots,k+n-1), \text{ где константа}$$
$$c = \frac{D(y_1,\dots,y_n|0,\dots,n-1)}{D(x_1,\dots,x_n|0,\dots,n-1)}.$$

В самом деле, по доказанной теореме

$$D_x = D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(x_1,\dots,x_n|0,\dots,n-1),$$
$$D_y = D(y_1,\dots,y_n|k,\dots,k+n-1) =$$
$$\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(y_1,\dots,y_n|0,\dots,n-1),$$

$$\frac{D_x}{D_y} = \frac{\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(x_1,\dots,x_n|0,\dots,n-1)}{\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(y_1,\dots,y_n|0,\dots,n-1)} =$$

$$\frac{D(x_1,\dots,x_n|0,\dots,n-1)}{D(y_1,\dots,y_n|0,\dots,n-1)} = \text{некоторой константе } c \text{ из } \mathbb{R},$$

откуда следует требуемое равенство.

7. Для нормированной системы решений $x_1(k),\dots,x_n(k)$ Казаратиан $D(x_1,\dots,x_n|0,\dots,n-1) = 1$. Тогда

$$D(x_1,\dots,x_n|k,\dots,k+n-1) =$$
$$\left((-1)^{nk}\prod_{j=0}^{k-1} a_0(j)\right)\cdot D(x_1,\dots,x_n|0,\dots,n-1) =$$
$$(-1)^{nk}\prod_{j=0}^{k-1} a_0(j).$$

Теорема. Общее решение неоднородного уравнения (R_1)

$$x_{\text{он}}(k) = C_1x_1(k) + \dots + C_nx_n(k) + x_{\text{чн}}(k) \qquad (9.3)$$

$\longleftrightarrow$ $x_1(k),\dots,x_n(k)$ есть ФСР для однородного уравнения R_0, а

$x_{чн}(k)$ есть какое-либо частное решение неоднородного уравнения R_1.

Доказательство. *Необходимость*. Пусть общее решение неоднородного уравнения (R_1) задается формулой (9.3). Тогда при $C_1 = \ldots = C_n = 0$ функция $x_{чн}(k)$ есть частное решение уравнения R_1. Положим в (9.3) $C_1=1$, $C_2 = \ldots = C_n = 0$. Тогда $x_1(k) + x_{чн}(k)$, как и $x_{чн}(k)$, есть решение неоднородного уравнения R_1, и тогда их разность $x_1(k)$ есть решение однородного уравнения R_0. Аналогично можно показать, что $x_2(k),\ldots,x_n(k)$ есть решения однородного ЛРУ R_0. Покажем, что $x_1(k),\ldots,x_n(k)$ есть ФСР для R_0. Пусть $y(k)$ есть некоторое частное решение для R_2. Тогда для некоторых $C_1^0,\ldots,C_n^0$ из $\mathbb{R}$ $y(k) = C_1^0x_1(k) +\ldots+ C_n^0x_n(k) + x_{чн}(k)$ $\forall k$. Поэтому неоднородная система линейных уранений

$$C_1x_1(k) +\ldots+ C_nx_n(k) = y(k) - x_{чн}(k),\ k=0,1,\ldots,n-1,$$

имеет решение $C_1^0,\ldots,C_n^0$ и потому ее определитель (Казаратиан) $D \neq 0$. Следовательно система решений $x_1(k),\ldots,x_n(k)$ линейно независима и потому она является ФСР.

Достаточность. Пусть $x_1(k),\ldots,x_n(k)$ есть ФСР для однородного уравнения R_0, а $x_{чн}(k)$ есть какое-либо частное решение неоднородного уравнения R_1. Покажем, что общее решение неоднородного уравнения R_1 выражается формулой (9.3).

1. При любых $C_1,\ldots,C_n$ из $\mathbb{R}$ функция (9.3) есть решение для R_1. Покажем, что для всякого решения $y(k)$ для R_1 найдутся константы $c_1,\ldots,c_n$ из $\mathbb{R}$, для которых $y(k)=c_1x_1(k)+\ldots+c_nx_n(k)+x_{чн}(k)$. Рассмотрим линейную систему

$$C_1x_1(k) +\ldots+ C_nx_n(k) = y(k) - x_{чн}(k),\ k=0,1,\ldots,n-1. \quad (9.4)$$

Ее определитель есть Казаратиан $D\neq0$ и потому система (9.4) имеет единственное решение $c_1,\ldots,c_n$. Тогда

$$y(k) = c_1x_1(k) +\ldots+ c_nx_n(k) + x_{чн}(k),\ k=0,1,\ldots,n-1.$$

Это равенство справедливо для всякого k. Теорема доказана.

9.3.1. Метод Лагранжа вариации произвольных постоянных вычисления частного решения неоднородного уравнения

Напомним, что частное решение $y(t_k)$ неоднородного уравнения (R_1) называется главным, если $y(t_k)=0$, $k=0,1,\ldots,n-1$.

Теорема (Лагранжа). Если $x_1(k),\ldots,x_n(k)$ есть ФСР для нормированного однородного ЛРУ R_0'

$$L(x(k)) = x(k+n)+a_{n-1}(k)x(k+n-1)+\ldots+a_0(k)x(k) = 0, \quad (R_0')$$

то главное частное решение $y(k)$ нормированного неоднородного ЛРУ R_1'

$$L(x(k)) = x(k+n)+a_{n-1}(k)x(k+n-1)+\ldots+a_0(k)x(k) = f(k) \quad (R_1')$$

может быть найдено с помощью формул:

$$y(0) = 0,$$

$$y(k) = \sum_{i=0}^{k-1} \frac{D(x_1,\ldots,x_n \mid i+1,\ldots,i+n-1,k)}{D(x_1,\ldots,x_n \mid i+1,\ldots,i+n-1,i+n)} f(k),\ k>0, \quad (9.5)$$

если выражение (9.5) имеет смысл.

Доказательство. Пусть $x_1(k),\ldots,x_n(k)$ – фундаментальная система решений уравнения R_0'. Общее решение однородного уравнения R_0' $y_{oo} = C_1x_1(k) +\ldots+ C_nx_n(k)$, где произвольные постоянные $C_1,\ldots,C_n$ пробегают $\mathbb{R}$. Будем искать главное частное решение $y(k)$ в виде суммы

$$y(k) = \sum_{i=1}^{n} c_i(k)x_i(k), \quad (9.6)$$

где $c_i(k)$ – пока неизвестные дискретные функции вида $\mathbb{N}\to\mathbb{R}$. Подставим (9.6) в неоднородное уравнение R_1':

$$y(k+n) + \sum_{j=1}^{n} a_{n-j}(k)y(k+n-j) = f(k). \text{ Получим}$$

$$\underbrace{\sum_{i=1}^{n} c_i(k+n)x_i(k+n)}_{y(k+n)} + \sum_{j=1}^{n} a_{n-j}(k) \underbrace{\sum_{i=1}^{n} c_i(k+n-j)x_i(k+n-j)}_{y(k+n-j)} = f(k). \quad (9.7)$$

Будем искать функции $c_i(k)$, удовлетворяющими $\forall k\in\mathbb{N}$ следующим соотношениям:

$$\sum_{i=1}^{n} c_i(k+1)x_i(k+1) = \sum_{i=1}^{n} c_i(k)x_i(k+1),$$

$$\sum_{i=1}^{n} c_i(k+2)x_i(k+2) = \sum_{i=1}^{n} c_i(k)x_i(k+2), \qquad (9.8)$$

$$\dots$$

$$\sum_{i=1}^{n} c_i(k+n-1)x_i(k+n-1) = \sum_{i=1}^{n} c_i(k)x_i(k+n-1),$$

то есть соотношениям

$$\sum_{i=1}^{n} c_i(k+n-j)x_i(k+n-j) = \sum_{i=1}^{n} c_i(k)x_i(k+n-j), \quad j=1,\dots,n. \qquad (9.8')$$

Так как последнее соотношение в (9.7) справедливо $\forall k \in \mathbb{N}$, то возьмем его с на единицу большим аргументом:

$$\sum_{i=1}^{n} c_i(k+n)x_i(k+n) = \sum_{i=1}^{n} c_i(k+1)x_i(k+n),$$

а потом отнимем и добавим к равенству равные слагаемые. Получим

$$\sum_{i=1}^{n} c_i(k+n)x_i(k+n) = \sum_{i=1}^{n} c_i(k+1)x_i(k+n) - \sum_{i=1}^{n} c_i(k)x_i(k+n) + \sum_{i=1}^{n} c_i(k)x_i(k+n). \qquad (9.9)$$

Объединим в (9.9) справа два первых слагаемых:

$$\sum_{i=1}^{n} c_i(k+n)x_i(k+n) = \sum_{i=1}^{n} \Delta^1 c_i(k,k+1)x_i(k+n) +$$

$$\sum_{i=1}^{n} c_i(k)x_i(k+n). \qquad (9.9')$$

Подставим правую часть в (9.9′) вместо первой суммы в (9.7), подставим правую часть в (9.8′) вместо третьей суммы в (9.7) и получим

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1)x_i(k+n) + \sum_{i=1}^{n} c_i(k)x_i(k+n) + \sum_{j=1}^{n} a_{n-j}(k) \sum_{i=1}^{n} c_i(k)x_i(k+n-j) = f(k). \qquad (9.10)$$

Объединим в (9.10) вторую и третью суммы и получим

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1)x_i(k+n) + \sum_{i=1}^{n} c_i(k) \cdot \left[x_i(k+n) + \sum_{j=1}^{n} a_{n-j}(k)x_i(k+n-j) \right] = f(k). \qquad (9.10')$$

Содержимое квадратных скобок равно нулю, ибо в однородное уравнение R_0' подставлено его решение $x_i(k)$. Тогда (9.10′) начинает следующую систему линейных уравнений (5 формул).

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1)x_i(k+n) = f(k).$$

Объединим суммы первого уравнения в (9.8) и получим

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1)x_i(k+1) = 0.$$

В уравнении 1 из (9.8) заменим k на $k+1$ и из результата вычтем уравнение 2 из (9.8):

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1)x_i(k+2) = 0.$$

В уравнении 2 из (9.8) заменим k на $k+1$ и из результата вычтем уравнение 3 из (9.8):

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1) x_i(k+3) = 0.$$

И так далее. Наконец, в уравнении $n-2$ из (9.8) заменим k на $k+1$ и из результата вычтем уравнение $n-1$ из (9.8):

$$\sum_{i=1}^{n} \Delta^1 c_i(k,k+1) x_i(k+n-1) = 0.$$

В результате получаем систему из n линейных уравнений с n неизвестными $\Delta^1 c_1(k,k+1),\ldots,\Delta^1 c_n(k,k+1)$:

$$\begin{aligned}
&\sum_{i=1}^{n} x_i(k+1)\cdot \Delta^1 c_i(k,k+1) = 0,\\
&\sum_{i=1}^{n} x_i(k+2)\cdot \Delta^1 c_i(k,k+1) = 0,\\
&\sum_{i=1}^{n} x_i(k+3)\cdot \Delta^1 c_i(k,k+1) = 0, \qquad (9.11)\\
&\ldots\\
&\sum_{i=1}^{n} x_i(k+n-1)\cdot \Delta^1 c_i(k,k+1) = 0.\\
&\sum_{i=1}^{n} x_i(k+n)\cdot \Delta^1 c_i(k,k+1) = f(k),
\end{aligned}$$

Определитель получившейся системы (9.11) есть Казаратиан

$$D(x_1,..,x_n\,|\,k+1,\ldots,k+n) =$$

$$\begin{vmatrix} x_1(k+1) & x_2(k+1) & \ldots & x_n(k+1)\\ & & \ldots & \\ x_1(k+n-1) & x_2(k+n-1) & \ldots & x_n(k+n-1)\\ x_1(k+n) & x_2(k+n) & \ldots & x_n(k+n) \end{vmatrix}.$$

Казаратиан D отличен от нуля, иначе функции фундаментальной системы решений $x_1,\ldots,x_n$ линейно зависимы. Решим систему (9.11) и тогда

$$\Delta^1 c_j(k,k+1) = \frac{\begin{vmatrix} x_1(k+1) & \ldots & 0 & \ldots & x_n(k+1)\\ & \ldots & & \ldots & \\ x_1(k+n-1) & \ldots & 0 & \ldots & x_n(k+n-1)\\ x_1(k+n) & \ldots & f(k) & \ldots & x_n(k+n) \end{vmatrix}}{D(x_1,..,x_n\,|\,k+1,\ldots,k+n)},$$

$j = 1,2,\ldots,n.$

Разложим определитель по столбцу j:

$$\Delta^1 c_j(k,k+1) = c_j(k+1) - c_j(k) =$$

$$(-1)^{n+j} f(k) \frac{D_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n\,|\,k+1,\ldots,k+n-1)}{D(x_1,..,x_n\,|\,k+1,\ldots,k+n)} \qquad (9.12)$$

Числитель (9.12) есть Казаратиан размерности $n-1$. Из (9.12)

$$c_j(k+1) = c_j(k) + (-1)^{n+j}\frac{D_j}{D} f(k).$$

Разрешая это уравнение рекурсивным образом, получаем

$$c_j(k+1) =$$

$$\underbrace{c_j(k-1) + (-1)^{n+j}\frac{D_j}{D} f(k-1)}_{c_j(k)} + (-1)^{n+j}\frac{D_j}{D} f(k) = \ldots =$$

$$c_j(0) + \sum_{i=0}^{k} (-1)^{n+j}\frac{D_j}{D} f(i), \quad \text{откуда } c_j(k+1) =$$

$$c_j(0) + \sum_{i=0}^{k} (-1)^{n+j}\frac{D_j(x_1,..,x_n\,|\,k+1,\ldots,k+n-1)}{D(x_1,..,x_n\,|\,k+1,\ldots,k+n)} f(t_i) \qquad (9.13)$$

Подставляем в (9.6) найденные в (9.13) $c_j(k)$ при $k \geqslant 1$, затем разделим сумму на два слагаемых и во втором слагаемом поменяем порядок суммирования. Тогда

$$y(k) = \sum_{j=1}^{n} c_j(k)\, x_j(k) = \tag{9.14}$$

$$\sum_{j=1}^{n} \left\{ c_j(0) + \sum_{i=0}^{k-1} (-1)^{n+j} \frac{D_j(x_1,..,x_n \mid i+1,\ldots,i+n-1)}{D(x_1,..,x_n \mid i+1,\ldots,i+n)} f(i) \right\} \cdot x_j(k) =$$

$$\sum_{j=1}^{n} c_j(0) x_j(k) +$$

$$\sum_{i=0}^{k-1} \left\{ \sum_{j=1}^{n} (-1)^{n+j} \frac{D_j(x_1,..,x_n \mid i+1,\ldots,i+n-1)}{D(x_1,..,x_n \mid i+1,\ldots,i+n)} x_j(k) \right\} f(i).$$

Вынесем не зависящий от j определитель $D(x_1,..,x_n \mid i+1,\ldots, i+n)$ в знаменателе за знак суммы по j:

$$y(k) = \sum_{j=1}^{n} c_j(0) x_j(k) +$$

$$\sum_{i=0}^{k-1} \left\{ \frac{\sum_{j=1}^{n} (-1)^{n+j} \cdot D_j(x_1,..,x_n \mid i+1,\ldots,i+n-1) \cdot x_j(k)}{D(x_1,..,x_n \mid i+1,\ldots,i+n)} \right\} f(i).$$

Внутренняя сумма есть разложение определителя $D(x_1,..,x_{j-1}, x_j, x_{j+1},\ldots,x_n \mid i+1,\ldots,i+n-1,k)$ порядка n по последней строке $x_1(k),\ldots,x_n(k)$. Тогда

$$y(k) = \underbrace{\sum_{j=1}^{n} c_j(0) x_j(k)}_{\text{слагаемое 1}} + \underbrace{\sum_{i=0}^{k-1} \frac{D(x_1,..,x_n \mid i+1,\ldots,i+n-1,k)}{D(x_1,..,x_n \mid i+1,\ldots,i+n)} f(i)}_{\text{слагаемое 2}}.$$

По определению главного частного решения $y(0) = y(1) = \ldots =$

$y(n-1) = 0$. При $k=1,2,\ldots,n-1$ слагаемое 2 обращается в ноль, ибо детерминант числителя имеет две одинаковые строки. При $k=0$ потребуем, чтобы $\sum_{j=1}^{n} c_j(0) x_j(k) = 0$. Тогда для главного частного решения $y(k)$ получаем систему уравнений:

$$y(k) = \sum_{j=1}^{n} c_j(0) x_j(k) = 0, \quad k=0,1,\ldots,n-1. \tag{9.15}$$

Определитель однородной системы (9.15) есть не равный нулю Кааратиан, поэтому система имеет только нулевое решение

$$c_j(0) = 0, \quad j=1,\ldots,n.$$

Тогда слагаемое 1 в (9.14) обращается в ноль, главное решение $y(k)$ есть слагаемое 2 и

$$y(0) = 0,$$

$$y(k) = \sum_{i=0}^{k-1} \frac{D(x_1,\ldots,x_n \mid i+1,\ldots,i+n-1,k)}{D(x_1,\ldots,x_n \mid i+1,\ldots,i+n-1,i+n)} f(k), \quad k>0.$$

Теорема доказана.

Следствие. Если $x_1(k),\ldots,x_n(k)$ есть фундаментальная система решений однородного уравнения (R_0), то для общего решения неоднородного уравнения (R_1) справедлива формула

$$y_{\text{он}}(k) = y(k) + C_1 x_1(k) + \ldots + C_n x_n(k).$$

Если $x_1(k),\ldots,x_n(k)$ есть нормированная фундаментальная система решений для (2), то

$$y(k) = y(0)x_1(k) + \ldots + y(n-1)x_n(k) +$$

$$\begin{cases} 0, \text{ если } k=0, \\ \sum_{i=0}^{k-1} \frac{D(x_1,..,x_n \mid i+1,\ldots,i+n-1,k)}{D(x_1,..,x_n \mid i+1,\ldots,i+n)} f(i), \text{ если } k>0, \end{cases}$$

если сумма при $k>0$ имеет смысл.

Замечание. 1. Формулы данной теоремы имеют смысл, если в уравнениях $(R_0),(R_1)$ коэффициент $a_0(k)$ нигде не обращается в ноль.

2. Из формулы (9.14) следует, что любое частное решение есть сумма двух слагаемых в (9.14).

3. Если коэффициенты $a_i(k)$ нигде не обращаются в нуль, то выражение для $y(k)$ существует.

4. Получить явно заданные (без помощи рекурсии) ФСР для однородного ЛРУ R_0 и какое-либо частное решение для неоднородного ЛРУ R_1 есть часто непростая задача.

9.4. Линейные рекуррентные уравнения с постоянными коэффициентами

Определение. Уравнения R_0 и R_1

$$L(x(k))=a_n x(k+n)+a_{n-1}x(k+n-1)+\ldots+a_0 x(k)=0, \qquad (R_0)$$

$$L(x(k))=a_n x(k+n)+a_{n-1}x(k+n-1)+\ldots+a_0 x(k)=f(k), \qquad (R_1)$$

где $a_i\in\mathbb{R}$, $i=0,1,\ldots,n$; $a_n\neq 0$; $f(k)\not\equiv 0$ – известная функция, $x(k)$ – неизвестная функция, называются *линейными рекуррентными уравнениями* (ЛРУ) *порядка n*, *однородным и неоднородным* соответственно *с постоянными коэффициентами*.

Уравнения R_0 и R_1 называют также *стационарными ЛРУ* (СЛРУ) однородным и неоднородным соответственно.

Определение. Выражение

$$L(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} +\ldots+ a_1\lambda + a_0 \qquad (9.16)$$

называется *характеристическим полиномом*, а выражение $L(\lambda)=0$ *характеристическим уравнением* для однородного СЛРУ R_0 (равно как и для неоднородного СЛРУ R_1).

Теорема. Если

$\lambda_1,\ldots,\lambda_p$ –вещественные корни характеристического уравнения,
$l_1,\ldots,l_p$ – их кратности,
$\mu_1,\ldots,\mu_s$ – группа комплексных корней $\mu_j=\alpha_j+i\beta_j, j=1,2,\ldots,s$,
$\bar{\mu}_1,\ldots,\bar{\mu}_s$ – группа комплексных корней $\bar{\mu}_j=\alpha_j-i\beta_j, j=1,2,\ldots,s$,
$r_1,\ldots,r_s$ – их кратности,
ρ_j,φ_j – модуль и аргумент комплексного числа $\mu_j, j=1,\ldots,s$,

то функции

$$x_j(k) = k^{m_j}\lambda_j^k,\ j=1,\ldots,p;\ m_j=0,\ldots,l_j-1,$$
$$y_j(k) = k^{m_j}\rho_j^k\cos(\varphi_j k),\ j=1,\ldots,s;\ m_j=0,\ldots,r_j-1,$$
$$z_j(k) = k^{m_j}\rho_j^k\sin(\varphi_j k),\ j=1,\ldots,s;\ m_j=0,\ldots,r_j-1,$$

составляют ФСР для однородного СЛРУ R_0.

Доказательство. Пусть

$$\varphi(k) = u(k)\lambda^k, \qquad (9.17)$$

где $u(k)$ есть пока не определенная функция, а λ комплексный параметр. Тогда

$$\begin{cases} \varphi(k+1) = u(k+1)\lambda^{k+1} = (u(k)+\Delta^1 u(k))\lambda^{k+1}, \\ \varphi(k+2) = u(k+2)\lambda^{k+2} = (u(k)+2\Delta^1 u(k)+\Delta^2 u(k))\lambda^{k+2}, \\ \quad\ldots \\ \varphi(k+n) = u(k+n)\lambda^{k+n} = (u(k) + C_n^1\Delta^1 u(k) +\ldots+ \\ \qquad\qquad C_n^n\Delta^n u(k))\lambda^{k+n}. \end{cases} \qquad (9.18)$$

Подставим функцию φ из (9.18) в (R_0) вместо x и вынесем λ^k за скобки.

$$\lambda^k\Big[a_n\cdot\Big(u(k) + C_n^1\Delta^1 u(k) +\ldots+ C_n^n\Delta^n u(k)\Big)\lambda^n +$$
$$a_{n-1}\cdot\Big(u(k) + C_{n-1}^1\Delta^1 u(k) +\ldots+ C_{n-1}^n\Delta^{n-1} u(k)\Big)\lambda^{n-1} +$$
$$\ldots$$
$$a_1\cdot\Big(u(k) + C_1^1\Delta^1 u(k)\Big)\lambda +$$
$$a_0\cdot C_0^0\Delta^1 u(k)\Big]\lambda^0 = 0.$$

Перегруппируем слагаемые.

$$\lambda^k\Big[\Big(a_n\lambda^n + a_{n-1}\lambda^{n-1} +\ldots+ a_0\Big)u(k) +$$
$$\lambda\Big(a_n n\lambda^{n-1} + a_{n-1}(n-1)\lambda^{n-2} +\ldots+ a_1\Big)\Delta^1 u(k) +$$
$$\frac{\lambda^2}{2!}\Big(a_n n(n-1)\lambda^{n-2} + a_{n-1}(n-1)(n-2)\lambda^{n-3} +\ldots+ a_2\Big)\Delta^2 u(k) +$$
$$\ldots$$
$$\frac{\lambda^{n-1}}{(n-1)!}\Big(a_n n(n-1)\ldots 3\cdot 2\cdot\lambda + a_{n-1}(n-1)!\Big)\Delta^{n-1}u(k) +$$
$$\frac{\lambda^n}{n!}\Big(a_n n!\Big)\Delta^n u(k)\Big] =$$
$$\lambda^k\Big(L(\lambda)u(k) + \frac{\lambda}{1!}L'(\lambda)\Delta^1 u(k) +\ldots+ \frac{\lambda^n}{n!}L^{(n)}(\lambda)\Delta^n u(k)\Big) =$$
$$L\Big(u(k)\cdot\lambda^k\Big) = 0.$$

Пусть λ есть корень характеристического уравнения (9.16) кратности $q \geqslant 1$. Тогда

$$L(\lambda) = L'(\lambda) = \ldots = L^{(q-1)}(\lambda) = 0,\ L^{(q)}(\lambda) \neq 0 \text{ и}$$

$$\frac{\lambda^q}{q!}L^{(q)}(\lambda)\Delta^q u(k) + \ldots + \frac{\lambda^n}{n!}L^{(n)}(\lambda)\Delta^n u(k) = 0. \qquad (9.19)$$

Из уравнения (9.19) находим $u(k)$ и решение φ в виде (9.17). Решениями уравнения (9.19) будут, в частности, функции

$$u(k)=1,\ u(k)=k,\ u(k)=k^2,\ \ldots,\ u(k)=k^{q-1}.$$

Тогда имеем q решений, отвечающих корню λ кратности q: λ^k, $k\lambda^k$, $k^2\lambda^k, \ldots, k^{q-1}\lambda^k$. Если $\lambda = \rho e^{i\varphi} = \rho(\cos\varphi + i\sin\varphi)$ есть комплексный корень кратности q, то вещественная и мнимая части в $k^s\lambda^k$, $s=0,\ldots,q-1$, есть вещественные решения

$$\rho^k \cos k\varphi,\ k\rho^k \cos k\varphi,\ k^2\rho^k \cos k\varphi, \ldots,\ k^{q-1}\rho^k \cos k\varphi,$$
$$\rho^k \sin k\varphi,\ k\rho^k \sin k\varphi,\ k^2\rho^k \sin k\varphi, \ldots,\ k^{q-1}\rho^k \sin k\varphi.$$

Все функции линейно независимы, так как у них разная асимптотика роста ρ^k. Их линейную независимость можно показать также с помощью определителя Казарати при $k=0,1,\ldots,n-1$.

Теорема доказана.

Замечание. 1. Если $x_1(k),\ldots,x_n(k)$ есть ФСР для однородного СЛРУ R_0, то его общее решение

$$x_{oo}(k) = C_1x_1(k) + \ldots + C_nx_n(k),$$

где произвольные постоянные $C_1,\ldots,C_n$ пробегают $\mathbb{R}$ независимо друг от друга. Если $x_{чн}$ есть какое-либо частное решение неоднородного СЛРУ R_1, то его общее решение

$$x_{он}(k) = x_{чн}(k) + C_1x_1(k) + \ldots + C_nx_n(k).$$

2. Поиски (без помощи рекурсии) частного решения неоднородного уравнения есть часто непростая задача. Частное решение неоднородного СЛРУ R_1 с правой частью – квазиполиномом $f(k) = P_m(k)\cdot\lambda^k$ может быть найдено в виде $k^rQ_m(k)\lambda^k$, где r есть кратность корня λ характеристического уравнения.

3. Уравнение

$$a_nx(k+n) + a_{n-1}x(k+n-1) + \ldots + a_{n-r}x(k+n-r) = f(k),$$
$$k = 0,1,2,\ldots$$

допускает понижение порядка. Это уравнение имеет характеристическое уравнение

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots + a_{n-r}\lambda^{n-r} =$$

$$\lambda^{n-r}(a_n\lambda^r + a_{n-1}\lambda^{r-1} + \ldots + a_{n-r+1}\lambda + a_{n-r}) = 0$$

с корнем $\lambda=0$ кратности $n-r$. Решение исходного уравнения совпадает с решением уравнения

$$a_ny(k+r) + a_{n-1}y(k+r-1) + \ldots + a_{n-r}y(k) = f(k-(n-r)),$$
$$k = n-r, n-r+1, \ldots$$

порядка r с характеристическим уравнением

$$a_n\lambda^r + a_{n-1}\lambda^{r-1} + \ldots + a_{n-r+1}\lambda + a_{n-r} = 0$$

без нулевых корней.

В самом деле, из первого уравнения

$$x(k+n) = (f(k) - a_{n-1}x(k+n-1) - \ldots - a_{n-r}x(k+n-r))/a_n,$$
$$k = 0,1,2,\ldots$$

Из второго уравнения

$$y(k+r) = (f(k-(n-r)) - a_{n-1}y(k+r-1) - \ldots - a_{n-r}y(k))/a_n,$$
$$k = n-r, n-r+1, \ldots$$

При указанных значениях k обе последовательности одинаковы. Вместо второго уравнения удобней решить уравнение

$$z(k+r) = (f(k) - a_{n-1}z(k+r-1) - \ldots - a_{n-r}z(k))/a_n,$$
$$k = 0,1,2,\ldots$$

и тогда $y(k+r) = z(k)$, $k = 0,1,2,\ldots$

Примеры. 1. $x(k+2) - 4x(k+1) + 3x(k)=0$,

$\lambda^2 - 4\lambda + 3 = 0$, $\lambda_1=1, \lambda_2=3$, $x_{oo} = C_1(\lambda_1)^k + C_2(\lambda_2)^k = C_11^k + C_23^k = C_1 + C_23^k$, $C_1, C_2 \in \mathbb{R}$.

2. $x(k+2) - 3x(k) = 0$, $\lambda^2 - 3=0$, $\lambda_1=\sqrt{3}$, $\lambda_2=-\sqrt{3}$,

$x_{oo} = C_1(\lambda_1)^k + C_2(\lambda_2)^k = C_1(\sqrt{3})^k + C_2(-\sqrt{3})^k$, $C_1, C_2 \in \mathbb{R}$.

3. $x(k+2) - x(k+1) - x(k) = 0$,

$$\lambda^2 - \lambda - 1 = 0,\ \lambda_1 = \frac{1+\sqrt{5}}{2},\ \lambda_2 = \frac{1-\sqrt{5}}{2},$$

$$x_{oo} = C_1\left(\frac{1+\sqrt{5}}{2}\right)^k + C_2\left(\frac{1-\sqrt{5}}{2}\right)^k,\ C_1, C_2 \in \mathbb{R}.$$

4. $x(k+2) + 2x(k+1) + x(k) = 0$,

$\lambda^2 + 2\lambda + 1 = 0$, $\lambda = -1$ кратность 2,

$x_{oo} = C_1(-\lambda)^k + C_2k(-\lambda)^k = C_1(-1)^k + C_2k(-1)^k =$
$(-1)^k(C_1 + C_2k)$, $C_1, C_2 \in \mathbb{R}$.

5. $x(k+3) + 10x(k+2) + 32x(k+1) + 32x(k) = 0$,
$\lambda^3 + 10\lambda^2 + 32\lambda + 32 = 0$,
$(\lambda+4)^2(\lambda+2) = 0$, $\lambda_1=-4$ кратности 2, $\lambda_2=-2$,

$x_{oo} = C_1(\lambda_1)^k + C_2k(\lambda_1)^k + C_3(\lambda_3)^k =$
$C_1(-2)^k + C_2k(-2)^k + C_3(-2)^k$, $C_1, C_2, C_3 \in \mathbb{R}$.

6. $x(k+3) + 3x(k+2) + 3x(k+1) + x(k) = 0$,
$\lambda^3 + 3\lambda^2 + 3\lambda + 1 = 0$, $(\lambda+1)^3=0$, $\lambda_1=-1$ кратности 3,
$x_{oo} = C_1(\lambda)^k + C_2k(\lambda)^k + C_3k^2(\lambda)^k =$
$C_1(-1)^k + C_2k(-1)^k + C_3k^2(-1)^k = (-1)^k(C_1 + kC_2 + k^2C_3)$,
$C_1, C_2, C_3 \in \mathbb{R}$.

7. $x(k+2) - 4x(k+1) + 3x(k) = 0$, $x(0)=10$, $x(1)=16$.
$\lambda^2 - 4\lambda + 3 = 0$, $\lambda_1=1$, $\lambda_2=3$,
$x_{oo} = C_1 \cdot 1^k + C_2 \cdot 3^k = C_1 + C_2 3^k$,

$$\begin{cases} x(0) = C_1+C_2 = 10, \\ x(1) = C_1+C_2 3 = 16, \end{cases} \quad \begin{cases} C_1=7, \\ C_2=3, \end{cases} \quad x(k) = 7 + 3 \cdot 3^k = 7 + 3^{k+1}.$$

8. $64x(k+8) + 48x(k+6) + 12x(k+4) + x(k+2) = 0$,

$64\lambda^8 + 48\lambda^6 + 12\lambda^4 + \lambda^2 = 0$, $\lambda^2((2\lambda)^2+1)^2(4\lambda^2+2\lambda+1) = 0$,
$\lambda=0$, кратность 2,

$$\lambda = \frac{1}{2}i = \frac{1}{2}e^{(\pi/2)i} = \frac{1}{2}\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right), \text{ кратность 2, } \rho=\frac{1}{2}, \ \varphi=\frac{\pi}{2},$$

$$\bar{\lambda} = -\frac{1}{2}i, \text{ комплексно сопряженный корень, кратность 2,}$$

$$\lambda = -\frac{1}{4} + i\frac{\sqrt{3}}{4}, \text{ кратность 1, } \rho = \sqrt{\frac{1}{16} + \frac{3}{16}} = \frac{1}{2},$$

$$\operatorname{tg}\varphi = \frac{y}{x} = \frac{\sqrt{3}/4}{-1/4} = -\sqrt{3}, \ \varphi = \frac{\pi}{2} + \frac{\pi}{6} = \frac{2\pi}{3},$$

$$\lambda = -\frac{1}{4} - i\frac{\sqrt{3}}{4}, \text{ комплексно сопряженный корень, кратность 1,}$$

$$x_{oo} = C_1 \cdot 0^k + C_2k \cdot 0^k +$$
$$C_3\left(\frac{1}{2}\right)^k \cos\left(\frac{\pi}{2}k\right) + C_4\left(\frac{1}{2}\right)^k \sin\left(\frac{\pi}{2}k\right) +$$

$$C_5k\left(\frac{1}{2}\right)^k \cos\left(\frac{\pi}{2}k\right) + C_6k\left(\frac{1}{2}\right)^k \sin\left(\frac{\pi}{2}k\right) +$$
$$C_7\left(\frac{1}{2}\right)^k \cos\left(\frac{2\pi}{3}k\right) + C_8\left(\frac{1}{2}\right)^k \sin\left(\frac{2\pi}{3}k\right).$$

9. $x(k+1) - x(k) = k+1$, $x(0)=1$;
$\lambda - 1 = 0$, $\lambda=1$, $x_{oo} = C \cdot 1^k = C$,
$x_{oн} = x_{чн} + x_{oo} = x_{чн} + C$, $f(k) = k+1 = 1^k(k+1)$,
$x_{чн}(k) = k(ak+b) \cdot 1^k$, $x_{чн}(k+1) = (k+1)(a(k+1)+b)$.
Подставляем $x_{чн}(k)$, $x_{чн}(k+1)$ в исходное уравнение:
$(k+1)(a(k+1)+b) - k(ak+b) = k+1$, $2ak + (a+b) = k+1$.
Приравниваем коэффициенты при одинаковых степенях k:

$2a=1$, $a+b=1$; $a=1/2$, $b=1/2$;

$$x_{чн}(k) = \frac{k(k+1)}{2}; \quad x_{oн}(k) = x_{чн}(k) + C = \frac{k(k+1)}{2} + C,$$

$$1 = x(0) = \frac{0(0+1)}{2} + C = C, \ C=1; \ x(k) = \frac{k(k+1)}{2} + 1.$$

Ответ. $x(k) = \dfrac{k(k+1)}{2} + 1$.

10. $x(k+5) - 6x(k+4) + 9x(k+3) = 3k-1$, $k=0,1,2,\ldots$
$x(0)=1$, $x(1)=0$, $x(2)=1$, $x(3)=3$, $x(4)=0$.

Характеристическое уравнение
$\lambda^5 - 6\lambda^4 + 9\lambda^3 = 0$, $\lambda^3(\lambda^2 - 6\lambda + 9) = 0$.
Корни: $\lambda=0$ кратности 3, $\lambda=3$ кратности 2.
Переходим к эквивалентному уравнению

$y(k+2) - 6y(k+1) + 9y(k) = 3(k-3) - 1 = 3k-10$, $k=3,4,\ldots$,
$y(0)=1$, $y(1)=0$, $y(2)=1$, $y(3)=3$, $y(4)=0$.

Решение $y(k)$ этого уравнения, начиная с $k=3$, совпадет с решением $z(k)$ уравнения

$z(k+2) - 6z(k+1) + 9z(k) = 3k-1$, $k=0,1,2,\ldots$,
$z(0)=3$, $z(1)=0$,

в том смысле, что решение $y(k+3) = z(k)$, $k=0,1,2,\ldots$
Характеристическое уравнение $\lambda^2 - 6\lambda + 9 = 0$.
Корень $\lambda=3$ кратности 2.

$$z_{oн}(k) = z_{чн}(k) + z_{oo}(k) = z_{чн}(k) + C_1 \cdot 3^k + C_2k \cdot 3^k.$$

$z_{чн}(k) = z(k) = ak+b$, $z(k+1) = a(k+1) + b = ak+a+b$,
$z(k+2) = a(k+2) + b = ak+2a+b$.

Подставляем $z(k), z(k+1), z(k+2)$ в уравнение

$z(k+2) - 6z(k+1) + 9z(k) = 3k-1$ и получаем:

$$ak+2a+b - 6(ak+a+b) + 9(ak+b) = 3k-1,$$

$$4ak-4a+4b = 3k-1, \quad \begin{cases} 4a=3, \\ -4a+4b=-1, \end{cases} \quad \begin{cases} a=3/4, \\ b=2/4. \end{cases} \quad z_{чн}(k) = \frac{3}{4}k+\frac{2}{4},$$

$$z_{он}(k) = \frac{3k+2}{4} + C_1 \cdot 3^k + C_2 k \cdot 3^k.$$ Находим C_1, C_2.

$$\begin{cases} z_{он}(0) = \dfrac{3\cdot 0+2}{4} + C_1 \cdot 3^0 + C_2 \cdot 0 \cdot 3^0 = \dfrac{2}{4} + C_1 + C_2 \cdot 0 \cdot 3^0 = 3, \\ z_{он}(1) = \dfrac{3\cdot 1+2}{4} + C_1 \cdot 3^1 + C_2 \cdot 1 \cdot 3^1 = \dfrac{5}{4} + 3C_1 + 3C_2 = 0, \end{cases}$$

$$\begin{cases} C_1 = \dfrac{10}{4}, \\ 3C_1+3C_2 = -\dfrac{5}{4}, \end{cases} \quad 3C_2 = -\frac{5}{4} - 3C_1 = -\frac{5}{4} - \frac{30}{4} = -\frac{35}{4},$$

$$C_1 = \frac{10}{4}, \; C_2 = -\frac{35}{4\cdot 3},$$

$$z(k) = \frac{3k+2}{4} + \frac{10}{4}\cdot 3^k - \frac{35}{4\cdot 3}k \cdot 3^k, \; k=0,1,2,\ldots$$

Ответ. $x(0)=1$, $x(1)=0$, $x(2)=1$,

$$x(k) = \frac{3k+2}{4} + \frac{10}{4}\cdot 3^k - \frac{35}{4\cdot 3}k \cdot 3^k, \; k=3,4,5,\ldots$$

Часть 2. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

10. ПОРОЖДЕНИЕ КОМБИНАТОРНЫХ КОНФИГУРАЦИЙ И ИХ ПЕРЕСЧЕТ

10.1. Размещения, перестановки, сочетания

Перестановка n-элементного множества M есть упорядоченный набор длины n, составленный из попарно различных элементов множества M. Обозначим через P_M множество всех перестановок из n элементов и через P_n число всех перестановок из n элементов. Например, если $M = \{a,b,c\}$, то $P_M = \{(a,b,c), (a,c,b), (b,a,c), (b,c,a), (c,a,b), (c,b,a)\}$; $P_n = 6$.

Сочетание из n элементов по r есть r-элементное подмножество n-элементного множества M.

Обозначим через C_M^r множество всех сочетаний из n элементов по r и через C_n^r (или через $\binom{n}{r}$) число всех сочетаний из n элементов по r. Например, если $M = \{a,b,c\}$, то

$C_M^1 = \{(a), (b), (c)\}$; $C_M^2 = \{(a,b), (a,c), (b,c)\}$;

$C_3^1 = |C_M^1| = 3$; $C_3^2 = |C_M^2| = 3$.

Размещение из n элементов по r есть упорядоченный набор, состоящий из r попарно различных элементов, взятых из n-элементного множества M.

Обозначим через A_M^r множество всех размещений из n элементов по r и через A_n^r число всех размещений из n элементов по r. Заметим, что для n-элементного множества M множество перестановок $P_M = A_M^n$, а их число $P_n = A_n^n$.

Пример. $M = \{a,b,c\}$; $A_M^1 = \{(a),(b),(c)\}$; $A_M^2 = \{(a,b), (a,c), (b,c), (b,a), (c,a), (c,b)\}$; $A_3^1 = |A_M^1| = 3$; $A_3^2 = |A_M^2| = 6$.

В размещениях, перестановках, сочетаниях элементов некоторого n-элементного множества могут допускаться повторы элементов. Будем называть их размещениями, перестановками, сочетаниями с повторениями. Обозначим через $\hat{A}_M^r$, $\hat{P}_M$, $\hat{C}_M^r$ – множества всех размещений, перестановок, сочетаний с повторениями, а через $\hat{A}_n^r$, $\hat{P}_n$, $\hat{C}_n^r$ – их число. Например, если $M =$

$\{a,b,c\}$, то

$\hat{C}_M^2 = \{(a,a),(b,b),(c,c),\ (a,b),(a,c),(b,c)\}$;

$\hat{C}_3^2 = |\hat{C}_M^2| = 6$; $\hat{A}_M^2 = \{(a,a),(b,b),(c,c),(a,b),(a,c),$

$(b,c),(b,a),\ (c,a),(c,b)\}$; $\hat{A}_3^2 = 9$.

Чтобы подчеркнуть, что размещение, перестановка, сочетание содержит r элементов, будем говорить r-размещение, r-перестановка, r-сочетание. Размещения, перестановки, сочетания, составленные из элементов некоторого множества M, называются комбинаторными конфигурациями. Всякая конфигурация $(a_1,a_2,\ldots,a_r)$ для множества M лежит в декартовом произведении $M\times M\times\ldots\times M$, состоящем из r сомножителей. Мощности множеств комбинаторных конфигураций называются комбинаторными числами.

10.2. Правило суммы и правило произведения

Правило суммы. Пусть конечное множество M разбито на два непересекающихся подмножества M_1 и M_2 (в объединении дающих все множество M). Тогда мощность $|M| = |M_1| + |M_2|$.

Правило произведения. Пусть в некотором множестве объект a может быть выбран n способами, и после этого (т.е. после выбора объекта a) объект b может быть выбран m способами. Тогда объект ab может быть выбран $n\cdot m$ способами.

Оба правила допускают индуктивное обобщение. Если конечное множество M допускает разбиение на r попарно непересекающихся подмножеств $M_1,M_2,\ldots,M_r$, то мощность $|M| = |M_1| + |M_2| +\ldots+ |M_r|$.

Если объект a_1 может быть выбран k_1 способами, затем (после выбора объекта a_1) объект a_2 может быть выбран k_2 способами и так далее и, наконец, объект a_r может быть выбран k_r способами, то объект $a_1a_2\ldots a_r$ может быть выбран $k_1\cdot k_2\cdot\ldots\cdot k_r$ способами.

10.3. Подсчет числа размещений, перестановок, сочетаний

10.3.1. Число размещений без повторений

Теорема. Число размещений без повторений из n элементов по r $A_n^r = n(n-1)(n-2)\ldots(n-r+1) = n!/(n-r)!$.

Доказательство. В r-размещении $(a_1,a_2,\ldots,a_r)$ n-элементного множества M элемент a_1 можно выбрать n способами. После

этого элемент a_2 можно выбрать $n-1$ способами (из оставшихся $n-1$ элементов множества M). Затем элемент a_3 можно выбрать $n-2$ способами. И так далее. Наконец, элемент a_r можно выбрать $n-r+1$ способами. По правилу произведения $A_n^r = n(n-1)\cdot(n-2)\ldots(n-r+1)$; умножив и разделив правую часть равенства на $(n-r)(n-r-1)\ldots3\cdot2\cdot1$, получим $A_n^r = n!/(n-r)!$.

Следствие. Число перестановок без повторений $P_n = n!$.

10.3.2. Число размещений с повторениями

Теорема. Число размещений с повторениями $\hat{A}_n^r = n^r$.

Доказательство. В r-размещении $(a_1,a_2,\ldots,a_r)$ элемент a_1 можно в n-элементном множестве M выбрать n способами, элемент a_2 – тоже n способами, наконец, элемент a_r – n способами. По правилу произведения $\hat{A}_n^r = n^r$.

Следствие. Число перестановок с повторениями $\hat{P}_n = n^n$.

10.3.3. Число сочетаний без повторений

Теорема. Число сочетаний без повторений $C_n^r = \dfrac{n!}{r!(n-r)!}$.

Доказательство. Каждому r–сочетанию $(a_1,a_2,\ldots,a_r)$ для n-элементного множества соответствует $r!$ перестановок. Тогда число размещений $A_n^r = C_n^r\cdot r!$, откуда и следует требуемая формула.

10.3.4. Число сочетаний с повторениями

Теорема. Число сочетаний с повторениями $\hat{C}_n^r = C_{n+r-1}^r$.

Доказательство. Каждому r-сочетанию из n-элементного множества M сопоставим набор $(k_1,k_2,\ldots,k_n)$ из натуральных чисел, указывающих число повторов каждого элемента из M в выбранном сочетании. При этом $k_1+k_2+\ldots+k_n = r$. Например, если $M = \{a,b,c,d,e\}$, то сочетанию (a,a,c,c,c,e,e) сопоставим набор $(2,0,3,0,2)$, т.е. элементы a,b,c,d,e множества M встречаются в сочетании (a,a,c,c,c,e,e) соответственно 2,0,3,0, 2 раз. Каждому полученному набору $(k_1,k_2,\ldots,k_n)$ сопоставим набор $(l_1,l_2,\ldots,l_n)$ с $l_i = k_i+1$, $i = 1,2,\ldots,n$. Тогда $l_1 + l_2+\ldots+l_n = k_1+k_2+\ldots+k_n+n = r+n$. Каждому полученному набору $(l_1,l_2,\ldots,l_n)$ взаимно однозначно соответствует разбиение числа $n+r$ на n ненулевых слагаемых $l_1,l_2,\ldots,l_n$. Разделим $n+r$ последовательно записанных звездочек вертикальными раз-

делительными черточками на n непустых частей, состоящих соответственно из $l_1, l_2, \ldots, l_n$ звездочек. Для нашего примера получим следующее разбиение:

* * * |	* |	* * * * |	* |	* * *
$l_1=3$	$l_2=1$	$l_3=4$	$l_4=1$	$l_5=3$
$k_1=2$	$k_2=0$	$k_3=3$	$k_4=0$	$k_5=2$

Каждому разбиению числа $n+r$ на n ненулевых слагаемых взаимно однозначно соответствует расстановка $n-1$ разделителей, которые можно расставить на $n+r-1$ пробелах между звездочками C_{n+r-1}^{n-1} способами. Следовательно, число сочетаний с повторениями $\hat{C}_n^r = C_{n+r-1}^{n-1} = C_{n+r-1}^r$.

10.3.5. Число перестановок данной спецификации

Поясним сначала смысл выражения: $P_n(k_1,k_2,\ldots,k_r)$, $k_1+k_2+\ldots+k_r = n$. Пусть имеем набор (к,к,ж,ж,ж,с) из шести шаров, из которых два красных, три желтых, один синий. Набор (2,3,1) называется спецификацией этого набора шаров. Возможна, например, перестановка (ж,с,к,ж,к,ж) шаров. Пусть $P_6(2,3,1)$ означает число всех перестановок спецификации (2,3,1). Пусть теперь имеем n элементов, из которых:

k_1 элементов вида 1;
k_2 элементов вида 2;
...
k_r элементов вида r;

причем все $k_i>0$ и $k_1+k_2+\ldots+k_r = n$. Пусть $P_n(k_1,k_2,\ldots,k_r)$ означает число всех перестановок спецификации $(k_1,k_2,\ldots,k_r)$.

Теорема. $P_n(k_1,k_2,\ldots,k_r) = n!/(k_1!k_2!\ldots k_r!)$.

Доказательство. k_1 элементов вида 1 можно разместить на n местах $C_n^{k_1}$ способами;

k_2 элементов вида 2 – на оставшихся $n-k_1$ местах $C_{n-k_1}^{k_2}$ способами;

k_3 элементов вида 3 – на оставшихся $n-k_1-k_2$ местах $C_{n-k_1-k_2}^{k_3}$ способами;

...

k_r элементов вида r – на оставшихся $n-k_1-k_2-\ldots-k_{r-1}$ местах $C_{n-k_1-k_2-\ldots-k_{r-1}}^{k_r}$ способами.

По правилу произведения $P_n(k_1,k_2,\ldots,k_n) =$

$$C_n^{k_1} \cdot C_{n-k_1}^{k_2} \cdot C_{n-k_1-k_2}^{k_3} \cdot \ldots \cdot C_{n-k_1-k_2-\ldots-k_{r-1}}^{k_r} =$$

$$\frac{n!}{k_1!(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \cdot \ldots \cdot \frac{(n-k_1-\ldots-k_{r-1})!}{k_r!(n-k_1-\ldots-k_r)!} =$$

$$n!/(k_1!\ k_2!\ \ldots\ k_r!).$$

10.3.6. Число размещений данной спецификации

Пусть $A_n^r(k_1,k_2,\ldots,k_n)$, $k_1+k_2+\ldots+k_n = r$, означает число всех размещений спецификации $(k_1,k_2,\ldots,k_n)$.

Размещение из n разнотипных элементов по r спецификации $(k_1,k_2,\ldots,k_n)$ есть набор из r этих элементов, из которых

k_1 элементов вида 1;
k_2 элементов вида 2;
...
k_n элементов вида n;

причем $k_1+k_2+\ldots+k_n = r$ и все $k_i \geqslant 0$ (т.е. некоторые k_i могут равняться нулю). Пусть $k_{i_1},k_{i_2},\ldots,k_{i_p}$ есть ненулевые числа среди $k_1,k_2,\ldots,k_n$. Очевидно, что $k_{i_1}+k_{i_2}+\ldots+k_{i_p} = r$. Тогда с учетом равенства $0! = 1$ число всех размещений

$$A_n^r(k_1,k_2,\ldots,k_n) = P_r(k_{i_1},k_{i_2},\ldots,k_{i_p}) =$$

$$r!/(k_{i_1}!\ k_{i_2}!\ \ldots\ k_{i_p}!) = r!/(k_1!\ k_2!\ \ldots\ k_n!).$$

Например, имеем семь шаров: красный, оранжевый, желтый, зеленый, голубой, синий, фиолетовый. Тогда

1	2	3	4	5	6	7	– число шаров ($n = 7$);
к	о	ж	з	г	с	ф	– цвета шаров;
k_1	k_2	k_3	k_4	k_5	k_6	k_7	– число шаров одного цвета;
2	0	3	0	0	1	0	$k_{i_1}+k_{i_2}+\ldots+k_{i_p} = 6$.

Тогда $A_7^6(k_1,k_2,\ldots,k_7) = P_6(k_{i_1},k_{i_2},k_{i_3}) = P_6(k_1,k_3,k_6) = r!/(k_{i_1}!\ k_{i_2}!\ k_{i_3}!) = 6!/(2!\cdot 3!\cdot 1!) = 60$.

11. ПРОИЗВОДЯЩИЕ ФУНКЦИИ ДЛЯ КОМБИНАТОРНЫХ КОНФИГУРАЦИЙ И ИХ ЧИСЕЛ

11.1. Аппарат формальных степенных рядов

Аппарат формальных степенных рядов является достаточно универсальным методом порождения и пересчета комбинаторных конфигураций

Определение. Производящая функция для множества (числа) всех комбинаторных конфигураций определенного типа, построенных на основе множества $M = \{x_1, x_2, \ldots, x_n\}$, есть функция $f(t, x_1, x_2, \ldots, x_n)$ (функция $g(t)$), в формальном разложении которой в ряд по степеням t коэффициент при t^r есть все комбинаторные конфигурации (число всех комбинаторных конфигураций) из n элементов по r рассматриваемого типа.

Пример. $M = \{x_1, x_2, x_3\}$; $|M| = 3$. Функция $f(t, x_1, x_2, x_3) =$

$$\prod_{k=1}^{3} (1+x_k t) = (1+x_1 t)\cdot(1+x_2 t)\cdot(1+x_3 t) =$$

$$1\cdot t^0 + (x_1+x_2+x_3)\cdot t + (x_1x_2+x_1x_3+x_2x_3)t^2 + x_1x_2x_3\cdot t^3$$

есть производящая функция для сочетаний из трех элементов по r $(r = 0,1,2,3)$. Положим $x_1 = x_2 = x_3 = 1$; тогда $g(t) = f(t,1,1,1) = (1 + t)^3 = 1 + 3t + 3t^2 + t^3$ есть производящая функция для числа сочетаний из трех элементов: $C_3^0 = 1$; $C_3^1 = 3$; $C_3^2 = 3$; $C_3^3 = 1$.

11.2. Производящие функции для сочетаний

11.2.1. Сочетания без повторений

Теорема. Пусть множество $M = \{x_1, x_2, \ldots, x_n\}$. Функции

$$f(t, x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} (1+x_k t) \text{ и}$$

$$g(t) = f(t,1,1,\ldots,1) = (1 + t)^n$$

являются производящими функциями для сочетаний и для числа сочетаний из n элементов.

Доказательство. Разложим функцию f по степеням t:

$$f(t, x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} (1+x_k t) = (1+x_1 t)(1+x_2 t)\ldots(1+x_n t) =$$

$$\sum_{(a_1, a_2, \ldots, a_n)\in E_2^n} (x_1 t)^{a_1}\cdot(x_2 t)^{a_2}\cdot\ldots\cdot(x_n t)^{a_n} =$$

$$\sum_{r=0}^{n} \left(\sum_{\substack{(a_1, a_2, \ldots, a_n)\in E_2^n \\ a_1+a_2+\ldots+a_n=r}} (x_1^{a_1}\cdot x_2^{a_2}\cdot\ldots\cdot x_n^{a_n}) \right) \cdot t^r,$$

где $(a_1, a_2, \ldots, a_n)$ пробегают все наборы длины n из 0 и 1.

Члены $x_1 t, x_2 t, \ldots, x_n t$ называются кодирующими множителями. Функция

$$h_r(x_1, x_2, \ldots, x_n) = \sum_{\substack{(a_1, a_2, \ldots, a_n)\in E_2^n \\ a_1+a_2+\ldots+a_n=r}} (x_1^{a_1}\cdot x_2^{a_2}\cdot\ldots\cdot x_n^{a_n}),$$

являющаяся коэффициентом при t^r, перечисляет все сочетания без повторений из n элементов по r. Следовательно, функция $f(t, x_1, x_2, \ldots, x_n)$ есть производящая функция для сочетаний без повторений из n элементов множества M.

В функции $f(t, x_1, x_2, \ldots, x_n)$ положим $x_1 = x_2 = \ldots = x_n = 1$. Тогда функция $g(t) = f(t,1,1,\ldots,1) = (1 + t)^n =$

$$\sum_{r=0}^{n} \left(\sum_{\substack{(a_1, a_2, \ldots, a_n)\in E_2^n \\ a_1+a_2+\ldots+a_n=r}} (1^{a_1}\cdot 1^{a_2}\cdot\ldots\cdot 1^{a_n}) \right) \cdot t^r = \sum_{r=0}^{n} C_n^r\, t^r,$$

в которой $h_r(1,1,\ldots,1)$ есть число всех сочетаний из n элементов по r, является производящей функцией (энумератором) для числа сочетаний из n элементов.

Замечание. 1. $(1 + t)^n = \sum_{r=0}^{n} C_n^r\, t^r$.

2. Положим в предыдущем равенстве $t=a/b$. Умножим результат на b^n и получим $(a+b)^n = \sum_{r=0}^{n} C_n^r\, a^r\, b^{n-r}$.

3. $2^n = \sum_{r=0}^{n} C_n^r$. 4. $(1-t)^n = \sum_{r=0}^{n} (-1)^r C_n^r t^r$. 5. $0 = \sum_{r=0}^{n} (-1)^r C_n^r$.

6. $\left(\sum_{r=0}^{k} a_r\right)^n = \sum_{n_1+\ldots+n_k=n} \frac{n!}{n_1! n_2! \ldots n_k!} a_1^{n_1} a_2^{n_2} \ldots a_k^{n_k}$.

11.2.2. Сочетания с повторениями с ограничениями на число повторений

Теорема. Пусть множество $M = \{x_1, x_2, \ldots, x_n\}$. Функции

$$f(t, x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} (1 + x_k t + (x_k t)^2 + \ldots + (x_k t)^{c_k}) \quad \text{и}$$

$$g(t) = f(t, 1, 1, \ldots, 1) = \prod_{k=1}^{n} (1 + t + t^2 + \ldots + t^{c_k})$$

являются производящими функциями для сочетаний и для числа сочетаний из n элементов соответственно. Причем в каждом сочетании элемент x_k встречается не более чем c_k раз, $k = 1,2, \ldots, n$.

Доказательство.

$$f(t, x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} (1 + x_k t + (x_k t)^2 + \ldots + (x_k t)^{c_k}) =$$

$$((x_1 t)^0 + (x_1 t)^1 + (x_1 t)^2 + \ldots + (x_1 t)^{c_1}) \times$$
$$((x_2 t)^0 + (x_2 t)^1 + (x_2 t)^2 + \ldots + (x_2 t)^{c_2}) \times$$
$$\ldots$$
$$((x_n t)^0 + (x_n t)^1 + (x_n t)^2 + \ldots + (x_n t)^{c_n}) =$$

$$\sum_{\substack{(a_1, a_2, \ldots, a_n) \in N^n \\ 0 \leqslant a_k \leqslant c_k,\ k = 1, 2, \ldots, n}} (x_1 t)^{a_1} \cdot (x_2 t)^{a_2} \cdot \ldots \cdot (x_n t)^{a_n} =$$

$$\sum_{r=0}^{c_1 + c_2 + \ldots + c_n} \left(\sum_{\substack{(a_1, a_2, \ldots, a_n) \in N^n \\ a_1 + a_2 + \ldots + a_n = r \\ 0 \leqslant a_k \leqslant c_k,\ k = 1, 2, \ldots, n}} (x_1^{a_1} \cdot x_2^{a_2} \cdot \ldots \cdot x_n^{a_n}) \right) \cdot t^r.$$

Функция

$$h_r(x_1, x_2, \ldots, x_n) = \sum_{\substack{(a_1, a_2, \ldots, a_n) \in N^n \\ a_1 + a_2 + \ldots + a_n = r \\ 0 \leqslant a_k \leqslant c_k,\ k = 1, 2, \ldots, n}} (x_1^{a_1} \cdot x_2^{a_2} \cdot \ldots \cdot x_n^{a_n}),$$

являющаяся коэффициентом при t^r, перечисляет все сочетания из n элементов по r. При этом в каждом сочетании элемент x_k

встречается не более чем $a_k \leqslant c_k$ раз, $k = 1,2, \ldots, n$. Следовательно, $f(t, x_1, x_2, \ldots, x_n)$ есть искомая производящая функция для сочетаний с повторениями из n элементов множества M, причем в каждом сочетании элемент x_k встречается не более чем c_k раз, $k = 1,2, \ldots, n$. В функции $f(t, x_1, x_2, \ldots, x_n)$ положим $x_1 = x_2 = \ldots = x_n = 1$. Тогда функция

$$g(t) = f(t, 1, 1, \ldots, 1) =$$
$$\sum_{r=0}^{c_1 + c_2 + \ldots + c_n} \left(\sum_{\substack{(a_1, a_2, \ldots, a_n) \in N^n \\ a_1 + a_2 + \ldots + a_n = r \\ 0 \leqslant a_k \leqslant c_k,\ k = 1, 2, \ldots, n}} (1^{a_1} \cdot 1^{a_2} \cdot \ldots \cdot 1^{a_n}) \right) \cdot t^r,$$

в которой число $h_r(1,1, \ldots, 1)$ есть число всех сочетаний из n элементов по r, причем каждое сочетание содержит элемент x_k не более чем c_k раз, $k = 1,2, \ldots, n$, является производящей функцией (энумератором) для числа сочетаний с повторениями из n элементов, причем каждое сочетание содержит элемент x_k не более чем c_k раз, $k = 1,2, \ldots, n$.

11.2.3. Сочетания с повторениями без ограничений на число повторений

Теорема. Пусть множество $M = \{x_1, x_2, \ldots, x_n\}$. Функции

$$f(t, x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} (1 + x_k t + (x_k t)^2 + \ldots) \quad \text{и}$$

$$g(t) = f(t, 1, 1, \ldots, 1) = \prod_{k=1}^{n} (1 + t + t^2 + \ldots)$$

являются производящими функциями для сочетаний и для числа сочетаний из n элементов с повторениями.

Доказательство. $f(t, x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} (1 + x_k t + (x_k t)^2 + \ldots) =$

$$((x_1 t)^0 + (x_1 t)^1 + (x_1 t)^2 + \ldots) \times$$
$$((x_2 t)^0 + (x_2 t)^1 + (x_2 t)^2 + \ldots) \times$$
$$\ldots$$
$$((x_n t)^0 + (x_n t)^1 + (x_n t)^2 + \ldots) =$$

$$\sum_{(a_1,a_2,\ldots,a_n)\in N^n} (x_1t)^{a_1}\cdot(x_2t)^{a_2}\cdot\ldots\cdot(x_nt)^{a_n} =$$

$$\sum_{r=0}^{\infty}\left(\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n\\ a_1+a_2+\ldots+a_n=r}} (x_1^{a_1}\cdot x_2^{a_2}\cdot\ldots\cdot x_n^{a_n})\right)\cdot t^r.$$

Здесь бесконечная сумма берется по всем наборам $(a_1,a_2,\ldots,a_n)$ неотрицательных целых чисел. Функция

$$h_r(x_1,x_2,\ldots,x_n) = \sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n\\ a_1+a_2+\ldots+a_n=r}} (x_1^{a_1}\cdot x_2^{a_2}\cdot\ldots\cdot x_n^{a_n}),$$

являющаяся коэффициентом при t^r, перечисляет все сочетания с повторениями из n элементов по r. Следовательно, $f(t,x_1,x_2,\ldots,x_n)$ есть искомая производящая функция для сочетаний с повторениями из n элементов множества M.

В функции $f(t,x_1,x_2,\ldots,x_n)$ положим $x_1 = x_2 = \ldots = x_n = 1$. Тогда функция

$$g(t) = f(t,1,1,\ldots,1) = (1 + t + t^2 + \ldots)^n =$$

$$\sum_{r=0}^{\infty}\left(\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n\\ a_1+a_2+\ldots+a_n=r}} (1^{a_1}\cdot 1^{a_2}\cdot\ldots\cdot 1^{a_n})\right)\cdot t^r,$$

в которой $h_r(1,1,\ldots,1)$ есть число всех сочетаний с повторениями из n элементов по r, является производящей функцией (энумератором) для числа сочетаний с повторениями из n элементов.

Замечание. $g(t) = f(t,1,1,\ldots,1) = (1 + t + t^2 + \ldots)^n =$

$$\left(\frac{1}{1-t}\right)^n = (1-t)^{-n} =$$

$$\sum_{r=0}^{\infty}\frac{(-n)(-n-1)(-n-2)\ldots(-n-r+1)}{r!}\cdot(-t)^r =$$

$$\sum_{r=0}^{\infty}\frac{n(n+1)(n+2)\ldots(n+r-1)}{r!}t^r = \sum_{r=0}^{\infty}\frac{A_{n+r-1}^r}{r!}t^r =$$

$$\sum_{r=0}^{\infty}\frac{(n+r-1)!}{(n+r-1-r)!\cdot r!}t^r = \sum_{r=0}^{\infty}C_{n+r-1}^r t^r.$$

11.3. Производящие функции для размещений с повторениями

Теорема. Пусть множество $M = \{x_1,x_2,\ldots,x_n\}$. Функции

$$f(t,x_1,x_2,\ldots,x_n) = \prod_{k=1}^{n}\left(1 + \frac{x_kt}{1!} + \frac{(x_kt)^2}{2!} + \ldots + \frac{(x_kt)^{c_k}}{c_k!}\right) \text{ и}$$

$$g(t) = f(t,1,1,\ldots,1) = \prod_{k=1}^{n}\left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \ldots + \frac{t^{c_k}}{c_k!}\right)$$

являются производящими функциями для размещений и для числа размещений из n элементов; в каждом размещении элемент x_k встречается не более чем c_k раз, $k = 1, 2,\ldots,n$.

Доказательство.

$$f(t,x_1,x_2,\ldots,x_n) = \prod_{k=1}^{n}\left(1 + \frac{x_kt}{1!} + \frac{(x_kt)^2}{2!} + \ldots + \frac{(x_kt)^{c_k}}{c_k!}\right) =$$

$$\left(\frac{(x_1t)^0}{0!} + \frac{x_1t}{1!} + \frac{(x_1t)^2}{2!} + \ldots + \frac{(x_1t)^{c_1}}{c_1!}\right)\times$$

$$\left(\frac{(x_2t)^0}{0!} + \frac{x_2t}{1!} + \frac{(x_2t)^2}{2!} + \ldots + \frac{(x_2t)^{c_2}}{c_2!}\right)\times$$

$$\ldots$$

$$\left(\frac{(x_nt)^0}{0!} + \frac{x_nt}{1!} + \frac{(x_nt)^2}{2!} + \ldots + \frac{(x_nt)^{c_n}}{c_n!}\right) =$$

$$\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n\\ 0\leqslant a_k\leqslant c_k,\ k=1,\ldots,n}} \frac{(x_1t)^{a_1}}{a_1!}\cdot\frac{(x_2t)^{a_2}}{a_2!}\cdot\ldots\cdot\frac{(x_nt)^{a_n}}{a_n!} =$$

$$\sum_{r=0}^{c_1+c_2+\ldots+c_n}\left(\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n\\ a_1+a_2+\ldots+a_n=r\\ 0\leqslant a_k\leqslant c_k,\ k=1,2,\ldots,n}}\frac{x_1^{a_1}\,x_2^{a_2}\ldots x_n^{a_n}}{a_1!\,a_2!\ldots a_n!}\right)\cdot t^r.$$

Здесь сумма берется по всем наборам $(a_1,a_2,\ldots,a_n)$, неотрицательных целых чисел, для которых $a_k\leqslant c_k$, $k = 1,2,\ldots,n)$.

Функция

$$h_r(x_1,x_2,\ldots,x_n) = \sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n \\ a_1+a_2+\ldots+a_n=r \\ 0\leqslant a_k\leqslant c_k,\ k=1,2,\ldots,n}} \frac{x_1^{a_1}\,x_2^{a_2}\,\ldots\,x_n^{a_n}}{a_1!\,a_2!\,\ldots\,a_n!},$$

являющаяся коэффициентом при t^r, перечисляет сочетания из n элементов по r. При этом в каждом сочетании элемент x_k встречается $a_k \leqslant c_k$ раз, $k = 1,2,\ldots,n$. В числителе сочетание $(x_1^{a_1},x_2^{a_2},\ldots,x_n^{a_n})$ из r элементов, среди которых имеется a_k элементов x_k, $k=1,2,\ldots,n$, породит s размещений спецификации $A_n^r(a_1,a_2,\ldots,a_n)$, причем $s = \dfrac{r!}{a_1!\,a_2!\,\ldots\,a_n!}$.

Тогда функция

$$g(t) = f(t,1,1,\ldots,1) = \prod_{k=1}^{n}\left\{1 + \frac{t}{1!} + \frac{t^2}{2!} + \ldots + \frac{t^{c_k}}{c_k!}\right\} =$$

$$\sum_{r=0}^{c_1+c_2+\ldots+c_n}\left(\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n \\ a_1+a_2+\ldots+a_n=r \\ 0\leqslant a_k\leqslant c_k,\ k=1,2,\ldots,n}} \frac{1}{a_1!\,a_2!\,\ldots\,a_n!}\right)\cdot t^r =$$

$$\sum_{r=0}^{c_1+c_2+\ldots+c_n}\left(\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n \\ a_1+a_2+\ldots+a_n=r \\ 0\leqslant a_k\leqslant c_k,\ k=1,2,\ldots,n}} \frac{r!}{a_1!\,a_2!\,\ldots\,a_n!}\right)\cdot \frac{t^r}{r!} =$$

$$\sum_{r=0}^{c_1+c_2+\ldots+c_n}\left(\sum_{\substack{(a_1,a_2,\ldots,a_n)\in N^n \\ a_1+a_2+\ldots+a_n=r \\ 0\leqslant a_k\leqslant c_k,\ k=1,2,\ldots,n}} A_n^r(a_1,a_2,\ldots,a_n)\right)\cdot \frac{t^r}{r!}.$$

Следовательно, $f(t,x_1,x_2,\ldots,x_n)$ есть искомая производящая функция для размещений с повторениями из n элементов множества M (с учетом сделанного выше замечания), причем в каждом размещении элемент x_k встречается не более чем c_k раз, $k = 1,2,\ldots,n$. В функции $f(t,x_1,x_2,\ldots,x_n)$ положим $x_1 = x_2 = \ldots = x_n = 1$. Тогда функция $g(t)$ является производящей функцией (энумератором) для числа размещений с повторениями из n элементов, причем каждое размещение содержит элемент x_k не более чем c_k раз ($k = 1,2,\ldots,n$). Коэффициент

при $\dfrac{t^r}{r!}$ есть число размещений из n элементов по r, причем в каждом размещении элемент x_k встречается не более c_k раз, $k = 1,2,\ldots,n$.

Замечание. Для числа размещений с повторениями без ограничений на число повторений производящая функция

$$g(t) = \left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \ldots\right)^n =$$

$$e^{nt} = \sum_{r=0}^{\infty} n^r\,\frac{t^r}{r!}.$$

Коэффициент n^r при $\dfrac{t^r}{r!}$ равен числу всех размещений с неограниченными повторениями из n элементов по r, т.е. числу всех наборов длины r из n элементов.

12. КОМБИНАТОРНО ЛОГИЧЕСКИЙ АППАРАТ

12.1. Включения и исключения

Пусть имеем N объектов $1,2,\ldots,N$, которые могут обладать или не обладать n свойствами $1,2,\ldots,n$. Примем следующие обозначения.

$N(0)$ (или $N(\neg 1,\neg 2,\ldots,\neg n)$) – число объектов, не обладающих ни одним из свойств $1,2,3,\ldots,n$.

$N(i_1,i_2,\ldots,i_r)$ – число объектов, обладающих свойствами $i_1,i_2,\ldots,i_r$ (возможно и другими свойствами).

$N_{=k}$ – число объектов, обладающих в точности k свойствами.

$N_{\geqslant k}$ – число объектов, обладающих не менее чем k свойствами.

Найдем формулы для вычисления всех этих чисел.

Пример. Пусть имеем шесть объектов 1 – 6, которые могут обладать или не обладать пятью свойствами 1 – 5 (табл.12.1).

Заметим, что число объектов, не обладающих ни одним из свойств 1 – 4, равно $N(\neg 1,\neg 2,\neg 3,\neg 4) = 3$; сюда включаются объекты, дополнительно обладающие свойством 5 (объекты 5,6) и не обладающие свойством 5 (объект 2). Число объектов, не обладающими свойствами 1 – 4 и обладающими свойством 5, равно $N(\neg 1,\neg 2,\neg 3,\neg 4,5) = 2$ (объекты 5,6). Тогда число объектов, не обладающих ни одним из свойств 1 – 5 равно $N(0) = N(\neg 1,$

$\neg 2, \neg 3, \neg 4, \neg 5) = N(\neg 1, \neg 2, \neg 3, \neg 4) - N(\neg 1, \neg 2, \neg 3, \neg 4, 5) = 3-2 = 1.$

В общем случае

$N(0) = N(\neg 1, \neg 2, \ldots, \neg n-1) - N(\neg 1, \neg 2, \ldots, \neg n-1, n).$

Таблица 12.1

		свойства				
		1	2	3	4	5
объекты	1	1	1	0	1	0
	2	0	0	0	0	0
	3	1	1	1	1	1
	4	1	0	1	0	0
	5	0	0	0	0	1
	6	0	0	0	0	1

$N(1) = 3;\ N(1,2) = 2;\ N(2,4) = 2;$
$N(2) = 2;\ N(1,3) = 2;\ N(2,5) = 1;$
$N(3) = 2;\ N(1,4) = 2;\ N(3,4) = 1;$
$N(4) = 2;\ N(1,5) = 1;\ N(3,5) = 1;$
$N(5) = 3;\ N(2,3) = 1;\ N(4,5) = 1;$

$N(1,2,3) = 1;\ N(1,4,5) = 1;\ N(1,2,3,4) = 1;\ N(1,2,3,4,5) = 1.$
$N(1,2,4) = 2;\ N(2,3,4) = 1;\ N(1,2,3,5) = 1;$
$N(1,2,5) = 1;\ N(2,3,5) = 1;\ N(1,2,4,5) = 1;$
$N(1,3,4) = 1;\ N(2,4,5) = 1;\ N(1,3,4,5) = 1;$
$N(1,3,5) = 1;\ N(3,4,5) = 1;\ N(2,3,4,5) = 1;$

Теорема. Справедлива следующая формула включений и исключений. $N(0) = N - \sum_{1 \leqslant i \leqslant n} N(i) + \sum_{1 \leqslant i_1 < i_2 \leqslant n} N(i_1, i_2) - \ldots +$

$$(-1)^r \sum_{1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant n} N(i_1, i_2, \ldots, i_r) + \ldots +$$

$$(-1)^n\ N(1,2,\ldots,n).$$

Доказательство. Пусть

$$S_r = \sum_{1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant n} N(i_1, i_2, \ldots, i_r).$$

Далее индукция по числу свойств n.

Базис. $n = 1$ (одно свойство). $N(0) = N - N(1) = N - S_1$.

Предположение индукции. Допустим, что формула включений и исключений справедлива для $n-1$ свойств.

Шаг индукции. Покажем, что формула включений и исключений справедлива для n свойств: $N(0) = N - S_1 + S_2 - \ldots + (-1)^n S_n$. По предположению индукции для свойств $1,2,\ldots,n-1$ имеем

$$N(\neg 1, \neg 2, \ldots, \neg n-1) = N - S_1 + S_2 - \ldots + (-1)^{n-1} S_{n-1} =$$

$$N - \sum_{1 \leqslant i \leqslant n-1} N(i) + \sum_{1 \leqslant i_1 < i_2 \leqslant n-1} N(i_1, i_2) - \ldots +$$

$$(-1)^{n-1}\ N(1,2,\ldots,n-1).$$

Эта формула справедлива и для n свойств при фиксировании последнего свойства n:

$$N(\neg 1, \neg 2, \ldots, \neg n-1, n) = N(n) - \sum_{1 \leqslant i \leqslant n-1} N(i,n) +$$

$$\sum_{1 \leqslant i_1 < i_2 \leqslant n-1} N(i_1, i_2, n) - \ldots + (-1)^{n-1}\ N(1,2,\ldots,n-1,n).$$

Вычтем последнюю формулу из предпоследней:

$$N(\neg 1, \neg 2, \ldots, \neg n-1) - N(\neg 1, \neg 2, \ldots, \neg n-1, n) = N(0) =$$

$$N - (\sum_{1 \leqslant i \leqslant n-1} N(i) + N(n)) +$$

$$(\sum_{1 \leqslant i < j \leqslant n-1} N(i,j) + \sum_{1 \leqslant i \leqslant n-1} N(i,n)) - \ldots +$$

$$\ldots$$

$$(-1)^{n-1}\ (N(1,2,\ldots,n-1) + \sum_{1 \leqslant i_1 < \ldots < i_{n-2} \leqslant n-1}$$

$$N(i_1, \ldots, i_{n-2}, n)) + (-1)^n\ N(1,2,\ldots,n) =$$

$$\sum_{r=0}^{n} (-1)^r \sum_{1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant n} N(i_1, i_2, \ldots, i_r).$$

Замечание. Аналогично можно показать справедливость двух следующих формул:

$$N_{=k} = \sum_{j=0}^{n-k} (-1)^j\ C_{k+j}^{j} \sum_{1 \leqslant i_1 < i_2 < \ldots < i_{k+j} \leqslant n} N(i_1, i_2, \ldots, i_{k+j});$$

$$N_{\geqslant k} = \sum_{j=0}^{n-k} (-1)^j\ C_{k-1+j}^{k-1} \sum_{1 \leqslant i_1 < i_2 < \ldots < i_{k+j} \leqslant n} N(i_1, i_2, \ldots, i_{k+j}).$$

Для примера из табл.22.1 найдем $N(0)$, $N_{=2}$, $N_{\geqslant 2}$:

$$N(0) = N - \sum_{1 \leqslant i \leqslant 5} N(i) + \sum_{1 \leqslant i_1 < i_2 \leqslant 5} N(i_1, i_2) -$$

$$\sum_{1\leqslant i_1<i_2<i_3\leqslant 5} N(i_1,i_2,i_3) + \sum_{1\leqslant i_1<i_2<i_3<i_4\leqslant 5} N(i_1,i_2,i_3,i_4) -$$

$$N(1,2,3,4,5) = 6 - 12 + 14 - 11 + 5 - 1 = 1;$$

$$N_{=2} = \sum_{j=0}^{5-2} (-1)^j C_{2+j}^{j} \sum_{1\leqslant i_1<i_2<\ldots<i_{2+j}\leqslant 5} N(i_1,i_2,\ldots,i_{2+j}) =$$

$$1\cdot 14 - 3\cdot 11 + 6\cdot 5 - 10\cdot 1 = 14 - 33 + 30 - 10 = 1;$$

$$N_{\geqslant 2} = \sum_{j=0}^{5-2} (-1)^j C_{2-1+j}^{j} \sum_{1\leqslant i_1<i_2<\ldots<i_{2+j}\leqslant 5} N(i_1,i_2,\ldots,i_{2+j}) =$$

$$1\cdot 14 - 2\cdot 11 + 3\cdot 5 - 4\cdot 1 = 14 - 22 + 15 - 4 = 3.$$

Пример. Найдем число положительных натуральных чисел, не больших 1000 и не делящихся ни на одно из чисел 3,5,7. Выделим следующие свойства, которыми могут обладать или не обладать числа $1,2,\ldots,1000$.

Свойство 1: число n делится на 3.
Свойство 2: число n делится на 5.
Свойство 3: число n делится на 7.

В табл.22.2 приведены свойства, им удовлетворяющие множества чисел и их число.

$$N(0) = 1000 - \sum_{1\leqslant i\leqslant 3} N(i) + \sum_{1\leqslant i<j\leqslant 3} N(i,j) - N(1,2,3) =$$

$$1000 - (333 + 200 + 142) + (66 + 47 + 28) - 9 = 457.$$

12.2. Приложения формулы включений и исключений

12.2.1. Задача о беспорядках

Пусть имеем множество $M = \{1,2,\ldots,n\}$ из n элементов, и пусть подстановка (т.е. взаимно однозначная функция) $s : M \to M$. Подстановка s обладает свойством i, если $s(i) = i$, т.е. подстановка s элемент i переводит в себя. Подстановка s есть беспорядок, если $s(i) \neq i$ $\forall i \in M$. Например, подстановка $s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$ обладает свойствами 2 и 4 и не обладает свойствами 1 и 3. Подстановка $s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$ обладает свойством 3 и не

обладает свойствами 1,2,4. Подстановка $s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ является беспорядком.

Таблица 9.2

Свойства	Множество чисел	Число
1	$\{3k : k = 1,2,\ldots,333\}$	$N(1) = 333$
2	$\{5k : k = 1,2,\ldots,200\}$	$N(2) = 200$
3	$\{7k : k = 1,2,\ldots,142\}$	$N(3) = 142$
1,2	$\{15k : k = 1,2,\ldots,66\}$	$N(1,2) = 66$
1,3	$\{21k : k = 1,2,\ldots,47\}$	$N(1,3) = 47$
2,3	$\{35k : k = 1,2,\ldots,28\}$	$N(2,3) = 28$
1,2,3	$\{105k : k = 1,2,\ldots,9\}$	$N(1,2,3) = 9$

Пусть подстановка s обладает свойствами $i_1,i_2,\ldots,i_r$ (возможно, что и другими свойствами). Тогда

$$s = \begin{pmatrix} 1 & 2 & \ldots & i_1 & \ldots & i_2 & \ldots & i_r & \ldots & n \\ s(1) & s(2) & \ldots & i_1 & \ldots & i_2 & \ldots & i_r & \ldots & s(n) \end{pmatrix}.$$

Число таких подстановок равно $(n - r)! = N(i_1,i_2,\ldots,i_r)$. По методу включений и исключений число подстановок s, не обладающих ни одним из свойств $1,2,\ldots,n$ (т.е. число беспорядков) $N(0) =$

$$N(\neg 1,\neg 2,\ldots,\neg n) = \sum_{r=0}^{n} (-1)^r \sum_{1\leqslant i_1<i_2<\ldots<i_r\leqslant n} N(i_1,i_2,\ldots,i_r) =$$

$$\sum_{r=0}^{n} (-1)^r \sum_{\{i_1,i_2,\ldots,i_r\}\subseteq\{1,2,\ldots,n\}} (n - r)! =$$

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

Окончательно число беспорядков $N(0) = n! \sum_{r=0}^{n} (-1)^r \frac{1}{r!}$.

Замечание. $N(0) = n! \sum_{r=0}^{n} (-1)^r \frac{1}{r!} \approx n! \sum_{r=0}^{\infty} (-1)^r \frac{1}{r!} =$ $n!e^{-1}$, т.е. число беспорядков $N(0) \approx \frac{n!}{e}$.

12.2.2. Задача о встречах

Определим число подстановок

$$s = \begin{pmatrix} 1 & 2 & \dots & i_1 & \dots & i_2 & \dots & i_k & \dots & n \\ s(1) & s(2) & & i_1 & & i_2 & & i_k & & s(n) \end{pmatrix},$$

обладающих ровно k свойствами $s(i) = i$. Подсчет числа таких подстановок сведем к задаче о беспорядках. Искомое число

$$N_{=k} = C_n^k \cdot \left\{ (n-k)! \cdot \sum_{r=0}^{n-k} (-1)^r \frac{1}{r!} \right\},$$

где множитель C_n^k дает число способов, которыми можно выбрать k элементов (из данных n элементов), обладающих свойством $s(i)=i$, а второй множитель дает число беспорядков для остальных $n-k$ элементов (не обладающих свойством $s(i)=i$).

Так как $C_n^k = \dfrac{n!}{k!(n-k)!}$, то $N_{=k} = \dfrac{n!}{k!} \sum\limits_{r=0}^{n-k} (-1)^r \dfrac{1}{r!}$.

Число $N_{=k}$ можно интерпретировать как число встреч $n-k$ лиц из данных n лиц.

Замечание. $N_{=k} \approx \dfrac{n!}{k!} \sum\limits_{r=0}^{\infty} (-1)^r \dfrac{1}{r!} = \dfrac{n!}{k!e}$.

Часть 3

АЛГЕБРА ЛОГИКИ И ПРЕДИКАТЫ

13. АЛГЕБРА ЛОГИКИ

13.1. Функции алгебры логики

Пусть $E_2 = \{0,1\}$ – двухэлементное множество. *Набор* длины n из 0 и 1 есть последовательность длины n, составленная из 0 и 1.

Пример.

(0); (1) – наборы длины 1;

(0,0);(0,1);(1,0);(1,1) – наборы длины 2;

(0,0,0); (0,0,1); (0,1,0); (0,1,1); (1,0,0); (1,0,1); (1,1,0);

(1,1,1) – наборы длины 3;

$(0,0,\dots,0,0);(0,0,\dots,1);\dots;(a_1,a_2,\dots,a_n);\dots$;

$(1,1,\dots,1,1)$ – наборы длины n.

Пусть E_2^n есть множество всех наборов из 0 и 1 длины n.

Теорема. Число $h(n)$ всех наборов из E_2^n равно 2^n.

Доказательство. Индукция по n.

Базис. $n = 1$. $h(1) = 2$.

Предположение индукции. Пусть $h(n) = 2^n$.

Шаг индукции. Покажем, что $h(n+1) = 2^{n+1}$. Разобьем все наборы длины $n+1$ на два класса: класс наборов, начинающихся с 0, и класс наборов, начинающихся с 1.

0,0,...,0,0	1,0,...,0,0
0,0,...,0,1	1,0,...,0,1
0,0,...,1,0	1,0,...,1,0
...	...
0,1,...,1,1	1,1,...,1,1

Число всех наборов длины $n+1$, начинающихся с 0 (так же как и число всех наборов, начинающихся с 1), равно числу всех наборов длины n и по предположению индукции равно 2^n. Тогда $h(n+1) = 2^n + 2^n = 2 \cdot 2^n = 2^{n+1}$. Теорема доказана.

Определение. *Функция алгебры логики* (или *булева функция*) есть функция, аргументы и значения которой принимают лишь два значения 0 и 1.

Класс всех функций алгебры логики обозначим через P_2. Вместо слов "функция алгебы логики" будем иногда говорить

просто "функция".

Замечание. n-местная булева фуннкция f есть отображение $f: E_2^n \to E_2$.

Теорема. Число всех n-местных функций алгебры логики равно 2^{2^n}.

Доказательство. В табл.13.1 перечислены все n-местные функции.

Число всех строк равно 2^n, т.е. равно числу всех наборов длины n из 0 и 1. Число всех функций алгебры логики от n переменных равно 2^{2^n}, т.е. равно числу всех наборов длины 2^n из 0 и 1.

Теорема доказана.

В табл.13.2 приведены некоторые часто используемые в пракике функции: 0 – константа нуль; 1 – константа единица; x – тождественная функция; $\bar{x}$ – отрицание (обозначается также через $\neg x$); $x \vee y$ – дизъюнкция; $x \,\&\, y$ – конъюнкция (обозначается также через $x \cdot y$ или xy); $x \to y$ – импликация; $x + y$ – сложение (по модулю два); $x \equiv y$ – эквивалентность (равносильность); $x \mid y$ – штрих Шеффера; $x \uparrow y$ – стрелка Пирса.

13.2. Формулы. Реализация функций формулами

Пусть $F=\{f_1^{n_1}, f_2^{n_2}, \ldots\}$ есть множество функциональных символов указанной вверху местности и $\{x_1, x_2, \ldots\}$ есть множество символов переменных. Верхние индексы $n_1, n_2, \ldots$ могут опускаться, если их значение предполагается известным.

Определение. *Формула* (алгебры логики) над F определяется индуктивно следующим образом.

1. Символ переменной есть формула над F.

Таблица 13.1

x_1	x_2	...	x_{n-1}	x_n	f_0	f_1	f_2	...	f_r
0	0	...	0	0	0	0	0	...	1
0	0	...	0	1	0	0	0	...	1
		...						...	
1	1	...	1	0	0	0	1	...	1
1	1	...	1	1	0	1	0	...	1

Таблица 13.2

x	y	0	1	$x \vee y$	$x \,\&\, y$	$x \to y$	$x+y$	$x \equiv y$	$x \mid y$	$x \uparrow y$	x	$\bar{x}$	x	0	1
0	0	0	1	0	0	1	0	1	1	1	0	1	0	0	1
0	1	0	1	1	0	1	1	0	1	0	1	0	1	0	1
1	0	0	1	1	0	0	1	0	1	0					
1	1	0	1	1	1	1	0	1	0	0					

2. Если f есть функциональный символ арности m из F и $x_1, \ldots, x_m$ есть символы различных переменных, то выражение $f(x_1, \ldots, x_m)$ есть формула над F и $\{x_1, \ldots, x_m\}$ есть множество ее переменных.

3. Если $A(x_1, \ldots, x_m)$ есть формула над F, где $\{x_1, \ldots, x_m\}$ есть множество ее переменных, и если каждое из выражений $A_1, \ldots, A_m$ есть либо формула над F, либо символ переменной, то выражение $A(A_1, \ldots, A_m)$ есть формула над F, причем множество ее переменных есть объединение множеств переменных формул $A_1, \ldots, A_m$.

Пример. $F = \{f_1(x,y), f_2(x,y,z), f_3(x)\}$. Следующие выраеия являются формулами над F:

$f_1(x,y)$; $f_2(x,y,z)$; $f_3(x)$; $f_1(t,z)$; $f_2(t,t,t)$;

$f_3(f_2(x,y,x))$; $f_1(f_2(y,t,z), f_3(f_1(x,)))$.

Сопоставим каждому функциональному символу f^n из F некоторую функцию $f: E_2^n \to E_2$. Если множество функциональных символов $F = \{f_1^{n_1}, f_2^{n_2}, \ldots\}$, то пусть множество соответствующих функций $F = \{f_1^{n_1}, f_2^{n_2}, \ldots\}$.

Между F и F существует взаимно однозначное соответствие, при котором $f_i^{n_i}$ соответствует $f_i^{n_i}$, $i=1,2,\ldots$

Каждой формуле над F сопоставим функцию индуктивно следующим образом.

1. Переменной x сопоставим тождественную функцию x.

2. Формуле $f(x_1, \ldots, x_m)$ над F сопоставим функцию $f(x_1, \ldots, x_m)$ из F.

3. Если формуле $A(x_1, \ldots, x_m)$ над F сопоставлена функция $f(x_1, \ldots, x_m)$, а формулам $A_1, \ldots, A_m$ над F сопоставлены функции $f_1, \ldots, f_m$, то формуле $A(A_1, \ldots, A_m)$ над F сопоставим функцию $f(f_1, \ldots, f_m)$ из P_2.

Таким образом каждая формула над F реализует некоторую функцию из P_2. Пусть формула $A(x_1, \ldots, x_n)$ реализует некото-

Таблица 13.3

x y	$f(x,y)$
0 0	0
0 1	0
1 0	1
1 1	0

рую функцию $f(x_1,\ldots,x_n)$, и пусть $a = (a_1,\ldots,a_n)$ есть набор длины n из 0 и 1. Тогда значение формулы A на наборе a есть $f(a)$, т.е. $A(a_1,\ldots,a_n) = f(a_1,\ldots,a_n)$.

В дальнейшем формулу $A(x_1,\ldots,x_m)$ будем отождествлять с функцией $f(x_1,\ldots,x_n)$, которую формула A реализует и обозначать эту функцию через $A(x_1,\ldots,x_n)$. Говоря о формуле A над F, будем говорить просто о формуле A, не упоминая об F, если из контекста ясно, о каком множестве F идет речь.

Имея в виду взаимно однозначное соответствие между F и F, вместо слов "формула $A(x_1,\ldots,x_m)$ над множеством функциональных символов F" будем говорить "формула $A(x_1,\ldots,x_n)$ над множеством функций F".

Пример. Формула $xy \vee \bar{x}$ реализует функцию $f(x,y)$, приведенную в табл.13.3.

Определение. Функция f есть *суперпозиция* над F, если f реализуется некоторой формулой над F.

Определение. Пусть A есть некоторая формула над множеством функций F из P_2. Если A есть $f(x_1,\ldots,x_m)$ из F, то единственной *подформулой* формулы A является она сама. Если A есть формула $f(A_1,\ldots,A_m)$, где $f \in F$, а $A_1,\ldots,A_m$ – некоторые формулы над F, то подформулами формулы A являются она сама и все подформулы формул $A_1,\ldots,A_m$.

Замечание. Пусть $A(B)$ означает, что B есть подформула формулы A.

Определение. Класс функций F называется функционально *замкнутым*, если вместе с любыми своими функциями он содержит и любую их суперпозицию.

Определение. Множество функций $[F]$ называется *замыканием* класса функций F, если оно содержит все суперпозиции функций над множеством F и не содержит никаких других функций.

Замечание. 1. $F \subseteq [F]$.

2. $[[F]] = [F]$.

3. $F_1 \subseteq F_2$ влечет $[F_1] \subseteq [F_2]$.

4. Множество функций F замкнуто, если $F = [F]$.

Определение. Система G функций из замкнутого класса F *полна в* F (является *порождающей системой для* F), если $[G] = F$. Система функций H *полна* (в P_2), если $[H] = P_2$. Полная в F система функций G называется *базисом* в F, если никакая собственная подсистема в G не является полной в F.

13.3. Равносильные преобразования формул

Пусть A_1 и A_2 – формулы, а $x_1,\ldots,x_n$ есть полный список (множество) их переменных. Формулы A_1 и A_2 называются *равносильными* (равными), если для любого набора значений аргументов $x_1,\ldots,x_n$ они принимают одинаковые значения.

Пример. $A_1(x,y) = \overline{xy \vee \bar{x}}$; $A_2(x,y) = \overline{x \to y}$; $A_3(x) = x$; $A_4(x,y) = x \vee y$. В табл.13.4 приведены значения формул A_1–A_4, из которой видно, что $A_1 = A_2$, $A_2 \neq A_3$; $A_2 \neq A_4$; $A_3 \neq A_4$.

В инженерной практике наиболее распространены представления функций формулами, построенными с помощью конъюнкции, дизъюнкции, отрицания, констант 0 и 1, т.е. формулами над $F = \{x\&y, x \vee y, \neg x, 0, 1\}$. Такие формулы называются *булевыми*. Иногда в F включают импликацию.

Примем соглашение об опускании скобок в соответствии со следующим приоритетом операций: $\neg$, & , $\vee$, $\to$. Укажем некоторые свойства операций & , $\vee$, $\neg$. Эти операции (как и их свойства) называются булевыми.

Пусть A, B, C – произвольные формулы над F. Тогда справедливы следующие свойства булевых операций.

1. Идемпотентность.
   $A \,\&\, A = A$; $A \vee A = A$.
2. Коммутативность.
   $A \,\&\, B = B \,\&\, A$; $A \vee B = B \vee A$.

Таблица 13.4

x y	A_1	A_2	A_3	A_4
0 0	0	0	0	0
0 1	0	0	0	1
1 0	1	1	1	1
1 1	0	0	1	1

3. Ассоциативность.
   $A \,\&\, (B \,\&\, C) = (A \,\&\, B) \,\&\, C$;
   $A \vee (B \vee C) = (A \vee B) \vee C$.
4. Правило поглощения.
   $A \,\&\, (A \vee B) = A$; $A \vee A \,\&\, B = A$.
5. Дистрибутивность.
   $A \,\&\, (B \vee C) = A \,\&\, B \vee A \,\&\, C$;
   $A \vee B \,\&\, C = (A \vee B) \,\&\, (A \vee C)$.
6. Инволюция. $\bar{\bar{A}} = A$.
7. Свойства констант.
   $A \,\&\, 1 = A$; $A \vee 0 = A$;
   $A \,\&\, 0 = 0$; $A \vee 1 = 1$.
8. Закон исключенного третьего и закон противоречия.
   $A \vee \bar{A} = 1$; $A \,\&\, \bar{A} = 0$.
9. Правила де Моргана.
   $\overline{A \,\&\, B} = \bar{A} \vee \bar{B}$; $\overline{A \vee B} = \bar{A} \,\&\, \bar{B}$.
10. Связь импликации и дизъюнкции.
   $A \to B = \bar{A} \vee B$.

Все эти равенства устанавливаются непосредственной проверкой.

Правило подстановки (замены равным). Если

1) A, C есть формулы;
2) B есть подформула формулы A, то есть A есть $A(B)$;
3) $B = C$, то $A(B) = A(C)$.

Коротко правило подстановки записывают так:

$$\frac{A(B),\ B = C}{A(B) = A(C)}.$$

Примем без доказательства следующее утверждение.

Теорема. Если A и B – булевы формулы и $A = B$, то с помощью булевых равенств 1 – 9 и правила подстановки от формулы A можно перейти к формуле B за конечное число шагов.

Эта теорема широко используется при упрощении формул.

Пример. $\overline{\bar{x}y} \vee xy = \bar{\bar{x}} \vee \bar{y} \vee xy = x \vee \bar{y} \vee xy = x \vee xy \vee \bar{y} = x \vee \bar{y}$.

Замечание. Пусть M – некоторое множество и $\mathcal{P}(M)$ – множество всех подмножеств множества M. Если A, B, C – произвольные подмножества из M, и $\bar{A}$ интерпретируется как $M - A$

(т.е. $\bar{A}$ – дополнение A до M), $A \,\&\, B$ как $A \cap B$, $A \vee B$ как $A \cup B$, 0 как пустое множество $\varnothing$, а 1 есть все множество M, то при таком теоретико-множественном понимании операций –, &, $\vee$ булевы свойства 1 – 9 останутся справедливыми.

Множество M, в котором определены операции $\neg$, &, $\vee$ и константы 0 и 1, удовлетворяющие аксиомам 1 – 9, называется *булевой алгеброй*. Обозначим булеву алгебру через $(M,\&,\vee,-,0,1)$. Тогда системы $(\{0,1\},\&,\vee,\neg,0,1)$ и $(\mathcal{P}(M),\cap,\cup,\neg,\varnothing,M)$ являются булевыми алгебрами.

Множество M, в котором определены две операции & и $\vee$, удовлетворяющие аксиомам 1 – 4, называется *решеткой*. Решетка *дистрибутивна*, если дополнительно выполняется аксиома 5 дистрибутивности.

Пусть множество $M' = \{0,1,2,01,02,12,012\}$. Элемент 012 понимаем как 0&1&2. Аналогично другие элементы из M'. Множество M состоит из множества M' и всех дизъюнкций попарно различных элементов множества M'. При этом ни одно дизъюнктивное слагаемое, рассматриваемое как множество своих сомножителей, не содержится в другом его дизъюнктивном слагаемом. Например, элементом множества M является дизъюнкция 01 $\vee$ 12 $\vee$ 02. Множество M с операциями & и $\vee$, удовлетворяющими аксиомам 1 – 5, образуют (свободную) дистрибутивную решетку с образующими 0,1,2. Приведем пример преобразований в такой решетке. Решеточное выражение (0$\vee$2)(01$\vee$12$\vee$012) = 001 $\vee$ 012 $\vee$ 0012 $\vee$ 201 $\vee$ 212 $\vee$ 2012 = 01 $\vee$ 012 $\vee$ 012 $\vee$ 012 $\vee$ 12 $\vee$ 012 = 01 $\vee$ 12.

13.4. Нормальные формы

Элементарной конъюнкцией называется конъюнкция, составленная из попарно различных переменных или отрицаний переменных.

Пример. x, y, xy, $\bar{x}_1 x_2 \bar{x}_3$.

Дизъюнктивной нормальной формой (ДНФ) называется дизъюнкция попарно различных элементарных конъюнкций.

Пример. xy, $xy \vee x$, $x_1 \bar{x}_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3$.

Элементарной дизъюнкцией называется дизъюнкция, составленная из попарно различных переменных или отрицаний переменных.

Пример. x, $x \vee \bar{y}$, $\bar{x} \vee y \vee \bar{z}$.

Конъюнктивной нормальной формой (КНФ) называется конъюнкция попарно различных элементарных дизъюнкций.

Пример. x, $x \vee \bar{y}$, $(x \vee y)(x \vee \bar{y} \vee z)(\bar{x} \vee y \vee \bar{z})$.

Введем следующие обозначения:

$$x^c = \begin{cases} x, & \text{если } c = 1; \\ \bar{x}, & \text{если } c = 0; \end{cases}$$

$$\&_{i=1}^{n} A_i = A_1 \& A_2 \& \ldots \& A_n; \quad \vee_{i=1}^{n} A_i = A_1 \vee A_2 \vee \ldots \vee A_n.$$

Напомним, что знак $\leftrightarrow$ означает "тогда и только тогда". Заметим, что $x^c = 1 \leftrightarrow x = c$;

$$x_1^{c_1} \& \; x_2^{c_2} \& \; \ldots \; \& x_n^{c_n} = 1 \leftrightarrow x_1 = c_1, \; x_2 = c_2, \; \ldots, \; x_n = c_n.$$

Лемма (о разложении функции по компонентам). Всякая функция алгебры логики допускает представление:

$$f(x_1, \ldots, x_n) = \bigvee_{(c_1, \ldots, c_k)} f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n)\, x_1^{c_1} \ldots x_k^{c_k}, \tag{13.1}$$

где дизъюнкция берется по всем наборам $(c_1, \ldots, c_k)$ из 0 и 1.

Доказательство. Пусть левая часть (13.1) равна 1. Тогда для всякого набора $x_1, \ldots, x_k, x_{k+1}, \ldots, x_n$ длины n из 0 и 1 последовательно получаем следующее:

$$f(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = 1 \leftrightarrow$$

$$f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n) = 1, \; c_1 = x_1, \ldots, c_k = x_k \leftrightarrow$$

$$f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n) = 1, \; x_1^{c_1} = 1, \ldots, x_k^{c_k} = 1 \leftrightarrow$$

$$f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n) = 1, \; x_1^{c_1} \ldots x_k^{c_k} = 1 \leftrightarrow$$

$$f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n)\, x_1^{c_1} \ldots x_k^{c_k} = 1 \leftrightarrow$$

$$\bigvee_{(c_1, \ldots, c_k)} f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n) x_1^{c_1} \ldots x_k^{c_k} = 1.$$

Равенство (13.1) доказано. Лемма установлена.

Замечание. Функция $f(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n)$ называется *компонентой функции* $f(x_1, \ldots, x_n)$.

13.4.1. Совершенные нормальные формы

Теорема (О СДНФ). Всякая не равная тождественному нулю функция $f(x_1, \ldots, x_n)$ допускает представление

$$f(x_1, \ldots, x_n) = \bigvee_{f(c_1, \ldots, c_n) = 1} x_1^{c_1} \ldots x_n^{c_n}, \tag{13.2}$$

где дизъюнкция берется по всем наборам $\mathbf{c} = (c_1, \ldots, c_n)$ из 0 и 1, для которых $f(\mathbf{c}) = 1$.

Доказательство. Пусть $f(x_1, \ldots, x_n) \not\equiv 0$. Согласно лемме о разложении функции по компонентам при $k = n$ получаем

$$f(x_1, \ldots, x_n) = \bigvee_{(c_1, \ldots, c_n)} f(c_1, \ldots, c_n)\, x_1^{c_1} \ldots x_n^{c_n}.$$

Из правой части равенства удалим все нулевые дизъюнктивные члены, в которых $f(c_1, \ldots, c_n) = 0$. Тогда получим

$$f(x_1, \ldots, x_n) = \bigvee_{f(\mathbf{c}) = 1} f(c_1, \ldots, c_n)\, x_1^{c_1} \ldots x_n^{c_n}.$$

Так как в этой формуле в любом дизъюнктивном члене элемент $f(c_1, \ldots, c_n) = 1$, то она принимает вид

$$f(x_1, \ldots, x_n) = \bigvee_{f(\mathbf{c}) = 1} x_1^{c_1} \ldots x_n^{c_n}.$$

Теорема доказана.

Определение. Правая часть представления (13.2) называется *совершенной дизъюнктивной нормальной формой* (СДНФ) функции f.

Каждое слагаемое в СДНФ называется *конституентой единицы*.

Конституента единицы $x_1^{c_1} x_2^{c_2} \ldots x_n^{c_n} = 1$ на единственном наборе $x_1 = c_1, \; x_2 = c_2, \ldots, x_n = c_n$.

Пример. СДНФ для функции f, приведенной в табл.13.5, имеет вид: $f(x,y,z) = \bar{x}\bar{y}\bar{z} \vee \bar{x}yz \vee x\bar{y}z \vee xyz$.

Дизъюнктивные слагаемые $\bar{x}\bar{y}\bar{z}$, $\bar{x}yz$, $x\bar{y}z$, xyz являются конституентами единицы.

Замечание. Всякую функцию алгебры логики можно реализовать формулой, построенной с помощью конъюнкции, дизъюнкции и отрицания. Поэтому множество функций $F = \{\&, \vee, \neg\}$ составля-

Таблица 13.5

x	0	0	0	0	1	1	1	1
y	0	0	1	1	0	0	1	1
z	0	1	0	1	0	1	0	1
f	1	0	0	1	0	1	0	1

ет полную систему. Так как $x \vee y = \overline{\bar{x} \,\&\, \bar{y}}$, то система функций $F = \{\&, \neg\}$ полна. Так как $x \,\&\, y = \overline{\bar{x} \vee \bar{y}}$, то система $F = \{\vee, \neg\}$ полна. Система $F = \{x \mid y\}$, состоящая из единственной функции – штриха Шеффера, полна, ибо $\bar{x} = x \mid x$, а $x \,\&\, y = (x \mid y) \mid (x \mid y)$. Система $F = \{x \uparrow y\}$, состоящая только из стрелки Пирса, полна, так как $\bar{x} = x \uparrow x$, а $x \vee y = (x \uparrow y) \uparrow (x \uparrow y)$. Очевидно, что $x \uparrow y = \overline{x \vee y}$.

Теорема (о единственности СДНФ). Для всякой функции, не равной тождественному нулю, существует единственная СДНФ.

Доказательство. Существование СДНФ для функции $f \not\equiv 0$ вытекает из предыдущей теоремы. Покажем, что эта СДНФ единственная. В самом деле, имеется $2^{2^n}-1$ n-местных функций, не равных нулю тождественно. Подсчитаем число различных СДНФ от n переменных. Пусть C_n^k означает число сочетаний из n элементов по k. Тогда число одночленных СДНФ $x_1^{c_1}x_2^{c_2}\ldots x_n^{c_n}$ равно $C_2^1 n$. Число k-членных СДНФ равно $C_2^k n$. Число n-членных СДНФ равно $C_2^n n$. Число всех различных СДНФ

$$C_2^1 n + C_2^2 n + \ldots + C_2^k n + \ldots + C_2^n n = 2^{2^n} - 1.$$

Итак, $2^{2^n}-1$ функций реализуются посредством $2^{2^n}-1$ СДНФ, т.е. каждая функция реализуется единственной СДНФ.

Теорема (о СКНФ). Всякая не равная тождественной единице функция $f(x_1,\ldots,x_n)$ допускает представление

$$f(x_1,\ldots,x_n) = \mathop{\&}_{f(c_1,\ldots,c_n)=0} (x_1^{\bar{c}_1} \vee \ldots \vee x_n^{\bar{c}_n}), \qquad (13.3)$$

где конъюнкция берется по всем наборам $\mathbf{c} = (c_1,\ldots,c_n)$ из 0 и 1, для которых $f(\mathbf{c}) = 0$.

Доказательство. Заметим, что $\overline{x^c} = x^{\bar{c}}$. Пусть функция $f(x_1, \ldots,x_n) \not\equiv 1$, тогда $\bar{f} \not\equiv 0$ и потому функция $\bar{f}$ допускает представление в виде СДНФ $\bar{f}(x_1,\ldots,x_n) = \mathop{\vee}_{\bar{f}(\mathbf{c})=1} x_1^{c_1} \ldots x_n^{c_n}$.

Отсюда

$$f(x_1,\ldots,x_n) = \overline{\mathop{\vee}_{f(\mathbf{c})=0} x_1^{c_1}\ldots x_n^{c_n}} =$$

$$\mathop{\&}_{f(\mathbf{c})=0} (\overline{x_1^{c_1}} \vee \ldots \vee \overline{x_n^{c_n}}) = \mathop{\&}_{f(\mathbf{c})=0} (x_1^{\bar{c}_1} \vee \ldots \vee x_n^{\bar{c}_n}).$$

Теорема доказана.

Правая часть в представлении (13.3) называется *совершенной конъюнктивной нормальной формой* (СКНФ) функции f.

Пример. КНФ для функции, приведенной в табл.13.6, имеет вид: $f(x,y,z) = (x \vee y \vee z)(x \vee \bar{y} \vee \bar{z})(\bar{x} \vee y \vee \bar{z})(\bar{x} \vee \bar{y} \vee \bar{z})$.

13.5. Минимизация нормальных форм

Минимальной ДНФ (МДНФ) функции $f(x_1,\ldots,x_n)$ называется ДНФ, реализующая функцию f и содержащая минимальное число символов переменных по сравнению со всеми другими ДНФ, реализующими функцию f.

Если для всякого набора $a = (a_1,\ldots,a_n)$ значений переменных условие $g(a) = 1$ влечет $f(a) = 1$, то функция g называется *частью* функции f (или функция f *накрывает* функцию g). Если при этом для некоторого набора $c = (c_1,\ldots,c_n)$ функция $g(c) = 1$, то говорят, что функция g накрывает единицу функции f на наборе c (или что g накрывает конституенту единицы $x_1^{c_1}\ldots x_n^{c_n}$ функции f). Заметим, что конституента единицы функции f есть часть функции f, накрывающая только одну единицу функции f.

Элементарная конъюнкция K называется *импликантом* функции f, если для всякого набора $a = (a_1,\ldots,a_n)$ из 0 и 1 условие $K(a) = 1$ влечет $f(a) = 1$.

Импликант K функции f называется *простым*, если выражение,

Таблица 13.6

x	0	0	0	0	1	1	1	1
y	0	0	1	1	0	0	1	1
z	0	1	0	1	0	1	0	1
f	0	1	1	0	1	0	1	0

получающееся из него удалением любых множителей, уже не есть импликант функции f.

Всякий импликант функции f есть часть функции f.

Теорема. Всякая функция реализуется дизъюнкцией всех своих простых импликант (ПИ).

Доказательство. Пусть $f(x_1,\ldots,x_n)$ есть функция, а $A = K_1 \vee \ldots \vee K_m$ – дизъюнкция всех ее простых импликант. Пусть $\mathbf{a} = (a_1,\ldots,a_n)$ – произвольный набор длины n из 0 и 1.

Если $A(\mathbf{a}) = 1$, то найдется дизъюнктивное слагаемое $K_i(\mathbf{a}) = 1$, что влечет $f(\mathbf{a}) = 1$, ибо K_i есть импликант функции f.

Если $f(\mathbf{a}) = 1$, то в СДНФ для функции f найдется элементарная конъюнкция K, равная на этом наборе единице. Один из простых импликантов K_j функции f получается удалением некоторых множителей из K и потому $K_j(\mathbf{a}) = 1$, а тогда $A(\mathbf{a}) = 1$.

Следовательно, $f = A$. Теорема доказана.

Сокращенная ДНФ функции f есть дизъюнкция всех простых импликант функции f. Всякая функция f реализуется своей сокращенной ДНФ. Для всякой функции, не равной тождественно нулю, существует единственная сокращенная ДНФ.

Пусть A и B – произвольные формулы. Из свойств булевых операций вытекают следующие обратимые правила преобразования ДНФ:

1) $\dfrac{A\cdot B \vee A\cdot\overline{B}}{A}$ – полное склеивание (развертывание);

2) $\dfrac{A\cdot B \vee A\cdot\overline{B}}{A \vee A\cdot B \vee A\cdot\overline{B}}$ – неполное склеивание;

3) $\dfrac{A \vee A\cdot B}{A}$ – поглощение;

4) $\dfrac{A \vee A}{A}$; $\dfrac{A \,\&\, A}{A}$ – идемпотентность (удаление дублирующих членов).

Теорема (Куайна). Если в СДНФ функции f провести все операции неполного склеивания, а затем все операции поглощения и удаления дублирующих членов, то в результате получится сокращенная ДНФ функции f.

Доказательство можно провести по следующему плану. Пусть имеем сокращенную ДНФ функции f. Проведем все операции развертывания к каждому простому импликанту для получения недостающих переменных в каждом дизъюнктивном слагаемом сокращенной ДНФ. В полученном выражении из нескольких одинаковых дизъюнктивных слагаемых оставим только по одному экземпляру. В результате получим СДНФ функции f. Теперь, исходя из полученной СДНФ, в обратном порядке проведем операции добавления одинаковых дизъюнктивных слагаемых (с помощью правил идемпотентности), неполного склеивания и поглощения. В итоге получим исходную сокращенную ДНФ.

13.5.1. Алгоритм Куайна построения сокращенной ДНФ

1. Получить СДНФ функции f.
2. Провести все операции неполного склеивания.
3. Провести все операции поглощения.

Пример. Построим сокращенную ДНФ для функции, приведенной в табл.13.7.

1. Строим СДНФ функции f:

$$f(x,y,z,t) = \overline{x}\overline{y}\overline{z}\overline{t} \vee \overline{x}\overline{y}\overline{z}t \vee \overline{x}\overline{y}z\overline{t} \vee \overline{x}\overline{y}zt \vee \overline{x}y\overline{z}t \vee x\overline{y}\overline{z}\overline{t} \vee x\overline{y}zt \vee xy\overline{z}\overline{t} \vee xy\overline{z}t \vee xyz\overline{t} \vee xyzt.$$

Занумеруем дизъюнктивные члены в полученной СДНФ в порядке от 1 до 11.

2. Проводим все операции неполного склеивания.

Первый этап склеиваний:

Слагаемые	Склеивание по	Результат
1,2	t	$\overline{x}\overline{y}\overline{z}$
1,3	z	$\overline{x}\overline{y}\overline{t}$

Таблица 13.7

x	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
y	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
z	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
t	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
f	1	1	1	1	0	1	0	0	1	0	1	0	1	1	1	1

1,6	x	$\bar{y}\bar{z}\bar{t}$
2,4	z	$\bar{x}\bar{y}t$
2,5	y	$\bar{x}\bar{z}t$
3,4	t	$\bar{x}\bar{y}z$
3,7	x	$\bar{y}z\bar{t}$
5,9	x	$y\bar{z}t$
6,7	z	$x\bar{y}\bar{t}$
6,8	y	$x\bar{z}\bar{t}$
7,10	y	$xz\bar{t}$
8,9	t	$xy\bar{z}$
8,10	z	$xy\bar{t}$
9,11	z	xyt
10,11	t	xyz

После первого этапа склеиваний (и возможных поглощений) получаем, что $f(x,y,z,t) = \bar{x}\bar{y}\bar{z} \vee \bar{x}\bar{y}\bar{t} \vee \bar{y}\bar{z}\bar{t} \vee \bar{x}\bar{y}t \vee \bar{x}\bar{z}t \vee \bar{x}\bar{y}z \vee \bar{y}z\bar{t} \vee x\bar{y}\bar{t} \vee y\bar{z}t \vee x\bar{z}\bar{t} \vee xz\bar{t} \vee xy\bar{z} \vee xy\bar{t} \vee xyt \vee xyz$. Пронумеруем дизъюнктивные члены в полученной ДНФ в порядке их следования от 1 до 15.

Второй этап склеиваний:

Слагаемые	Склеивание по	Результат
1,6	z	$\bar{x}\bar{y}$
2,4	t	$\bar{x}\bar{y}$
2,8	x	$\bar{y}\bar{t}$
3,7	z	$\bar{y}\bar{t}$
8,13	y	$x\bar{t}$
10,11	z	$x\bar{t}$
12,15	z	xy
13,14	t	xy

После второго этапа склеиваний и последующих поглощений получаем, что $f(x,y,z,t) = xy \vee \bar{x}\bar{y} \vee \bar{y}\bar{t} \vee x\bar{t} \vee \bar{x}\bar{z}t \vee y\bar{z}t$.

Это и будет сокращенной ДНФ для функции f, ибо дальнейшие склеивания невозможны.

13.5.2. Алгоритм построения сокращенной ДНФ с помощью КНФ

Пусть $f(x_1,\ldots,x_n)$ есть некоторая функция алгебры логики. Построим для f некоторую КНФ. Осуществим далее следующие преобразования.

1. В КНФ раскроем скобки и удалим дублирующие члены согласно равенствам $K \cdot K = K$, $K \vee K = K$; удалим дизъюнктивные слагаемые, содержащие одновременно переменную и ее отрицание. В результате получим дизъюнкцию конъюнкций, каждая из которых содержит только по одному элементу из каждой скобки КНФ.

2. В полученном выражении проведем все поглощения (согласно равенству $A \vee AB = A$), а затем удалим дублирующие члены.

В результате проведенных операций получим сокращенную ДНФ функции f. Покажем это.

Для каждой элементарной дизъюнкции D в КНФ и каждой элементарной конъюнкции K в сокращенной ДНФ (сокр.ДНФ) существует некоторый множитель вида x^a из K, содержащийся в D т.е.

$$\forall D \in \text{ДНФ}\ \forall K \in \text{сокр.ДНФ}\ \exists x^a \in K\ (x^a \in D).$$

Допустим противное: в КНФ существует элементарная конъюнкция D, в сокращенной ДНФ существует элементарная конъюнкция K, для которой всякий множитель вида x^a из K не входит в D. Не уменьшая общности, возьмем для простоты

$$K = x_1^{a_1} x_2^{a_2} \ldots x_k^{a_k}, \quad D = x_{k+1}^{a_{k+1}} \vee \ldots \vee x_r^{a_r}.$$

Положим $x_1 = a_1,\ldots,x_k = a_k$, $x_{k+1} = c_{k+1} \neq a_{k+1}$,..., $x_r = c_r \neq a_r$. Тогда $K(a_1,\ldots,a_k) = 1$ и потому $f(a_1,\ldots,a_k,c_{k+1},\ldots,c_r) = 1$. С другой стороны, $D(c_{k+1},\ldots,c_r) = 0$ и потому $f(a_1,\ldots,a_k,c_{k+1},\ldots,c_r) = 0$. Противоречие.

Пусть по-прежнему для простоты произвольный простой импликант K из сокращенной ДНФ равен $x_1^{a_1}\&x_2^{a_2}\&\ldots\&x_k^{a_k}$. Тогда элементы $x_1^{a_1},x_2^{a_2},\ldots,x_k^{a_k}$ попадут в не менее чем k скобок в КНФ. Если допустим, что этого нет, то при перемножении скобок из КНФ не получим дизъюнктивного слагаемого, который содержал бы множители $x_1^{a_1},x_2^{a_2},\ldots,x_k^{a_k}$, а потому, строя из ре-

зультата перемножения сокращенную ДНФ вычеркиванием лишних сомножителей, не получим простого импликанта K.

Так как $x_1^{a_1}, x_2^{a_2}, \ldots, x_k^{a_k}$ содержатся в k разных скобках КНФ, а всякая другая скобка, отличная от указанных k скобок, содержит хотя бы один элемент вида x^a из K, то при раскрытии скобок имеем простой импликант K. После проведения всех операций поглощения и удаления дублирующих множителей, останутся только простые импликанты из сокращенной ДНФ, ибо если предположить наличие в результате хотя бы одного дизъюнктивного слагаемого, отличного от всех простых импликантов сокращенной ДНФ, то можно подобрать такие значения переменных функции f, на которых все простые импликанты примут значение 0, а это дополнительное слагаемое – значение 1, чего быть не может.

Пример. Построим сокращенную ДНФ этим способом для функции f = 1111010010101111 из предыдущего примера:

$$f(x,y,z,t) = (x \vee \bar{y} \vee z \vee t)(x \vee \bar{y} \vee \bar{z} \vee t)(x \vee \bar{y} \vee z \vee \bar{t}) \,\&$$
$$(\bar{x} \vee y \vee z \vee \bar{t})(\bar{x} \vee y \vee \bar{z} \vee \bar{t}) = (x \vee \bar{y} \vee t)(x \vee \bar{y} \vee \bar{z} \vee \bar{t})(\bar{x} \vee y \vee \bar{t}) =$$
$$(x \vee \bar{y} \vee \bar{y} \vee t \vee \bar{z} \vee t)(\bar{x} \vee y \vee \bar{t}) = xy \vee x\bar{t} \vee \bar{x}\bar{y} \vee \bar{y}\bar{t} \vee \bar{x}\bar{y}t \vee \bar{x}\bar{z}t \vee y\bar{z}t.$$

Сокращенная ДНФ для функции

$$f(x,y,z,t) = xy \vee x\bar{t} \vee \bar{x}\bar{y} \vee \bar{y}\bar{t} \vee \bar{x}\bar{z}t \vee y\bar{z}t,$$

что совпадает с результатом предыдущего примера.

Тупиковой ДНФ (ТДНФ) функции f называется такая ДНФ ее простых импликант, из которой нельзя удалить ни одного импликанта, не изменив функции f.

Теорема. Всякая минимальная ДНФ некоторой функции является ее тупиковой ДНФ.

Доказательство. В МДНФ входят только простые импликанты, иначе некоторые множители в непростом импликанте можно удалить в противоречие с минимальностью исходной ДНФ. В МДНФ нет лишних импликант, иначе исходная ДНФ не является минимальной.

Вывод. Для получения МДНФ функции f необходимо построить все ТДНФ функции f и выбрать те из них, которые содержат минимальное число букв.

13.5.3. Построение всех тупиковых ДНФ

Пусть $f(x_1, \ldots, x_n)$ есть функция алгебры логики.

1. Построим СДНФ функции f и пусть $P_1, P_2, \ldots, P_s$ есть ее

конституенты (единицы).

2. Построим сокращенную ДНФ функции f и пусть $K_1, K_2, \ldots, K_m$ есть ее простые импликанты.

3. Построим матрицу покрытий простых импликант функции f ее конституентами единицы (табл.13.8), полагая, что

$$a_{ij} = \begin{cases} 1, & \text{если каждый множитель в } K_i \text{ является} \\ & \text{множителем в } P_j;\ (P_j \text{ есть часть для } K_i); \\ 0 & \text{в противном случае.} \end{cases}$$

4. Для каждого столбца j $(1 \leqslant j \leqslant s)$ найдем множество E_j всех тех номеров i строк, для которых $a_{ij} = 1$. Пусть $E_j = \{e_{j_1}, e_{j_2}, \ldots, e_{j,r_j}\}$. Составим выражение $A = \bigwedge_{j=1}^{s} (e_{j_1} \vee e_{j_2} \vee \ldots \vee e_{j,r_j})$. Назовем его решеточным выражением. Это выражение можно рассматривать как формулу, построенную в свободной дистрибутивной решетке с образующими $1,2,\ldots,m$ и с операциями & и $\vee$.

5. В выражении A раскроем скобки, приведя выражение A к равносильному выражению $B = \bigvee_i e_{i_1} \& e_{i_2} \& \ldots \& e_{i_s}$, где перечислены все конъюнкции $e_{i_1} \& e_{i_2} \& \ldots \& e_{i_s}$, элементы $e_{i_1}, e_{i_2}, \ldots, e_{i_s}$ которой взяты из скобок $1,2,\ldots,s$ соответственно в выражении A.

6. В выражении B проведем все операции удаления дублирующих членов и все операции поглощения. В результате получим равносильное выражение C, представляющее собой дизъюнкцию элементарных конъюнкций.

Таблица 13.8

N	P_1	P_2	$\ldots$	P_j	$\ldots$	P_s
K_1	a_{11}	a_{12}	$\ldots$	a_{1j}	$\ldots$	a_{1s}
K_2	a_{21}	a_{22}	$\ldots$	a_{2j}	$\ldots$	a_{2s}
			$\ldots$		$\ldots$	
K_i	a_{i1}	a_{i2}	$\ldots$	a_{ij}	$\ldots$	a_{is}
			$\ldots$		$\ldots$	
K_m	a_{m1}	a_{m2}	$\ldots$	a_{mj}	$\ldots$	a_{ms}

Утверждение. Каждая элементарная конъюнкция $i_1\&i_2\&\ldots\&i_r$ в C дает ТДНФ $K_{i1} \vee K_{i2} \vee \ldots \vee K_{ir}$ для f. Все ТДНФ для функции f исчерпываются элементарными конъюнкциями в выражении C.

Пример. Сокращенная ДНФ для функции f = 1111010010101111 имеет вид $f = xy \vee \bar{x}\bar{y} \vee \bar{y}\bar{t} \vee x\bar{t} \vee \bar{x}\bar{z}t \vee y\bar{z}t$.

Для функции f построим все минимальные ДНФ.

1. Строим матрицу покрытий (табл.13.9).
2. Строим решеточное выражение (по столбцам табл.13.9).
$E = (2\vee3)(2\vee5)(2\vee3)2(5\vee6)(3\vee4)(3\vee4)(1\vee4)(1\vee6)\&$
$(1\vee4)(1) = (2\vee3)(2\vee5)(5\vee6)(3\vee4)(1\vee4)(1\vee6)12 =$
$(5\vee6)(3\vee4)(1)(2) = 1235 \vee 1245 \vee 1236 \vee 1246.$
3. Строим все тупиковые ДНФ функции f.

$f = xy \vee \bar{x}\bar{y} \vee \bar{y}\bar{t} \vee \bar{x}\bar{z}t$, простые импликанты 1,2,3,5;

$f = xy \vee \bar{x}\bar{y} \vee x\bar{t} \vee \bar{x}\bar{z}t$, простые импликанты 1,2,4,5;

$f = xy \vee \bar{x}\bar{y} \vee \bar{y}\bar{t} \vee y\bar{z}t$, простые импликанты 1,2,3,6;

$f = xy \vee \bar{x}\bar{y} \vee x\bar{t} \vee y\bar{z}t$, простые импликанты 1,2,4,6.

Таблица 13.9

N	ПИ	Конституенты единицы функции f										
		$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$	x	x	x	x	x	x
		$\bar{y}$	$\bar{y}$	$\bar{y}$	$\bar{y}$	y	$\bar{y}$	$\bar{y}$	y	y	y	y
		$\bar{z}$	$\bar{z}$	z	z	$\bar{z}$	$\bar{z}$	z	$\bar{z}$	$\bar{z}$	z	z
		$\bar{t}$	t	$\bar{t}$	t	t	$\bar{t}$	$\bar{t}$	$\bar{t}$	t	$\bar{t}$	t
1	xy								+	+	+	+
2	$\bar{x}\bar{y}$	+	+	+	+							
3	$\bar{y}\bar{t}$	+		+			+	+				
4	$x\bar{t}$						+	+	+		+	
5	$\bar{x}\bar{z}t$		+			+						
6	$y\bar{z}t$					+				+		

4. Все найденные ТДНФ являются минимальными ДНФ.

13.5.4. Алгоритм минимизации функций в классе ДНФ

1. Строим СДНФ функции f.
2. Строим сокращенную ДНФ функции f.
3. С помощью матрицы покрытий и решеточного выражения строим все ТДНФ функции f.
4. Среди построенных ТДНФ выбираем все минимальные дизъюнктивные нормальные формы функции f.

13.5.5. Алгоритм минимизации функций в классе КНФ

Чтобы построить все минимальные КНФ (МКНФ) функции f, следует построить все МДНФ функции $\bar{f}$ и взять от каждой из них отрицание, для чего заменить знаки & на $\vee$, а $\vee$ на & (сохранив первоначальное распределение скобок) и над каждой буквой поставить знак отрицания. Полученные КНФ для функции f будут минимальными. В самом деле, если бы для f существовала КНФ с меньшим числом букв, то ее отрицание дало бы для $\bar{f}$ ДНФ с меньшим числом букв, чем в любой из минимальных ДНФ для $\bar{f}$. Противоречие с их минимальностью.

13.5.6. Алгоритм минимизации функций в классе нормальных форм

Пусть f – функция алгебры логики.

1. Строим все МДНФ функции f.
2. Строим все МКНФ функции f.
3. Из построенных минимальных форм выбираем простейшие (по числу букв).

Пример. В классе нормальных форм минимизировать функцию f = 01011110.

1. Строим СДНФ для функции f:
$$f(x,y,z) = \bar{x}\bar{y}z \vee \bar{x}yz \vee x\bar{y}\bar{z} \vee x\bar{y}z \vee xy\bar{z}.$$
2. Строим сокращенную ДНФ функции f:
$$f(x,y,z) = (x\vee y\vee z)(x\vee\bar{y}\vee z)(\bar{x}\vee\bar{y}\vee\bar{z}) =$$
$$(x\vee x\bar{y}\vee x\bar{z}\vee xy\vee y\bar{y}\vee yz\vee xz\vee\bar{y}\bar{z}\vee z\bar{z})(\bar{x}\vee\bar{y}\vee\bar{z}) = (x\vee z)(\bar{x}\vee\bar{y}\vee\bar{z}) =$$
$$x\bar{x} \vee x\bar{y} \vee x\bar{z} \vee \bar{x}z \vee \bar{y}z \vee z\bar{z} = \bar{x}z \vee \bar{y}z \vee x\bar{y} \vee x\bar{z}.$$
3. Строим матрицу покрытий (табл.13.10).

Решеточное выражение $E = (1 \vee 2)1(3 \vee 4)(2 \vee 3)4 = 134 \vee 124$.

4. Строим все тупиковые ДНФ функции f:

$f(x,y,z) = \bar{x}z \vee x\bar{y} \vee x\bar{z}$; $f(x,y,z) = \bar{x}z \vee \bar{y}z \vee x\bar{z}$.

5. Обе построенные ТДНФ являются минимальными.

6. Повторяем эти этапы для функции $\bar{f}$.

СДНФ: $\bar{f}(x,y,z) = \bar{x}\bar{y}\bar{z} \vee \bar{x}y\bar{z} \vee xyz$.

Сокращенная ДНФ: $\bar{f}(x,y,z) =$

$(x \vee y \vee \bar{z})(x \vee \bar{y} \vee \bar{z})(\bar{x} \vee y \vee z)(\bar{x} \vee y \vee \bar{z})(\bar{x} \vee \bar{y} \vee z) = (x \vee \bar{z})(\bar{x} \vee y)(\bar{x} \vee \bar{y} \vee z) =$

$(x \vee \bar{z})(\bar{x} \vee yz) = xyz \vee \bar{x}\bar{z}$.

Строим матрицу покрытий (табл.13.11).

Решеточный многочлен $E = 112 = 12$. Единственная тупиковая ДНФ (она же минимальная) для функции $\bar{f}(x,y,z) = \bar{x}\bar{z} \vee xyz$.

Минимальная КНФ функции $f(x,y,z) = (x \vee z)(\bar{x} \vee \bar{y} \vee \bar{z})$. Из построенных МДНФ и МКНФ выбираем простейшую:

$f(x,y,z) = (x \vee z)(\bar{x} \vee \bar{y} \vee \bar{z})$.

Пример. В классе нормальных форм минимизировать функцию f = 11011011.

1. СДНФ: $f(x,y,z) = \bar{x}\bar{y}\bar{z} \vee \bar{x}\bar{y}z \vee \bar{x}yz \vee x\bar{y}\bar{z} \vee xy\bar{z} \vee xyz$.

2. Сокращенная ДНФ: $f(x,y,z) = (x \vee \bar{y} \vee z)(\bar{x} \vee y \vee \bar{z}) =$

$xy \vee x\bar{z} \vee \bar{y}\bar{z} \vee \bar{x}z \vee yz \vee \bar{x}\bar{y}$.

3. Строим матрицу покрытий (табл.13.12).

$E = (3 \vee 6)(4 \vee 6)(4 \vee 5)(2 \vee 3)(1 \vee 2)(1 \vee 5) = 1246 \vee 1356 \vee 134 \vee 256 \vee 2345$.

Таблица 13.10

N	ПИ	$\bar{x}\bar{y}z$	$\bar{x}yz$	$x\bar{y}\bar{z}$	$x\bar{y}z$	$xy\bar{z}$
1	$\bar{x}z$	+	+			
2	$\bar{y}z$	+			+	
3	$x\bar{y}$			+	+	
4	$x\bar{z}$			+		+

Таблица 13.11

N	ПИ	$\bar{x}\bar{y}\bar{z}$	$\bar{x}y\bar{z}$	xyz
1	$\bar{x}\bar{z}$	+	+	
2	xyz			+

Таблица 13.12

N	ПИ	$\bar{x}\bar{y}\bar{z}$	$\bar{x}\bar{y}z$	$\bar{x}yz$	$x\bar{y}\bar{z}$	$xy\bar{z}$	xyz
1	xy					+	+
2	$x\bar{z}$				+	+	
3	$\bar{y}\bar{z}$	+			+		
4	$\bar{x}z$		+	+			
5	yz			+			+
6	$\bar{x}\bar{y}$	+	+				

4. Тупиковые ДНФ функции f:

$f(x,y,z) = xy \vee x\bar{z} \vee \bar{x}z \vee \bar{x}\bar{y}$;

$f(x,y,z) = xy \vee \bar{y}\bar{z} \vee yz \vee \bar{x}\bar{y}$;

$f(x,y,z) = xy \vee \bar{y}\bar{z} \vee \bar{x}z$;

$f(x,y,z) = x\bar{z} \vee yz \vee \bar{x}\bar{y}$;

$f(x,y,z) = x\bar{z} \vee \bar{y}\bar{z} \vee \bar{x}z \vee yz$.

5. Минимальные ДНФ функции f:

$f(x,y,z) = xy \vee \bar{y}\bar{z} \vee \bar{x}z$; $f(x,y,z) = x\bar{z} \vee yz \vee \bar{x}\bar{z}$.

6. Повторяем указанные выше этапы для функции $\bar{f}$.

СДНФ: $\bar{f}(x,y,z) = \bar{x}y\bar{z} \vee x\bar{y}z$.

Сокращенная ДНФ: $\bar{f}(x,y,z) = (x \vee y \vee z)(x \vee y \vee \bar{z})$ &

$(x \vee \bar{y} \vee \bar{z})(\bar{x} \vee y \vee z)(\bar{x} \vee \bar{y} \vee z)(\bar{x} \vee \bar{y} \vee \bar{z}) = (x \vee y)(x \vee \bar{y} \vee \bar{z})(\bar{x} \vee y \vee z)(\bar{x} \vee \bar{y}) =$

$(x \vee y\bar{z})(\bar{x} \vee \bar{y}z) = \bar{x}y\bar{z} \vee x\bar{y}z$.

Построенная сокращенная ДНФ функции $\bar{f}$ является для нее тупиковой и минимальной.

Минимальная КНФ функции $f(x,y,z) = (x \vee \bar{y} \vee z)(\bar{x} \vee y \vee \bar{z})$.

Построенные МДНФ и МКНФ имеют одно и то же число букв; все они составляют минимальные формы для f:

$f(x,y,z) = xy \vee \bar{y}\bar{z} \vee \bar{x}z$;

$f(x,y,z) = x\bar{z} \vee yz \vee \bar{x}\bar{z}$;

$f(x,y,z) = (x \vee \bar{y} \vee z)(\bar{x} \vee y \vee \bar{z})$.

13.6. Минимизация частично определенных функций

Пусть функция $f(x_1,\ldots,x_n)$ частично (не всюду) определена. Если f не определена на p наборах из 0 и 1, то существует 2^p возможностей для доопределения функции f. Полностью определенная функция $g(x_1,\ldots,x_n)$ есть доопределение функции f, если g совпадает с f на тех наборах из 0 и 1, на которых f определена.

Задача минимизации частично определенной функции f сводится к отысканию такого доопределения g функции f, которое имеет простейшую (по числу букв) минимальную форму.

Обозначим через $f_0(x_1,\ldots,x_n)$ и $f_1(x_1,\ldots,x_n)$ доопределения нулями и единицами соответственно частично определенной функции $f(x_1,\ldots,x_n)$.

Теорема. Минимальная ДНФ частично определенной функции $f(x_1,\ldots,x_n)$ есть минимальная по числу букв дизъюнкция импликант в сокращенной ДНФ доопределения $f_1(x_1,\ldots,x_n)$, которые в совокупности накрывают все конституенты единицы доопределения $f_0(x_1,\ldots,x_n)$.

Доказательство. Рассмотрим СДНФ некоторого доопределения $g(x_1,\ldots,x_n)$ функции $f(x_1,\ldots,x_n)$. Конституенты единицы, входящие в эту форму, войдут и в СДНФ доопределения f_1. Поэтому любой простой импликант функции g будет совпадать с некоторым импликантом функции f_1 или накрываться им. Самые короткие импликанты, накрывающие единицы функции f, есть импликанты функции f_1. Доопределение f_0 имеет минимальное количество конституент единицы в своей СДНФ, следовательно, и количество простых импликант функции f_1, потребных для накрытия этих конституент, будет наименьшим. Кратчайшая по числу букв ДНФ, составленная из простых импликант в сокращенной ДНФ функции f_1, накрывающих все конституенты единицы функции f_0, будет минимальной ДНФ, доопределяющей функцию f.

Так как единицы функции f_1 составлены из единиц функции f и единиц на наборах, на которых f не определена, то построенная ДНФ, накрывая все единицы функции f_0 (а, следовательно, и все единицы функции f), совпадает с минимальной ДНФ некоторого доопределения g функции f.

13.6.1. Алгоритм минимизации частично определенных функций в классе ДНФ

1. Строим СДНФ функции f_0.
2. Строим сокращенную ДНФ функции f_1.
3. С помощью матрицы покрытий конституент единицы функции f_0 простыми импликантами функции f_1 и решеточного выражения строим все тупиковые ДНФ (для некоторых доопределений функции f).
4. Среди полученных ТДНФ выбираем простейшие; они являются минимальными ДНФ (для некоторых доопределений функции f).

13.6.2. Алгоритм минимизации частично определенных функций в классе КНФ

Построение минимальных КНФ для частично определенной функции аналогично построению минимальных КНФ для всюду определенной функции.

Алгоритм минимизации частично определенных функций в классе нормальных форм аналогичен алгоритму минимизации в классе нормальных форм для всюду определенных функций.

Пример. В классе нормальных форм минимизировать частично определенную функцию $f(x,y,z,t) = 1{-}{-}{-}010010{-}01{-}{-}1$.

Решение. Минимизируем функцию f в классе ДНФ.

1. Строим сокращенную ДНФ для доопределения единицами f_1 функции f (табл.13.13).

$$f_1(x,y,z,t) = (x \vee \bar{y} \vee z \vee t)(x \vee \bar{y} \vee \bar{z} \vee t)(x \vee \bar{y} \vee \bar{z} \vee \bar{t})(\bar{x} \vee y \vee z \vee \bar{t})(\bar{x} \vee y \vee \bar{z} \vee \bar{t}) =$$

$$(x \vee \bar{y} \vee t)(xy \vee \bar{x}\bar{y} \vee \bar{z} \vee \bar{t})(\bar{x} \vee y \vee z \vee \bar{t}) =$$

$$(x \vee \bar{y} \vee t)(xy \vee \bar{x}\bar{y} \vee \bar{x}\bar{z} \vee yz \vee \bar{z}\bar{t} \vee \bar{x}\bar{t} \vee y\bar{t} \vee z\bar{t} \vee t) = (x \vee \bar{y} \vee t)(xy \vee \bar{x}\bar{y} \vee \bar{x}\bar{z} \vee y\bar{z} \vee \bar{t}) =$$

$$xy \vee xyz \vee x\bar{t} \vee \bar{x}\bar{y} \vee \bar{x}\bar{y}z \vee \bar{y}\bar{t} \vee xyt \vee \bar{x}\bar{y}\bar{t} \vee \bar{x}\bar{z}t \vee y\bar{z}t =$$

$$xy \vee x\bar{t} \vee \bar{y}\bar{t} \vee \bar{x}\bar{y} \vee \bar{x}\bar{z}t \vee y\bar{z}t.$$

2. Строим матрицу покрытий конституент единицы в СДНФ для доопределения нулями f_0 функции f с помощью построенной сокращенной ДНФ для f_1 (табл.13.14).

3. По табл.13.14 строим решеточный многочлен

$E = (2 \vee 4)(5 \vee 6)(3 \vee 4)(1 \vee 3)1 = 145 \vee 1235 \vee 146 \vee 1236$.

4. Строим все тупиковые ДНФ:

$g_1 = xy \vee \bar{y}\bar{t} \vee \bar{x}\bar{z}t$; $g_3 = xy \vee \bar{y}\bar{t} \vee y\bar{z}t$;

$g_2 = xy \vee \bar{x}\bar{y} \vee x\bar{t} \vee \bar{x}\bar{z}t$; $g_4 = xy \vee \bar{x}\bar{y} \vee x\bar{t} \vee y\bar{z}t$.

Таблица 13.13

x	y	z	t	f	f_0	f_1	$\bar{f}$	h_0	h_1
0	0	0	0	1	1	1	0	0	0
0	0	0	1	–	0	1	–	0	1
0	0	1	0	–	0	1	–	0	1
0	0	1	1	–	0	1	–	0	1
0	1	0	0	0	0	0	1	1	1
0	1	0	1	1	1	1	0	0	0
0	1	1	0	0	0	0	1	1	1
0	1	1	1	0	0	0	1	1	1
1	0	0	0	1	1	1	0	0	0
1	0	0	1	0	0	0	1	1	1
1	0	1	0	–	0	1	–	0	1
1	0	1	1	0	0	0	1	1	1
1	1	0	0	1	1	1	0	0	0
1	1	0	1	–	0	1	–	0	1
1	1	1	0	–	0	1	–	0	1
1	1	1	1	1	1	1	0	0	0

Таблица 13.14

N	ПИ	$\bar{x}\bar{y}\bar{z}\bar{t}$	$\bar{x}y\bar{z}t$	$x\bar{y}\bar{z}\bar{t}$	$xy\bar{z}\bar{t}$	$xyzt$
1	xy				+	+
2	$\bar{x}\bar{y}$	+				
3	$x\bar{t}$			+	+	
4	$\bar{y}\bar{t}$	+		+		
5	$\bar{x}\bar{z}t$		+			
6	$y\bar{z}t$		+			

5. Из построенных тупиковых ДНФ выбираем минимальные:

$$g_1 = xy \vee \bar{y}\bar{t} \vee \bar{x}\bar{z}t; \quad g_3 = xy \vee \bar{y}\bar{t} \vee y\bar{z}t.$$

Функции g_1 и g_3 есть минимальные доопределения функции f в классе ДНФ.

Минимизируем теперь функцию f в классе КНФ. Для этого минимизируем функцию $\bar{f}$ в классе ДНФ. Пусть h_0 и h_1 есть доопределения нулями и единицами соответственно функции $\bar{f}$.

1. Сокращенная ДНФ для $h_1 = (x \vee y \vee z \vee t)(x \vee \bar{y} \vee z \vee \bar{t})\&$
$(\bar{x} \vee y \vee z \vee t)(\bar{x} \vee \bar{y} \vee z \vee t)(\bar{x} \vee \bar{y} \vee \bar{z} \vee \bar{t}) = (x \vee z \vee y\bar{t} \vee \bar{y}t)(\bar{x} \vee z \vee t)(\bar{x} \vee \bar{y} \vee \bar{z} \vee \bar{t}) =$

$$(x \vee z \vee y\bar{t} \vee \bar{y}t)(\bar{x} \vee \bar{y}z \vee z\bar{t} \vee \bar{y}t \vee \bar{z}t) =$$

$$\bar{y}t \vee x\bar{y}z \vee xz\bar{t} \vee x\bar{z}t \vee \bar{x}z \vee \bar{y}z \vee z\bar{t} \vee \bar{x}y\bar{t} \vee yz\bar{t} = \bar{y}t \vee \bar{x}z \vee z\bar{t} \vee x\bar{z}t \vee \bar{x}y\bar{t} \vee \bar{y}z.$$

2. Матрица покрытий конституент единицы в СДНФ для h_0 с помощью простых импликант в сокращенной ДНФ для h_1 приведена в табл.13.15.

3. Решеточное выражение $E = 5(2 \vee 3 \vee 5)2(1 \vee 4)(1 \vee 6) = 25(1 \vee 46) = 125 \vee 2456$.

Таблица 13.15

N	ПИ	$\bar{x}y\bar{z}\bar{t}$	$\bar{x}yz\bar{t}$	$\bar{x}yzt$	$x\bar{y}\bar{z}t$	$x\bar{y}zt$
1	$\bar{y}t$				+	+
2	$\bar{x}z$		+	+		
3	$z\bar{t}$		+			
4	$x\bar{z}t$				+	
5	$\bar{x}y\bar{t}$	+	+			
6	$\bar{y}z$					+

4. Строим две тупиковые ДНФ:

$$g_5 = \bar{y}t \vee \bar{x}z \vee \bar{x}y\bar{t} \text{ и } g_6 = \bar{x}z \vee x\bar{z}t \vee \bar{x}y\bar{t} \vee \bar{y}z.$$

Минимальная ДНФ $g_5 = \bar{y}t \vee \bar{x}z \vee \bar{x}y\bar{t}$.

5. Функция $\bar{g}_5 = (y \vee \bar{t})(x \vee \bar{z})(x \vee \bar{y} \vee t)$ есть минимальное доопределение функции f в классе КНФ.

Найденные МДНФ g_1, g_3 и МКНФ $\bar{g}_5$ являются минимальными доопределениями функции f в классе нормальных форм.

Техническая реализация минимальных форм для функции часто проще, а потому дешевле реализации ее СДНФ (СКНФ). Следовательно, этап минимизации при конструировании логических схем является одним из важнейших.

13.7. Двойственные функции

Двойственной для функции $f(x_1,\ldots,x_n)$ называется функция

$$f^*(x_1,\ldots,x_n) = \bar{f}(\bar{x}_1,\ldots,\bar{x}_n).$$

Примеры. 1. $f = x \,\&\, y$; $f^* = \overline{\bar{x} \,\&\, \bar{y}} = x \vee y$.

2. $f = x \vee y$; $f^* = \overline{\bar{x} \vee \bar{y}} = x \,\&\, y$.

3. $f = x$; $f^* = \bar{\bar{x}} = x$.

4. $f = \bar{x}$; $f^* = \bar{\bar{x}} = \bar{x}$.

5. $f(x_1,\ldots,x_n) = 0$; $f^* = \bar{f}(\bar{x}_1,\ldots,\bar{x}_n) = \bar{0} = 1$.

6. $f(x_1,\ldots,x_n) = 1$; $f^* = \bar{f}(\bar{x}_1,\ldots,\bar{x}_n) = \bar{1} = 0$.

7. $f = x \to y$; $f^* = \overline{\bar{x} \to \bar{y}} = \overline{x \vee \bar{y}} = \overline{y \to x}$.

Заметим, что $(f^*(x_1,\ldots,x_n))^* = (\bar{f}(\bar{x}_1,\ldots,\bar{x}_n))^* = \bar{\bar{f}}(\bar{\bar{x}}_1,\ldots,\bar{\bar{x}}_n) = f(x_1,\ldots,x_n)$, т.е. $(f^*)^* = f$.

Теорема (о суперпозиции двойственных функций). Функция, двойственная суперпозиции функций, равна суперпозиции функций, двойственных к функциям, составляющим эту суперпозицию.

Доказательство. $(f(g_1,\ldots,g_m))^* =$

$\bar{f}(g_1(\bar{x}_{11},\ldots,\bar{x}_{1,n_1}),\ldots,g_m(\bar{x}_{m1},\ldots,\bar{x}_{m,n_m})) =$

$\bar{f}(\bar{\bar{g}}_1(\bar{x}_{11},\ldots,\bar{x}_{1,n_1}),\ldots,\bar{\bar{g}}_m(\bar{x}_{m1},\ldots,\bar{x}_{m,n_m})) =$

$\bar{f}(\bar{g}_1^*(x_{11},\ldots,x_{1,n_1}),\ldots,\bar{g}_m^*(x_{m1},\ldots,x_{m,n_m})) =$

$f^*(g_1^*,\ldots,g_m^*)$. Теорема доказана.

13.7.1. Принцип двойственности

Если функция f задана формулой, построенной с помощью &, $\vee$, $-$, 0, 1 и переменных, то по теореме о суперпозиции двойственных функций и ввиду того, что для функций $x\&y$, $x\vee y$, x, $\bar{x}$, 0, 1 двойственными являются функции $x\vee y$, $x\&y$, x, $\bar{x}$, 1, 0 соответственно, то f^* получается из f заменой & на $\vee$, $\vee$ на &, 0 на 1, 1 на 0 (при сохранении исходной расстановки скобок).

Пример. $\overline{\overline{(x\&y)\vee z}\ \&\ 1\ \&\ (x\vee 0)})^* = \overline{\overline{x\vee y\&z}\ \vee\ 0\ \vee\ (x\&1)}$.

Функция, совпадающая со своей двойственной, называется *самодвойственной*.

Если функция $f(x_1,\ldots,x_n)$ самодвойственна, то функция $\bar{f}$ тоже самодвойственна, так как $(\bar{f}(x_1,\ldots,x_n))^* =$

$$\bar{\bar{f}}(\bar{x}_1,\ldots,\bar{x}_n) = \overline{f^*}(x_1,\ldots,x_n) = \bar{f}(x_1,\ldots,x_n).$$

Теорема. Класс самодвойственных функций замкнут относительно суперпозиции.

Доказательство. Пусть функции $f, g_1,\ldots,g_m$ самодвойственны. Тогда $f^* = f$, $g_1^* = g_1,\ldots,$ $g_m^* = g_m$. Суперпозиция $h = f(g_1,\ldots,g_m)$ этих функций самодвойственна, ибо функция $h^* = (f(g_1\ldots,g_m))^* = f^*(g^*,\ldots,g^*) = f(g_1,\ldots,g_m) = h$ по теореме о суперпозиции двойственных функций.

Наборы $\mathbf{a} = (a_1,\ldots,a_n)$ и $\mathbf{a}' = (\bar{a}_1,\ldots,\bar{a}_n)$ из 0 и 1 называются противоположными.

Следствие. Чтобы функция была самодвойственной, необходимо и достаточно, чтобы на всяких двух противоположных наборах она принимала разные значения.

Доказательство. Функция $f(x_1,\ldots,x_n)$ самодвойственна $\leftrightarrow$

$$f(x_1,\ldots,x_n) = f^*(x_1,\ldots,x_n) \leftrightarrow \bar{f}(\bar{x}_1,\ldots,\bar{x}_n) =$$
$$f(x_1,\ldots,x_n) \leftrightarrow f(x_1,\ldots,x_n) \neq f(\bar{x}_1,\ldots,\bar{x}_n) \leftrightarrow$$

на всяких двух противоположных наборах функция f принимает разные значения.

Пример. f = 01001101, g = 01001111. Функция f самодвойственна, а функция g не самодвойственна, ибо

$$g(0,0,1) = g(1,\ 1,0).$$

Лемма (о несамодвойственной функции). Подстановкой функций x и $\bar{x}$ в несамодвойственную функцию можно получить одну из констант.

Доказательство. Пусть $f(x_1,\ldots,x_n)$ – несамодвойственная функция. Тогда существует набор $(a_1,\ldots,a_n)$, для которого $f(a_1,\ldots,\ a_n) = f(\bar{a}_1,\ldots,\bar{a}_n)$. Построим функцию $h(x)$, заменив единицы в $f(a_1,\ldots,a_n)$ на x, а нули – на $\bar{x}$. Так как $\bar{x} = x^0$, $x = x^1$, то $h(x) = f(x^{a_1},\ldots,x^{a_n})$. Заметим, что $0^{a_i} = \bar{a}_i$, $1^{a_i} = a_i$. Тогда $h(1) = f(1^{a_1},\ldots,1^{a_n}) = f(a_1,\ldots,a_n) = f(\bar{a}_1,\ldots,\bar{a}_n) = f(0^{a_1},\ldots,0^{a_n}) = h(0)$, т.е. $h(1) = h(0)$. Следовательно, функция $h(x)$ есть одна из констант.

13.8. Линейные функции

Арифметические функции в алгебре логики это сложение и умножение по модулю два. Вот эти функции:

x	y	$x+y$	$x\cdot y$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Вычитание (по модулю два) определяется как операция, обратная сложению, т.е. $x - y$ равно такому элементу z, для которого $x = y + z$. В частности, $0 - x$ равно такому y, что $0 = x + y$. Правая часть этого равенства равна нулю только при $y = x$. Поэтому $0-x = x$. Отсюда следует, что $-x = x$ (по модулю два).

Деление (по модулю два) определяется как операция, обрат-ная умножению, т.е. x/y равно такому z, для которого $x = y \cdot z$. В частности, $1/x$ равно такому y, для которого $1 = x \cdot y$. Правая часть равенства равна 1 только при $y = x = 1$. Так что обратный элемент $x^{-1} = 1/x$ возможен лишь при x=1 и равен 1.

Следующие свойства арифметических операций проверяются непосредственно:

1) $x+(y+z) = (x+y)+z$; 2) $x+y = y+x$;
3) $x+0=x$; 4) $x+(-x) = 0$;
5) $x(yz) = (xy)z$; 6) $xy = yx$;
7) $x \cdot 1 = x$; 8) $x \cdot (x^{-1}) = 1$;
9) $x(y+z) = xy+xz$.

Двухэлементное множество $\{0,1\}$ с операциями сложения и умножения по модулю два образуют коммутативное поле F.

Справедливы также следующие свойства:

10) $x + x = 0$; 11) $x \cdot x = x$.

Так что поле F имеет характеристику два, а операция умножения идемпотентна.

Следующие четыре равенства устанавливают связь между арифметическими и логическими операциями:

1) $\bar{x} = x + 1$; 3) $x \vee y = \overline{\bar{x} \,\&\, \bar{y}} = (x+1)(y+1)+1 = xy+x+y$;
2) $x \,\&\, y = x \cdot y$; 4) $x+y = \bar{x}y \vee x\bar{y}$;

Многочлен Жегалкина в поле F есть выражение

$$\sum_{(i_1,\ldots,i_n)\in E_2^n} a_{i_1,i_2,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n},$$

где $x^i = \begin{cases} x, & \text{если } i = 1, \\ 1, & \text{если } i = 0, \end{cases}$

а каждый коэффициент $a_{i_1,i_2,\ldots,i_n}$ равен 0 или 1.

Теорема (Жегалкина). Всякую функцию алгебры логики можно

представить единственным полиномом Жегалкина.

Доказательство. Каждый многочлен Жегалкина определяет некоторую функцию алгебры логики. Два различных многочлена определяют различные функции. Аналогично тому, как это делали при доказательстве единственности СДНФ, можно показать, что существует 2^{2^n} различных многочленов Жегалкина от n переменных.

Пример. Многочлен Жегалкина для функции

$$f(x,y,z) = \bar{x}\bar{y}z \vee \bar{x}yz \vee x\bar{y}z \vee xy\bar{z} \vee xyz = (x+1)(y+1)z + (x+1)yz + x(y+1)z + xy(z+1)+xyz = xyz+xz+yz+z + xyz+yz + xyz+xz + xyz+xy + xyz = xyz + xy + z.$$

Многочлен Жегалкина можно получить также с помощью треугольника Паскаля по единицам его левой стороны (табл.13.16) следующим образом. Построим многочлен Жегалкина для функции g = 10011110. Верхняя сторона треугольника есть фукция g. Любой другой элемент треугольника есть сумма по модулю два двух соседних элементов предыдущей строки. Левая сторона треугольника для функции g содержит шесть единиц. Многочлен Жегалкина будет содержать шесть слагаемых. Первая единица треугольника соответствует набору 000. Первое слагаемое многочлена есть 1. Третья снизу единица в левой стороне треугольника соответствует набору 101. В качестве слагаемого многочлена берем xz. Аналогично для других единиц треугольника. Слева от наборов показаны слагаемые многочлена Жегалкина. Тогда полином Жегалкина

$$g(x,y,z) = 1 + z + y + xz + xy + xyz.$$

Таблица 13.16

N	xyz	f g	Треугольник Паскаля
1	000	0 1	g = 1 0 0 1 1 1 1 0
z	001	1 0	1 0 1 0 0 0 1
y	010	0 0	1 1 1 0 0 1
yz	011	1 1	0 0 1 0 1
x	100	0 1	0 1 1 1
xz	101	1 1	1 0 0
xy	110	1 1	1 0
xyz	111	1 0	1

Функция $f(x_1,\ldots,x_n)$ называется *линейной*, если многочлен Жегалкина для нее имеет линейный относительно переменных вид: $f(x_1,\ldots,x_n) = a_1x_1 + \ldots + a_nx_n + a_{n+1}$, где каждое a_i равно 0 или 1.

Теорема. Класс линейных функций замкнут относительно суперпозиции.

Доказательство. Пусть линейные функции

$$f(x_1,\ldots,x_n) = a_1x_1 + \ldots + a_nx_n + a_{n+1},$$

$$g_i(x_1,\ldots,x_m) = b_{i1}x_1 + \ldots + b_{im}x_m + b_{i,m+1},$$

$$i = 1,2,\ldots,n,$$

для простоты зависят от одних и тех же переменных. Тогда их суперпозиция

$$f(g_1,\ldots,g_n) =$$

$$a_1(b_{11}x_1 + \ldots + b_{1m}x_m+b_{1,m+1}) +$$

$$a_2(b_{21}x_1 + \ldots + b_{2m}x_m+b_{2,m+1}) +$$

$$\ldots$$

$$a_n(b_{n1}x_1 + \ldots + b_{nm}x_m+b_{n,m+1}) + a_{n+1} =$$

$$(a_1b_{11} + \ldots + a_nb_{n1})x_1 + \ldots + (a_1b_{1m} +\ldots + a_nb_{nm})x_m +$$

$(a_1b_{1,m+1} + \ldots + a_nb_{n,m+1}) + a_{n+1}$ есть линейная функция.

Лемма (о нелинейной функции). Суперпозицией нелинейной функции, отрицания и константы 1 можно получить конъюнкцию.

Доказательство. Пусть $f(x_1,\ldots,x_n)$ – нелинейная функция. Тогда полином Жегалкина для нее содержит слагаемое, в котором присутствует произведенние x_ix_j. Будем считать для простоты, что x_1x_2 в многочлене Жегалкина является этим произведением. Произведя группировку слагаемых, функцию f представим в виде

$$f(x_1,\ldots,x_n) = x_1x_2\cdot h_0(x_3,\ldots,x_n) + x_1\cdot h_1(x_3,\ldots,x_n) +$$

$$x_2\cdot h_2(x_3,\ldots,x_n) + h_3(x_3,\ldots,x_n).$$

Функция h_0 не есть тождественный нуль, иначе в полиноме Жегалкина отсутствует слагаемое с произведением x_1x_2. Тогда существует набор $(a_3,\ldots,a_n)$ из 0 и 1, для которого $h_0(a_3,\ldots,a_n) = 1$. Пусть $h_1(a_3,\ldots,a_n) = a$, $h_2(a_3,\ldots,a_n) = b$, $h_3(a_3,\ldots,a_n) = c$. Тогда функция $g(x_1,x_2) = f(x_1,x_2,a_3,\ldots,a_n) = x_1x_2 + ax_1 + bx_2 + c$.

Построим функцию $h(x_1,x_2) = g(x_1+b,x_2+a) = (x_1+b)(x_2+a) + a(x_1+b) + b(x_2+a) + c = x_1x_2 + ax_1 + bx_2 + ab + ax_1 + ab + bx_2 + ab + c = x_1x_2 + d$, где $d = ab + c$. Если $d = 0$, то $h(x_1,x_2) = x_1x_2$. Если $d = 1$, то $h(x_1,x_2) = x_1x_2 + 1$ и тогда $x_1x_2 = \bar{h}(x_1,x_2)$. Лемма доказана.

Определение. Функция $f(x_1\ldots,x_n)$ *сохраняет константу* $a \in \{0,1\}$, если $f(a,\ldots,a) = a$.

Пример. Функция xy сохраняет 0, сохраняет 1. Функция $\bar{x} \to y$ сохраняет 1 и не сохраняет 0.

Теорема. Класс функций, сохраняющих константу, замкнут относительно суперпозиции.

Доказательство. Пусть функции $f(x_1\ldots,x_n)$ и $g_i(x_1,\ldots,x_m)$, $i = 1,2,\ldots,n$, сохраняют константу a. Тогда их суперпозиция $h = f(g_1,\ldots, g_n)$ сохраняет константу a, ибо $h(a,\ldots,a) = f(g_1(a,\ldots,a),\ldots,g_n(a,\ldots,a)) = f(a,\ldots,a) = a$.

13.9. Монотонные функции

Если $\mathbf{a} = (a_1,\ldots,a_n)$ и $\mathbf{b} = (b_1,\ldots,b_n)$ – наборы длины n из 0 и 1, то $\mathbf{a} \leqslant \mathbf{b}$, если $a_1 \leqslant b_1,\ldots,a_n \leqslant b_n$.

Пример. $(0,1,0) \leqslant (1,1,0)$. Наборы $(0,1)$ и $(1,0)$ несравнимы. Также несравнимы наборы $(0,1)$ и $(0,1,0)$.

Функция $f(x_1\ldots,x_n)$ называется *монотонной*, если для всяких наборов $\mathbf{a} = (a_1,\ldots,a_n)$, $\mathbf{b} = (b_1,\ldots,b_n)$ условие $\mathbf{a} \leqslant \mathbf{b}$ влечет $f(\mathbf{a}) \leqslant f(\mathbf{b})$.

Теорема. Класс монотонных функций замкнут относительно суперозиции.

Доказательство. Пусть $f(x_1,\ldots,x_m)$ и $g_i(x_1,\ldots,x_n)$, $i = 1,2,\ldots,n$ – монотонные функции (для простоты от одних и тех же пере-менных). Покажем, что их суперпозиция

$$h(x_1,\ldots,x_n) = f(g_1(x_1,\ldots, x_n),\ldots,g_m(x_1,\ldots,x_n))$$

монотонна. Пусть $\mathbf{c} = (c_1,\ldots, c_n)$, $\mathbf{d} = (d_1,\ldots,d_n)$ – произвольные наборы из 0 и 1, для которых $\mathbf{c} \leqslant \mathbf{d}$, т.е. $c_i \leqslant d_i$, $i = 1,2,\ldots,n$. Пусть $a_i = g_i(\mathbf{c})$, $b_i = g_i(\mathbf{d})$, $i = 1,2,\ldots,m$. Пусть $\mathbf{a} = (a_1,\ldots,a_m)$, $\mathbf{b} = (b_1,\ldots,b_m)$. Так как функции g_i монотонны и $\mathbf{c} \leqslant \mathbf{d}$, то $g_i(\mathbf{c}) \leqslant g_i(\mathbf{d})$, т.е. $a_i \leqslant b_i$, $i = 1,2,\ldots,m$. Следовательно, $\mathbf{a} \leqslant \mathbf{b}$. Отсюда в силу монотонности функции f получаем, что $f(\mathbf{a}) \leqslant f(\mathbf{b})$. Тогда

$$h(\mathbf{c}) = f(g_1(\mathbf{c}),\ldots, g_m(\mathbf{c})) = f(a_1,\ldots,a_m) = f(\mathbf{a}) \leqslant f(\mathbf{b}) =$$

$f(b_1,\ldots,b_m) = f(g_1(\mathbf{d}),\ \ldots,g_m(\mathbf{d})) = h(\mathbf{d})$.

Итак, для любых наборов **c** и **d** из 0 и 1 неравенство $\mathbf{c} \leqslant \mathbf{d}$ влечет $h(\mathbf{c}) \leqslant h(\mathbf{d})$, т.е. функция h монотонна. Теорема доказана.

Лемма (о немонотонной функции). Суперпозицией констант 0 и 1 и переменной x в немонотонную функцию можно получить отрицание.

Доказательство. Пусть $f(x_1\ldots,x_n)$ – немонотонная функция. Тогда существуют наборы $\mathbf{a} = (a_1,\ldots,a_n)$ и $\mathbf{b} = (b_1,\ldots,b_n)$, для которых $\mathbf{a} \leqslant \mathbf{b}$, но $f(\mathbf{a}) > f(\mathbf{b})$. Пусть $i_1,\ldots,i_k$ есть все те номера аргументов, для которых $a_{i_p} < b_{i_p}$, $p = 1,\ldots,k$. На всех остальных аргументных местах j имеем $a_j = b_j$. В выражении $f(a_1,\ldots,a_n)$ заменим нули на местах $i_1,\ldots,i_k$ на x. В результате получим функцию $g(x)$, для которой $g(0) = f(a) = 1$ и $g(1) = f(b) = 0$. Функция $g(x)$ является отрицанием.

Лемма. Функция монотонна тогда и только тогда, когда ее сокращенная ДНФ не содержит отрицаний.

Доказательство. Пусть простой импликант

$$I = x_{i_1}^{c_1} \ldots x_{i_{j-1}}^{c_{j-1}} \cdot \bar{x}_{i_j} \cdot x_{i_{j+1}}^{c_{j+1}} \ldots x_{i_t}^{c_t}$$

монотонной функции $f(x_1,\ldots,x_n)$ содержит отрицание переменной. Покажем, что выражение

$$J = x_{i_1}^{c_1} \ldots x_{i_{j-1}}^{c_{j-1}} \cdot x_{i_{j+1}}^{c_{j+1}} \ldots x_{i_t}^{c_t}$$

есть импликант функции f. Очевидно, что $I = J \cdot \bar{x}_{i_j}$.

Пусть $\mathbf{a} = (a_1,\ldots,a_n)$ есть произвольный набор из 0 и 1 длины n, для которого $J(\mathbf{a}) = 1$. Покажем, что $f(\mathbf{a}) = 1$. Возможны два случая. Пусть набор

$$\mathbf{a} \text{ есть } \mathbf{b} = (a_1,\ldots\ a_{i_{j-1}},0,a_{i_{j+1}},\ldots,a_n).$$

Тогда $I(\mathbf{b}) = J(\mathbf{b}) \cdot \bar{0} = 1$, откуда $f(\mathbf{b}) = 1$, ибо I – импликант для f. Так как **a** есть **b**, то $f(\mathbf{a}) = 1$.

Пусть теперь набор **a** есть

$$\mathbf{c} = (a_1,\ldots,a_{i_{j-1}},1,a_{i_{j+1}},\ldots,a_n).$$

Так как $\mathbf{b} \leqslant \mathbf{c}$, то $1 = f(\mathbf{b}) \leqslant f(\mathbf{c})$, откуда $f(\mathbf{c}) = 1$. Поэтому $f(\mathbf{a}) = 1$. В обоих случаях $J(\mathbf{a}) = 1$ влечет $f(\mathbf{a}) = 1$, т.е. J – импликант функции f. Противоречие с простотой импликанта I.

В обратную сторону доказательство очевидно.

Следствие. Функция монотонна тогда и только тогда, когда ее минимальная ДНФ не содержит отрицаний.

13.10. Теорема Поста о функциональной полноте

Теорема (Поста). Чтобы система функций из P_2 была функционально полной (в P_2), необходимо и достаточно, чтобы эта система содержала:

1) функцию, не сохраняющую 0;
2) функцию, не сохраняющую 1;
3) несамодвойственную функцию;
4) немонотонную функцию;
5) нелинейную функцию.

Доказательство. Пусть система функций F из P_2 полна в P_2. Допустим, что в F нет одной из указанных функций, например немонотонной. Тогда все функции в F монотонны, а так как класс монотонных функций замкнут относительно суперпозиции, то никакая суперозиция над F не даст немонотонной функции.

Пусть система F содержит все указанные функции:

$f_1(x_1,\ldots,x_{n_1})$ не сохраняет 0;
$f_2(x_1,\ldots,x_{n_2})$ не сохраняет 1;
$f_3(x_1,\ldots,x_{n_3})$ несамодвойственная функция;
$f_4(x_1,\ldots,x_{n_4})$ немонотонная функция;
$f_5(x_1,\ldots,x_{n_5})$ нелинейная функция.

Покажем, что суперпозицией функций системы F можно получить полную систему $G = \{x\&y,\bar{x}\}$.

1. Пусть $g(x) = f_1(x,\ldots,x)$. Тогда $g(0) = f_1(0,\ldots,0) = 1$. Далее возможны два случая:

а) $g(1) = 1$. Тогда $g(x) \equiv 1$. Функция $h(x) = f_2(g(x),\ldots,g(x)) = f_2(1,\ldots,1) = 0$, т.е. $h(x) \equiv 0$. Получили константы 0 и 1;

б) $g(1) = 0$. Тогда $g(x) = \bar{x}$. По лемме о несамодвойственной функции суперпозицией над $\{f_3,\bar{x}\}$ можно получить одну из констант, например 0. Тогда $f_1(0,\ldots,0) = 1$ есть другая константа.

В обоих случаях получили обе константы.

2. По лемме о немонотонной функции суперпозицией над $\{f_4, 0, 1\}$ можно получить отрицание.

3. По лемме о нелинейной функции суперпозицией над $\{f_5, 1, \bar{x}\}$ можно получить конъюнкцию. Теорема доказана.

Примеры.

1. $F = \{1, x*y\}$, где $x*y = 0010$.

Проверяем условия полноты:

1) константа 1 не сохраняет 0;
2) $x*y$ не сохраняет 1;
3) $x*y$ не самодвойственна, ибо $0*0 = 1*1$;
4) $x*y$ не монотонна, ибо $(1,0) \leqslant (1,1)$, но $1*0 > 1*1$;
5) $x*y$ не линейна, ибо $x*y = \bar{x}y = xy+y$.

Следовательно, система F по теореме Поста полна. Отрицание есть $1*x$. Конъюнкция $x\&y = \bar{x}*y$.

2. $F = \{0,1,\bar{x},m(x,y,z)\}$, где функция $m(x,y,z) = 00010111$ равна единице на тех и только тех наборах, в которых число единиц больше числа нулей.

Проверяем условия полноты:

1) 0 не сохраняет 1;
2) 1 не сохраняет 0;
3) константа 1 не самодвойственна;
4) отрицание не монотонно;
5) функция $m(x,y,z) = xyz+xy+xz+yz$ не линейна.

По теореме Поста система F полна. Конъюнкция $x\&y = m(x,y,0)$. Заметим, что система F избыточно полна. Ее подсистемы

$$\{1,\bar{x},\ m(x,y,z)\} \text{ и } \{0,\bar{x},m(x,y,z)\}$$

являются функционально полными системами.

13.10.1. Предполные классы

Замкнутый класс K из P_2 *предполон*, если K не является полным, но для всякой функции f не из K система $K \cup \{f\}$ полна.

Теорема. Следующие замкнутые классы функций предполны: класс T_0 функций, сохраняющих константу 0; класс T_1 функций, сохраняющих константу 1; класс M монотонных функций; класс L линейных функций; класс S самодвойственных функций.

Доказательство. Покажем, что класс монотонных функций предполон. Пусть функция $f \notin M$. Заметим, что функции $x\&y$, $x \vee y$, 0, 1 монотонны. По лемме о немонотонной функции отрицание $\bar{x} \in [\{f,0,1\}]$. Система $\{x\&y, x \vee y, \bar{x}\}$ полна.

Покажем, что класс T_1 предполон. Пусть функция $f(x_1,\ldots,x_n) \notin T_1$, т.е. $f(1,\ldots,1) = 0$. Пусть $g(x) = f(x,\ldots,x)$. Тогда $g(1) = f(1,\ldots,1)=0$. Заметим, что функции $x \vee y$, $x\&y$ лежат в T_1. Возможны следующие случаи:

а) $g(0) = 1$. Тогда $g(x) = \bar{x}$. Следовательно, система $\{x\&y, \bar{x}\} \subseteq [T_1 \cup \{f\}]$ и потому класс $T_1 \cup \{f\}$ полон;

б) $g(0) = 0$. Тогда $g(x) \equiv 0$. Очевидно, что $(x \to y) \in T_1$, $\bar{x}$ есть $x \to 0$. Получаем $\{x\&y,\bar{x}\} \subseteq [T_1 \cup \{f\}]$ и потому класс $T_1 \cup \{f\}$ полон.

Покажем, что класс T_0 предполон. Пусть функция $f(x_1,\ldots,x_n) \notin T_0$. Пусть $g(x) = f(x,\ldots,x)$ и $g(0) = f(0,\ldots,0) = 1$. Возможны следующие случаи.

$g(1) = 0$. Тогда $g(x) = \bar{x}$.

$g(1) = 1$. Тогда $g(x) \equiv 1$. Так как $x+y \in T_0$, то $\bar{x} = x+1$. В обоих случаях получено отрицание. Так как $x\&y \in T_0$, то $\{x\&y,\bar{x}\} \subseteq [T_0 \cup \{f\}]$. Следовательно, класс T_0 предполон.

Покажем, что класс L линейных функций предполон. Пусть $f(x_1,\ldots,x_n) \notin L$. Заметим, что функции $x+1$, 0, 1 лежат в L. По лемме о нелинейной функции $x\&y \in [\{f,0,1\}]$. Следовательно, $\{x\&y,\bar{x}\} = [L \cup \{f\}]$ и потому система $L \cup \{f\}$ полна.

Покажем, что класс S самодвойственных функций предполон. Пусть функция $f \notin S$. Заметим, что $\bar{x} \in S$. По лемме о несамодвойственной функции константы 0 и 1 лежат в $[\{f, \bar{x}\}]$. Функция $g(x,y,z) = xy \vee xz \vee yz \in S$. Тогда $g(x,y,1) = xy \vee x \vee y = x \vee y$. Система $\{x \vee y, \bar{x}\} \subseteq [S \cup \{f\}]$. Следовательно, система $S \cup \{f\}$ полна и потому класс S самодвойственных функций предполон. Теорема доказана.

Приведем следующую перефразировку теоремы Поста.

Теорема. Система функций полна, если и только если она не входит целиком ни в один из пяти предполных классов.

Замечание. Пост описал все замкнутые классы в P_2. Оказалось, что множество всех замкнутых классов в P_2 счетно и образует решетку по отношению к включению замкнутых классов. Эта решетка имеет единственный наибольший элемент – класс P_2 и три минимальных элемента: $O_1 = [\{x\}]$, $O_2 = [\{1\}]$, $O_3 = [\{0\}]$.

14. ФУНКЦИИ К-ЗНАЧНОЙ ЛОГИКИ

14.1. Функции и отношения

Пусть $E_k = \{0,1,\ldots,k-1\}$, $k \geqslant 2$. Функции k-значной логики есть функции вида

$$f(x_1,\ldots,x_n):\ E_k^n = \underbrace{E_k\times\ldots\times E_k}_{n\text{ раз}} \to E_k,\ n = 0,1,\ldots$$

h-арное отношение ρ на E_k есть некоторое подмножество из E_k^h.

Определение. Функция $f(x_1,x_2,\ldots,x_n)$ из P_k сохраняет h-арное отношение ρ на E_k (или h-арный предикат ρ), если для всяких n наборов из ρ:

$$\begin{array}{c}(a_1^0, a_1^1, \ldots, a_1^{h-1}) \in \rho,\\ (a_2^0, a_2^1, \ldots, a_2^{h-1}) \in \rho,\\ \ldots\\ (a_n^0\ a_n^1\ \ldots, a_n^{h-1}) \in \rho,\\ \mathbf{a}^0\ \mathbf{a}^1\quad \mathbf{a}^{h-1}\end{array}$$

набор $(f(\mathbf{a}^0),f(\mathbf{a}^1),\ldots,f(\mathbf{a}^{h-1})) \in \rho$.

Определение. Функция $f(x_1,\ldots,x_n)$ существенно зависит от переменной x_i, если $f(a_1,a_2,\ldots,a_{i-1},b,a_{i+1},\ldots,a_n) \neq f(a_1, a_2,\ldots,a_{i-1},c,a_{i+1},\ldots,a_n)$ для некоторых элементов $a_1,a_2, \ldots,a_{i-1},a_{i+1},\ldots,a_n,b,c$ из E_k. Функция f называется существенной, если у нее имеются по меньшей мере две переменные, от которых она зависит существенно.

Определение. Понятие суперпозиции функций над множеством $A \subseteq P_k$ определяется по индукции:

1) каждая функция f из A есть суперпозиция над A;

2) если $g_0(x_1,\ldots,x_n)$ есть суперпозиция над A и если для каждого $i=1,2,\ldots,n$ $g_i(x_{i1},\ldots,x_{i,m_i})$ есть либо суперпозиция над A, либо переменная x_{il} $(1 \leqslant l \leqslant m_i)$, то функция $g_0(g_1(x_{11},\ldots,x_{1,m_1}),\ldots,g_n(x_{n1},\ldots,x_{n,m_n}))$ является суперпозицией над A.

Определение. Множество A функций из P_k называется (функционально) замкнутым классом относительно операции суперпозиции (или просто замкнутым классом), если вместе с любыми функциями из A множеству A принадлежит и их суперпозиция.

Определение. Замыканием множества A из P_k называется множество всех суперпозиций над A.

Замыкание множества A обозначают через $[A]$. Очевидно, что $A \subseteq [A]$. Множество A есть замкнутый класс, если $[A] = A$.

Определение. Множество A функций из P_k называется полным в замкнутом классе $M \subseteq P_k$, если $[A] = M$. Система A называется полной (в P_k), если $[A] = P_k$.

Определение. Множество A из P_k называется предполным классом, если $A \neq P_k$ и для всякой функции f, не принадлежащей A, система функций $A \cup \{f\}$ является полной.

В табл.14.1 приведены примеры функций из P_k. Определим также следующие функции:

$x \vee y = \max(x,y)$, $x\&y = \min(x,y)$,

$x + y = rest(x+y,k)$, $x\cdot y = rest(x\cdot y,k)$,

где $rest(x,k)$ означает остаток от деления x на k, а выражения $x+y$ и $x\cdot y$ в скобках являются обычными суммой и произведением.

Опишем некоторые функционально полные системы.

Теорема. Система функций $\{0,1,\ldots,k-1,J_0(x),J_1(x),\ldots, J_{k-1}(x), \max(x,y), \min(x,y)\}$ полна (в P_k).

Доказательство. Справедливость теоремы вытекает из тождества

$$f(x_1,\ldots,x_n) = \bigvee_{(i_1,\ldots,i_n)\in E_k^n} J_{i_1}(x_1)\&\ldots\&J_{i_n}(x_n)\&f(i_1,\ldots,i_n). \qquad (14.1)$$

Докажем его. Пусть $(a_1,\ldots,a_n)$ – произвольный набор из E_k^n. Рассмотрим равенство (14.1) на этом наборе:

$$f(a_1,\ldots,a_n) = \bigvee_{(i_1,\ldots,i_n)\in E_k^n} J_{i_1}(a_1)\&\ldots\&J_{i_n}(a_n)\&f(i_1,\ldots,i_n). \qquad (14.2)$$

Таблица 14.1

x	$\bar{x}$	Nx	$J_i(x)$	$j_i(x)$
0	1	$k-1$	0	0
1	2	$k-2$	0	0
...	...	...	...	...
i	$i+1$	$k-i-1$	$k-1$	1
...	...	...	...	...
$k-2$	$k-1$	1	0	0
$k-1$	0	0	0	0

Если $(i_1,\ldots,i_n) \neq (a_1,\ldots,a_n)$, то при некотором j будет $i_j \neq a_j$ и тогда $J_{i_j}(a_j) = 0$. В правой части (14.2) останется лишь одно слагаемое – то, у которого $(i_1,\ldots,i_n) = (a_1,\ldots,a_n)$, т.е. слагаемое $(k-1)\&(k-1)\&\ldots\&(k-1)\&f(a_1,\ldots,a_n)$. Оно равно $f(a_1,\ldots,a_n)$. Поскольку в представлении (14.1) участвуют лишь функции, указанные в теореме, то приведенная система функций полна (в P_k).

Теорема. Система функций $\{x+1, \max(x,y)\}$ полна в P_k.

Доказательство. Построим функции $x+l$, $0 \leqslant l \leqslant k-1$. Имеем $x+2 = (x+1)+1$, $x+3 = (x+2)+1$, ..., $x+(k-1) = (x+(k-2))+1$, $x = (x+(k-1))+1$. Очевидно, что $x\vee(x+1)\vee(x+2)\vee\ldots\vee(x+(k-1)) \equiv k-1$. Теперь можно получить остальные константы: $0 = (k-1)+1$, $1 = 0+1,\ldots,k-2 = (k-3)+1$. Далее, $J_i(x) = 1+\max\{x+l: l\neq k-1-i\}$. В самом деле, пусть $x = i$. Тогда левая и правая части этого соотношения равны соответственно $J_i(i) = k-1$ и $1 + \max\{i+l : l \neq k-1-i\} = 1 + \max\{i+l: i+l \neq k-1\} = k-1$. Пусть $x = j \neq i$. Тогда слева имеем $J_i(j) = 0$, а справа $1 + \max\{j + l: l \neq k-1-i\} = 1 + \max\{j,j+1,\ldots,j+k-1-j,\ldots,j+k-1\} = 1+k-1 = 0$.

Построим функции

$$f_{i,s}(x) = \begin{cases} s, & \text{если } x = i, \\ 0, & \text{если } x \neq i, \end{cases} \quad i,s \in E_k.$$

Имеем $f_{i,s}(x) = s + 1 + \max\{J_i(x),k-1-s\}$.

Это равенство подтверждается серией графиков на рис.14.1.

Пусть $\varphi(x)$ есть произвольная унарная функция из P_k (рис.14.2).

Очевидно, что

$$f_{i,\varphi(i)}(x) = \begin{cases} \varphi(i), & \text{если } x = i, \\ 0, & \text{если } x \neq i. \end{cases}$$

Тогда $\varphi(x) = \max\{f_{i,\varphi(i)}(x) : i=0,1,\ldots,k-1\}$. Значит любая одноместная функция, в том числе и функция Nx, выражается через $x+1$ и $\max(x,y)$. Остается заметить, что $\min(x,y) = N(\max(Nx,Ny))$ (см. рис.14.2).

Следствие. Функция $\nu(x,y) = \max(x,y)+1$, называемая часто функцией Вебба, образует полную в P_k систему.

Доказательство. Имеем $x+1 = \nu(x,x)$ и $\max(x,y) = \max(x,y)+1+k-1 = \nu(x,y) + k-1$.

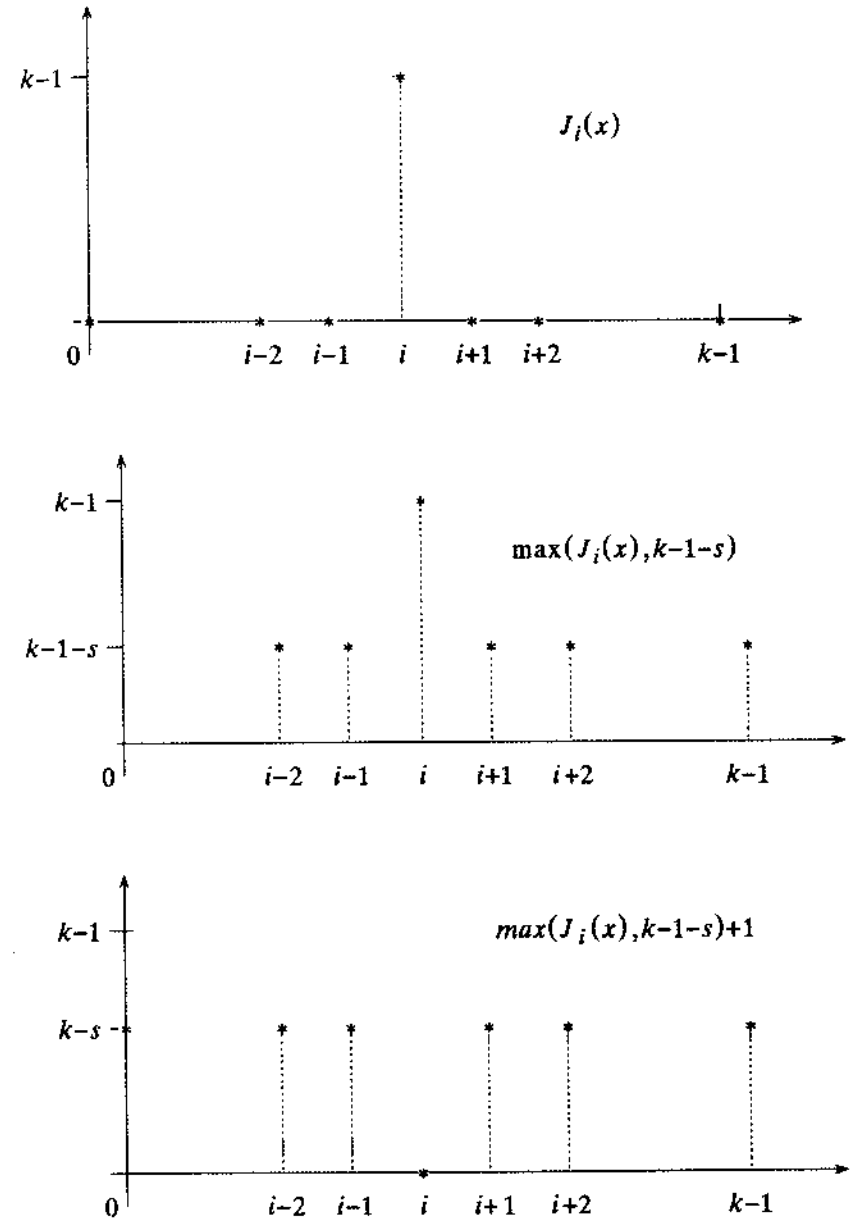
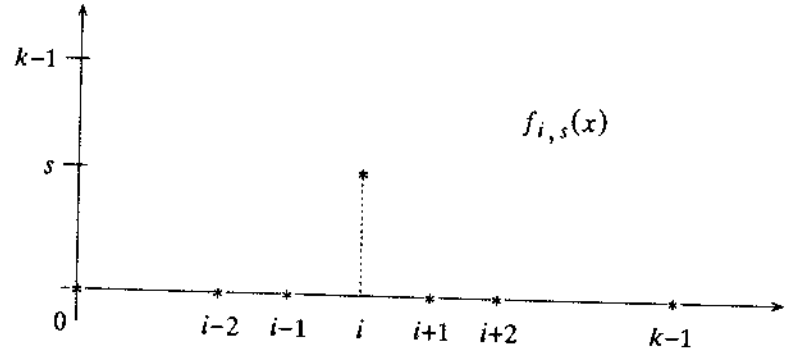


Рис.14.1



Продолжение рис.14.1.

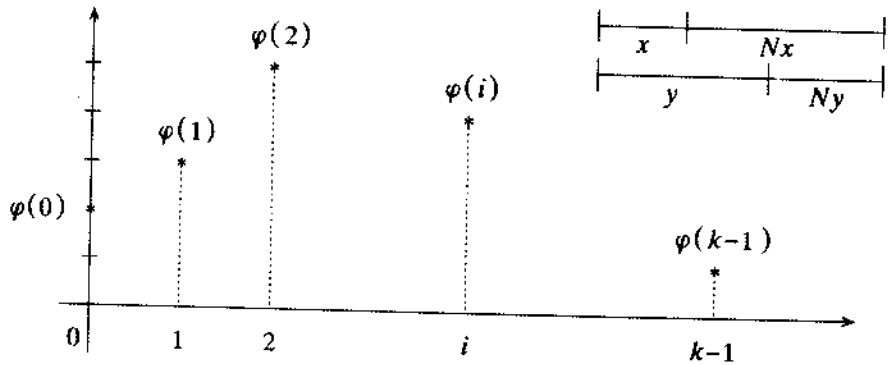


Рис.14.2

14.2. Самодвойственные функции

Определение. Пусть $s(x)$ есть подстановка на E_k. Функция $f^s(x_1,\ldots,x_n) = s^{-1}(f(s(x_1),\ldots,s(x_n)))$ называется двойственной к функции $f(x_1,\ldots,x_n)$ относительно $s(x)$.

Определение. Пусть $s(x) \in \mathfrak{S}_k$. Функция $f(x_1,\ldots,x_n) \in P_k$ называется самодвойственной относительно $s(x)$, если $f^s = f$, т.е. если она двойственна самой себе относительно $s(x)$.

Пусть $Z_k(s(x))$ есть множество всех самодвойственных относительно $s(x)$ функций из P_k.

Утверждение. Множество $Z_k(s(x))$ является замкнутым классом.

Теорема. Если подстановка $s(x)$ на E_k имеет только циклы одинаковой простой длины r, то класс $Z_k(s(x))$ самодвойственных относительно $s(x)$ функций предполон.

14.3. Монотонные функции

Определение. Пусть E есть некоторое множество. Частичным порядком на E называется бинарное отношение ρ, удовлетворяющее следующим аксиомам: $\forall\ a,b,c \in E$

1) $a \preccurlyeq a$;
2) $a \preccurlyeq b\ \&\ b \preccurlyeq a \to a = b$;
3) $a \preccurlyeq b\ \&\ b \preccurlyeq c \to a \preccurlyeq c$.

Если $a \preccurlyeq b$, то элементы a и b называются сравнимыми. В противном случае a и b несравнимы.

Множество E с заданным на нем частичным порядком называется частично упорядоченным.

Элемент e частично упорядоченного множества E называется наибольшим в E, если $\forall a \in E\ (e \succcurlyeq a)$. Элемент o из E называется наименьшим в E, если $\forall\ a \in E\ (o \preccurlyeq a)$. Если наибольший (равно как и наименьший) элемент в частично упорядоченном множестве существует, то он единственный. Элемент m из E называется максимальным в E, если $\forall a \in E$ условие a сравнимо с m влечет $m \succcurlyeq a$. Аналогично, элемент n минимален в E, если $\forall a \in E$ условие a сравнимо с n влечет $n \preccurlyeq a$).

Пусть $\mathbb{O}$ есть множество всех отношений частичного порядка с единственным максимальным и единственным минимальным элементами.

Теорема (В.В.Мартынюк). Пусть $\rho \in \mathbb{O}$. Тогда A_ρ есть предполный класс.

14.4. Линейные функции

Определение. Пусть $G = (E_k,+)$ есть абелева группа, определенная на E_k, с операцией сложения "+" и нулем o. Функция $f(x_1,\ \ldots,x_n)$ из P_k называется линейной относительно G, если $\forall \mathbf{a} = (a_1,\ldots,a_n) \in E_k^n$ и $\forall \mathbf{b} = (b_1,\ldots,b_n) \in E_k^n$ $f(\mathbf{a}+\mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}) - f(\mathbf{o})$, где $\mathbf{a}+\mathbf{b} = (a_1+b_1,\ldots,a_n+b_n)$, $\mathbf{o} = (o,\ldots,o) \in E_k^n$.

Обозначим через L_G множество всех функций, линейных относительно G.

Теорема. L_G есть предполный класс.

14.5. Функции, сохраняющие разбиение

Определение. Функция $f(x_1,\ldots,x_n)$ из P_k сохраняет разбиение $\mathfrak{E}$ множества E_k, если f сохраняет отношение эквивалентности ρ, порожденное этим разбиением.

Через $\mathbb{E}$ обозначим класс всех отношений эквивалентности на E_k, отличных от тривиальных отношений ι_2 и π_2, порождающих разбиения $<\{0\},\{1\},\ldots,\{k-1\}>_k$ и $<0,1,\ldots,k-1>_k$ соответственно.

Теорема. Если $\rho \in \mathbb{E}$, то A_ρ есть предполный класс.

14.6. Классы типа $\mathbb{C}$

Через ι_h, $h\geqslant 2$, обозначим h-арное отношение на E_k, состоящее из всех наборов множества E_k^h, каждый из которых имеет хотя бы два одинаковых элемента.

Определение. h-арное отношение ρ на E_k ($h\geqslant 1$) называется тотально рефлексивным, если $\rho \supseteq \iota_h$.

Всякое унарное отношение является тотально рефлексивным.

Через π_h обозначим h-арное отношение на E_k ($h\geqslant 1$), совпадающее с E_k^h. Через σ_h обозначается h-арное отношение на E_k ($h\geqslant 1$), состоящее из всех таких наборов множества E_k^h, в каждом из которых все элементы разные.

Подстановка на E_k есть взаимно однозначное отображение $s(x)$: $E_k \to E_k$. Через Θ_k обозначим множество всех подстановок на E_k.

Определение. Пусть $1 \leqslant h \leqslant k$, $s(x) \in \Theta_h$ и ρ есть h-арное отношение на E_k. Определим отношение ρ_s (s-сдвиг отношения ρ) с помощью соотношения

$$(a_0,\ldots,a_{h-1}) \in \rho_s \longleftrightarrow (a_{s(0)},\ldots,a_{s(h-1)}) \in \rho.$$

Далее, отношение

ρ симметрично относительно $s \longleftrightarrow \rho = \rho_s$;

ρ тотально симметрично $\longleftrightarrow \rho = \rho_s \quad \forall s \in \Theta_h$;

ρ асимметрично относительно $s \longleftrightarrow \rho \cap \rho_s = \varnothing$.

Через $\mathbb{T}_h$ ($2\leqslant h\leqslant k$) обозначим множество всех h-арных отношений на E_k, которые тотально симметричны, тотально рефлексивны и отличны от ι_h и π_h.

Определение. Отношение $\rho \subseteq E_k^h$ называется центральным, если $\rho \in \mathbb{T}_h$ и $\exists Z \subset E_k$ ($Z \neq \varnothing$ & $\forall a_0,a_1,\ldots,a_{h-2} \in E_k$ $\forall u \in Z$ $(a_0,\ldots,a_{h-2},u) \in \rho$). Множество Z называется центром отношения ρ. Каждый элемент из Z называется центральным.

Через $\mathbb{C}_h$ обозначим множество всех h-арных центральных отношений, $2 \leqslant h \leqslant k$.

Теорема. Если $\rho \in \mathbb{C}$, то A_ρ есть предполный класс.

14.7. Классы типа $\mathbb{B}$

Пусть $h \geqslant 2$, $s \geqslant 1$, $a \in E_h s$ и $[a^{(0)},a^{(1)},\ldots,a^{(s-1)}]$ есть h-адическая запись (развертка) числа a в системе счисления по основанию h (здесь $0 \leqslant a^{(i)} < h$, $i = 0,1,\ldots,s-1$, и $a^{(0)} + a^{(1)}h + \ldots + a^{(s-2)}h^{s-2} + a^{(s-1)}h^{s-1} = a$). Всякий набор $(a^{(0)},a^{(1)},\ldots,a^{(s-1)})$ из E_h^s можно рассматривать как развертку соответствующего числа a из $E_h s$ и наоборот. Таким образом, между элементами из $E_h s$ и наборами из E_h^s существует взаимно однозначное соответствие.

Определение. Определим h-адическое элементарное отношение ζ_s следующим образом. Набор $(a_0,a_1,\ldots,a_{h-1})$ из $E_h^h s$ лежит в ζ_s, если в h-адических записях

$$a_0 = [a_0^{(0)},a_0^{(1)},\ldots,a_0^{(j)},\ldots,a_0^{(s-1)}],$$
$$\ldots$$
$$a_l = [a_l^{(0)},a_l^{(1)},\ldots,a_l^{(j)},\ldots,a_l^{(s-1)}],$$
$$\ldots$$
$$a_{h-1} = [a_{h-1}^{(0)},a_{h-1}^{(1)},\ldots,a_{h-1}^{(j)},\ldots,a_{h-1}^{(s-1)}],$$

чисел a_l из $E_h s$, $l = 0,1,\ldots,h-1$ (в системе счисления по основанию h) каждый набор-столбец $(a_0^{(j)},a_1^{(j)},\ldots,a_{h-1}^{(j)})$ из E_h^s, где $a_l^{(j)}$ есть j-я цифра в h-адической записи числа a_l, $j = 0,1,\ldots,s-1$, лежит в ι_h (т.е. неразнозначен).

Определение. h-арное отношение ρ на E_k называется сильно гомоморфным прообразом h-арного отношения σ на E_l при отображении (сильном гомоморфизме)

$$p\colon E_k \to E_l, \text{ если } \forall a_0,a_1,\ldots,a_{h-1} \in E_k$$
$$(p(a_0),p(a_1),\ldots,p(a_{h-1})) \in \sigma \longleftrightarrow (a_0,a_1,\ldots,a_{h-1}) \in \rho.$$

h-арное отношение ρ на E_k называется сильно гомоморфным прообразом h-арного отношения σ на E_l, если существует сильный гомоморфизм $p\colon E_k \to E_l$, при котором ρ есть прообраз отношения σ.

Теорема. Если 1) $3 \leqslant h \leqslant k$, 2) h-арное отношение ρ на E_k есть сильно гомоморфный прообраз h-адического элементарного отношения ζ_s на E_{h^s} при отображении p: $E_k \to E_{h^s}$ $(h^s \leqslant k)$, то A_ρ является предполным классом.

Замечание. Итак, описаны все шесть семейств предполных классов в k-значной логике: 1) классы самодвойственных функций; 2) классы монотонных функций; 3) классы линейных функций; 4) классы функций, сохраняющих разбиения множества E_k; 5) классы типа $\mathbb{C}$; 6) классы типа $\mathbb{B}$.

Справедлива следующая теорема.

Теорема (Розенберга о функциональной полноте). Система F функций k-значной логики полна (в P_k), если и только если она не содержится целиком ни в одном из предполных классов перечисленных семейств.

14.8. Сравнение функций двузначной и многозначной логик

Отметим сходство и различие двузначных и многозначных логик.

1. P_2 и P_k $(k \geqslant 3)$ имеют конечное число предполных классов.

2. Множество замкнутых классов в P_2 счетно (взаимно однозначно с множеством натуральных чисел), и каждый замкнутый класс имеет конечный базис. Множество замкнутых классов в P_k $(k \geqslant 3)$ имеет мощность континуум (взаимно однозначно с множеством всех вещественных чисел). В P_k $(k \geqslant 3)$ существуют замкнутые классы, не имеющие конечного базиса.

3. Число предполных классов в P_2 равно пяти. Число предполных классов в P_k $(k \geqslant 3)$ растет с ростом k, но число типов предполных классов конечно и равно шести. Число предполных классов в каждом типе предполных классов растет с ростом k.

15. ЧАСТИЧНО УПОРЯДОЧЕННЫЕ МНОЖЕСТВА, РЕШЕТКИ, БУЛЕВЫ АЛГЕБРЫ

15.8. Отношение частичного порядка

Определение. Бинарное отношение $\sigma \subseteq A \times A$, определенное на множестве A, есть отношение частичного порядка (обозначение $a \leqslant b$), если оно удовлетворяет следующим аксиомам: $\forall x,y,z \in A$

1) $x \leqslant x$, рефлексивность,

2) $x \leqslant y \;\&\; y \leqslant x \to x = y$, антисимметричность,

3) $x \leqslant y \;\&\; y \leqslant z \to x \leqslant z$, транзитивность.

Пример. 1. Обычное отношение $x \leqslant y$ меньше или равно, определенное на множестве целых чисел есть отношение частичного порядка.

2. Отношение включения $A \subseteq B$ для подмножеств множества U есть отношение частичного порядка.

Определение. Элементы a,b из A сравнимы, если $a \leqslant b$ или $b \leqslant a$. В противном случае элементы a и b несравнимы.

Определение. Частично упорядоченное множество (ЧУМ) есть пара $(A,\leqslant)$, где A есть множество, а $x \leqslant y$ есть отношение частичного порядка, определенного на A.

Замечание. Частично упорядоченные множества, особенно конечные, удобно изображать диаграммами (Хассе).

Введем отношение "больше или равно", положив $x \geqslant y \longleftrightarrow y \leqslant x$. Отношение $\geqslant$ есть отношение частичного порядка.

Утверждение (принцип двойственности). Если в верном утверждении для ЧУМ $(A,\leqslant)$ знак $\leqslant$ заменить на знак $\geqslant$, то получим тоже верное утверждение.

Введем отношение "строго меньше" и "строго больше", положив $x<y \longleftrightarrow x \leqslant y \;\&\; x \neq y$ и $x>y \longleftrightarrow y \leqslant x$ соответственно.

Из аксиом частичного порядка можно вывести справедливость следующих соотношений: $\forall x,y,z \in A$

1) $x \neg< x$, 2) $x<y \;\&\; y<z \to x<z$.

Для отношений $<$ и $>$ справедлив принцип двойственности.

Определение. Частично упорядоченное множество $(A,\leqslant)$, для которого $\forall x,y \in A$ $(x \leqslant y \vee y \leqslant x)$, называется линейно упорядоченным.

Определение. Элемент $0 \in A$ в частично упорядоченном множестве $(A,\leqslant)$ называется наименьшим, если $\forall x \in A$ $0 \leqslant x$. Элемент I в A наибольший, если $\forall x \in A$ $x \leqslant I$.

Наибольший и наименьший элементы в ЧУМ называются универсальными границами.

Утверждение. Всякое ЧУМ $(A,\leqslant)$ имеет не более одного наименьшего и не более одного наибольшего элементов.

Доказательство. Пусть 0 и $0'$ есть два наименьших элемента в $(A,\leqslant)$. Тогда $0 \leqslant 0'$, ибо 0 есть наименьший элемент в A. $0' \leqslant 0$, ибо 0 есть наименьший элемент в A. Тогда $0 \leqslant 0' \;\&\; 0' \leqslant 0 \to 0 = 0'$, откуда $0=0'$.

Для наибольшего элемента доказательство аналогично.

Определение. Элемент l из A минимален в A, если $\neg\exists x \in A$ $x < l$. Элемент m из A максимален в A, если $\neg\exists x \in A$ $x > l$.

В ЧУМ возможно несколько минимальных и несколько максимальных элементов. Наименьший элемент является минимальным. Обратное не верно. Наибольший элемент является максимальным. Обратное не верно.

Определение. В ЧУМ $(A,\leqslant)$ элемент b непосредственно следует за a (обозначение $a \lessdot b$), если $a<b$ & $\neg\exists x\in A\ a<x<b$.

Утверждение. Если в конечном ЧУМ $(A,\leqslant)$ имеет место $a<b$, то множество A содержит цепь $a=x_0\lessdot x_1\lessdot\ldots\lessdot x_n=b$, в которой x_{i+1} непосредственно следует за x_i $\forall i=0,1,\ldots,n-1$.

Доказательство. Пусть число элементов $y\in A$ со свойством p: $a<y<b$ равно n. Далее индукция по n.

Базис. $n=0$. Тогда $a\lessdot b$ и утверждение справедливо.

Предположение индукции. Пусть утверждение справедливо $\forall a,b\in A$ с числом элементов y со свойством p меньше n.

Шаг индукции. Покажем, что утверждение справедливо $\forall a,b\in A$ с числом элементов y со свойством p равным n. Пусть для некоторого $c\in A$ имеем $a<c<b$. Тогда число элементов $z\in A$ с $a<z<c$ и $c<z<b$ меньше n. По предположению индукции существуют конечные цепи $a=x_0\lessdot x_1\lessdot\ldots\lessdot x_k=c$, $c=x_k\lessdot x_{k+1}\lessdot\ldots\lessdot x_n=b$ и цепь $a=x_0\lessdot x_1\lessdot\ldots\lessdot x_k\lessdot x_{k+1}\lessdot\ldots\lessdot x_n=b$ искомая.

Определение. Пусть $(A,\leqslant)$ есть ЧУМ и множество $B\subseteq A$.

1. Элемент $a\in A$ есть верхняя грань для B, если $\forall b\in B\ b\leqslant a$.

2. Совокупность B^{Δ} всех верхних граней для B называется верхним конусом для B.

3. Элемент $b^{\Delta}\in B^{\Delta}$ есть точная верхняя грань для B (обозначение $\sup B$), если b^{Δ} есть наименьший элемент в B^{Δ}, то есть если $\forall b\in B^{\Delta}\ b\leqslant b^{\Delta}$.

4. Элемент $a\in A$ есть нижняя грань для B, если $\forall b\in B\ a\leqslant b$.

5. Совокупность B^{∇} всех нижних граней для B называется нижним конусом для B.

6. Элемент $b^{\nabla}\in B^{\nabla}$ есть точная нижняя грань для B (обозначение $\inf B$), если b^{∇} есть наибольший элемент в B^{∇}, то есть если $\forall b\in B^{\nabla}\ b^{\nabla}\leqslant b$.

Пример. $A=\{0,1,\ldots,21\}$, $B=\{6,7,10,11\}$. ЧУМ $(A,\leqslant)$ изображено на рис.15.1.

Наибольший элемент в A не существует.

Наименьший элемент в A есть 0.

Множество максимальных элементов в A есть $\{19,20,21\}$.

Верхний конус для B есть $B^{\Delta}=\{21\}$.

Точная верхняя грань для B есть 21 (это наименьший элемент в B^{Δ}).

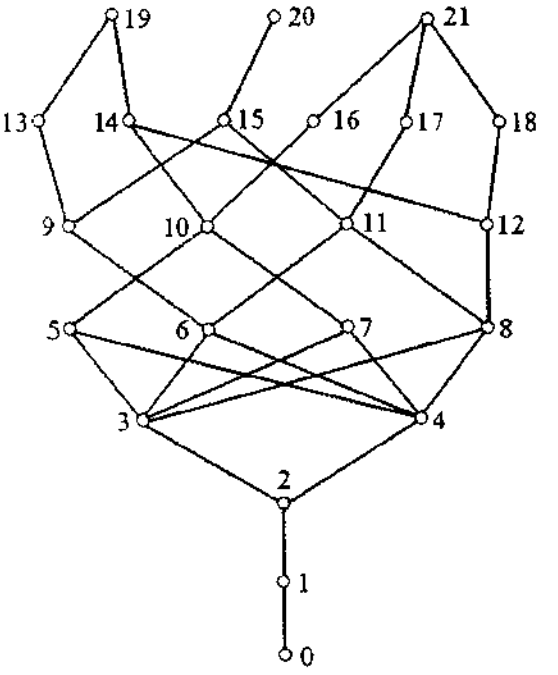


Рис.15.1

Множество минимальных элементов в A есть $\{0\}$.

Нижний конус для B есть $B^{\nabla}=\{0,1,2,3,4\}$.

Точная нижняя грань для B не существует (нет наибольшего элемента в B^{∇}).

Лемма Цорна. Если в ЧУМ $(A,\leqslant)$ любая цепь имеет верхнюю (нижнюю) грань, то A имеет максимальный (минимальный) элемент.

Замечание. $\forall x,y\in A$

$$z=\inf(x,y)\longleftrightarrow \begin{matrix} z\leqslant x \\ z\leqslant y \end{matrix}\ \&\ \forall u\left(\begin{matrix} u\leqslant x \\ u\leqslant y \end{matrix}\rightarrow u\leqslant z\right).$$

$$z = \sup(x,y) \longleftrightarrow \begin{matrix} z \geqslant x \\ z \geqslant y \end{matrix} \;\&\; \forall u \left(\begin{matrix} u \geqslant x \\ u \geqslant y \end{matrix} \rightarrow u \geqslant z \right).$$

15.9. Решетки

Определение. Решетка $L=(A,\vee,\wedge)$ есть множество A с двумя определенными на нем бинарными операциями (функциями):

$$x \vee y : A\times A \to A, \quad x \wedge y : A\times A \to A,$$

удовлетворяющими следующим аксиомам: $\forall x,y,z \in A$

$L1$. $x \vee x = x$, $x \wedge x = x$.
$L2$. $x \vee y = y \vee x$, $x \wedge y = y \wedge x$.
$L3$. $x \vee (y \vee z) = (x \vee y) \vee z$,
$x \wedge (y \wedge z) = (x \wedge y) \wedge z$.
$L4$. $x \vee (x \wedge y) = x$, $x \wedge (x \vee y) = x$.

Обозначение. Решетка $L=(A,\vee,\wedge)$, $x\&y$, $x \wedge y$, $x \cdot y$, xy.
Иногда вместо $x \wedge y$ пишут $x\&y$ или xy.

Решетка называется дистрибутивной, если для нее выполняются равенства $L1$–$L4$ и равенства

$L5$. $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$,
$x \wedge (y \vee z) = (x \vee y) \wedge (x \vee z)$.

Решетка называется модулярной, если для нее выполняются равенства $L1$–$L4$ и равенства

$L6$. $x \wedge (y \vee (x \wedge z)) = (x \wedge y) \vee (x \wedge z)$,
$x \vee (y \wedge (x \vee z)) = (x \vee y) \wedge (x \vee z)$.

Решетка L модулярна $\longleftrightarrow$ L не содержит подрешетки, изображенной на рис.15.2.

Замечание (принцип двойственности). Если в верном в решетке L равенстве заменить знак $\vee$ на знак $\wedge$, а знак $\wedge$ на знак $\vee$, то получим снова верное равенство.

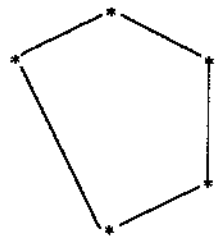


Рис.15.2

Лемма. В решетке $L=(A,\vee,\wedge)$ справедливо $x \wedge y = x \longleftrightarrow x \vee y = y$.

Доказательство.

(1) $x \wedge y = x$. (1) $x \vee y = y$.
(2) $y = y \vee xy$, $L4$. (2) $x = x(x \vee y)$, $L4$.
(3) $y = y \vee x$, (1,2). (3) $x = xy$, (1,2).
(4) $x \vee y = y$, (3),$L2$.

Введем отношение $x \leqslant y$, положив $x \leqslant y \longleftrightarrow xy = x$. Тогда по лемме $x \leqslant y \longleftrightarrow x \vee y = y$.

Теорема. В решетке $L=(A,\vee,\wedge)$ отношение $x \leqslant y$ есть отношение частичного порядка.

Доказательство. Покажем, что все три аксиомы частичного порядка для введенного отношения справедливы.

1. $x \leqslant x$, ибо $x \cdot x = x$ по аксиоме $L1$.
2. $x \leqslant y \;\&\; y \leqslant x \to x = y$. В самом деле,

$$\begin{matrix} x \leqslant y \\ y \leqslant x \end{matrix} \underset{def}{\longrightarrow} \begin{matrix} xy = x \\ yx = y \end{matrix} \underset{L2}{\longrightarrow} \begin{matrix} xy = x \\ xy = y \end{matrix} \longrightarrow x = y.$$

3. $x \leqslant y \;\&\; y \leqslant z \to x \leqslant z$. В самом деле,

$$\begin{matrix} x \leqslant y \\ y \leqslant z \end{matrix} \underset{def}{\longrightarrow} \begin{matrix} xy = x \\ yz = y \end{matrix} \longrightarrow xz \underset{x=xy}{=} (xy)z \underset{L3}{=} x(yz) \underset{yz=y}{=} xy \underset{xy=x}{=} x,$$

откуда $xz = x$, то есть $x \leqslant z$.

Следствие. Решетка $L=(A,\vee,\wedge)$ есть ЧУМ $(A,\leqslant)$ с отношением частичного порядка $x \leqslant y \longleftrightarrow xy = x$.

Замечание. Во всяком ЧУМ можно определить функции $x \vee y = \sup(x,y)$ и $x \wedge y = \inf(x,y)$.

Теорема. В ЧУМ $(A,\leqslant)$ справедливы следующие равенства (если участвующие в этих равенствах операнды определены):

$L1$. $x \vee x = x$, $x \wedge x = x$.
$L2$. $x \vee y = y \vee x$, $x \wedge y = y \wedge x$.
$L3$. $x \vee (y \vee z) = (x \vee y) \vee z$,
$x \wedge (y \wedge z) = (x \wedge y) \wedge z$.
$L4$. $x \vee (x \wedge y) = x$, $x \wedge (x \vee y) = x$.

Доказательство. В силу принципа двойственности достаточно доказать только одну часть этих равенств, например, левую.

$L1$. $x \vee x = \sup(x,x) = x$.
$L2$. $x \vee y = \sup(x,y) = \sup(y,x) = y \vee x$.
$L3$. $x \vee (y \vee z) = \sup(x,y,z) = (x \vee y) \vee z$.
$L4$. $x \vee xy = \sup(x, \underbrace{\inf(x,y)}_{\leqslant x,\ \leqslant y}) = x$.

Следствие. Если в ЧУМ $(A,\leqslant)$ операции $\sup(x,y)$ и $\inf(x,y)$

всюду определены, то по отношению к этим операциям ЧУМ есть решетка.

Вывод. 1. Всякая решетка $(A,\vee,\wedge)$ с отношением $x\leqslant y$, определенным как $xy=x$ (или $x\vee y=y$) есть частично упорядоченное множество $(A,\leqslant)$.

2. Если в ЧУМ $(A,\leqslant)$ операции $x\vee y=\sup(x,y)$ и $x\wedge y=\inf(x,y)$ всюду определены, то система $(A,\vee,\wedge)$ есть решетка.

Решетки, имеющие наименьший 0 и наибольший I элементы, называются решетками с универсальными границами. Возможны решетки с одной и двумя универсальными границами, возможны без обеих. Всякая конечная решетка имеет универсальные границы 0 и I.

Универсальные границы обладают следующими свойствами.

L7. $x\vee I=I$, $x\vee 0=a$; $x\wedge I=x$, $x\wedge 0=0$.

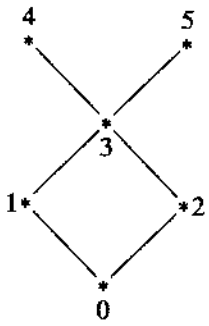
Дополнение элемента a решетки L есть такой элемент $\bar{a}$ в L, что $\bar{a}\vee a=I$, $\bar{a}\wedge a=0$.

Решетка, в которой каждый элемент имеет дополнение, называется решеткой с дополнениями. Дистрибутивная решетка с дополнениями называется булевой решеткой.

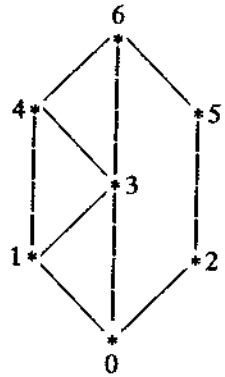
Пример. 1. Рис.15.3.

2. Система $(\mathbb{N},\leqslant)$ с отношением $m\leqslant n \longleftrightarrow m\,|\,n$ есть ЧУМ. Тогда функции $m\vee n=\sup(m,n)=\mathrm{LCM}(m,n)$, $m\wedge n=\inf(m,n)=\mathrm{GCD}(m,n)$. Система $(\mathbb{N},\vee,\wedge)$ есть решетка.

3. Пусть U есть некоторое множество, $\mathcal{P}(U)$ – множество всех подмножеств множества U. Система $(\mathcal{P}(U),\subseteq)$ есть ЧУМ.



ЧУМ не решетка, ибо $\neg\exists\sup(4,5)$



ЧУМ решетка

Рис.15.3

Тогда функции $A\vee B=\sup(A,B)=A\cup B$, $A\wedge B=\inf(A,B)=A\cap B$. Система $(\mathcal{P}(U),\vee,\wedge)$ есть решетка.

15.10. Изоморфизм решеток

Определение. Частично упорядоченные множества $(A_1,\leqslant_1)$ и $(A_2,\leqslant_2)$ изоморфны, если существует взаимно однозначное соответствие (1–1–соответствие) $\varphi: A_1\to A_2$, сохраняющее отношение порядка: $\forall x,y\in A_1\ (x\leqslant_1 y \longleftrightarrow \varphi(x)\leqslant_2\varphi(y))$.

Определение. Решетки $L_1=(A_1,\vee_1,\wedge_1)$, $L_2=(A_2,\vee_2,\wedge_2)$ изоморфны (обозначение. $L_1\cong L_2$), если существует взаимно однозначное соответствие $\varphi: A_1\to A_2$, для которого $\forall x,y\in A_1\ (\varphi(x\vee_1 y)=\varphi(x)\vee_2\varphi(y)$ и $\varphi(x\wedge_1 y)=\varphi(x)\wedge_2\varphi(y))$.

Теорема. Пусть $L_1=(A_1,\vee_1,\wedge_1)$, $L_2=(A_2,\vee_2,\wedge_2)$ есть две решетки. Пусть в решетках L_1 и L_2 частичные порядки соответственно

$x\leqslant_1 y \longleftrightarrow x\wedge_1 y=x$ (или $x\vee_1 y=y$),
$x\leqslant_2 y \longleftrightarrow x\wedge_2 y=x$ (или $x\vee_2 y=y$).
Тогда $(A_1,\vee_1,\wedge_1)\cong(A_2,\vee_2,\wedge_2) \longleftrightarrow (A_1,\leqslant)\cong(A_2,\leqslant_2)$.

Доказательство. $\to$. Пусть $L_1\cong L_2$. Тогда существует 1–1–соответствие $\varphi: A_1\to A_2$, сохраняющее операции $x\vee y$ и $x\wedge y$. Покажем, что отображение φ сохраняет частичный порядок $\leqslant$. В самом деле, пусть $x\leqslant_1 y$. Тогда $x\wedge_1 y=x$. Так как φ сохраняет операцию $\wedge$, то $\varphi(x\wedge_1 y)=\varphi(x)=\varphi(x)\wedge_2\varphi(y))$, откуда $\varphi(x)\leqslant_2\varphi(y)$. Тогда $x\leqslant_1 y\to\varphi(x)\leqslant_2\varphi(y)$.

Пусть теперь $\varphi(x)\leqslant_2\varphi(y)$. Для элементов x и y справедливо одно из трех: $x\leqslant_1 y$, $x>_1 y$, x и y несравнимы. Случай $x>_1 y$ невозможен, ибо тогда $\varphi(x)>_2\varphi(y)$, чего нет. Случай несравнимости x и y тоже невозможен, ибо тогда $x\neg\leqslant_1 y$ и $y\neg\leqslant_1 x$ и потому $x\wedge_1 y\neq x$, $x\wedge_1 y\neq y$, откуда $\varphi(x)\wedge_2\varphi(y)\neq\varphi(x)$, $\varphi(x)\wedge_2\varphi(y)\neq\varphi(y)$, то есть элементы $\varphi(x),\varphi(y)$ несравнимы, чего нет, ибо по условию элементы $\varphi(x),\varphi(y)$ сравнимы. Остается $x\leqslant_1 y$. Тогда $\varphi(x)\leqslant_2\varphi(y)\to x\leqslant_1 y$. Показано: $\forall x,y\in A_1\ (x\leqslant_1 y \longleftrightarrow\varphi(x)\leqslant_2\varphi(y))$, то есть $(A_1,\leqslant)\cong(A_2,\leqslant)$.

$\leftarrow$. Пусть $(A_1,\leqslant)\cong(A_2,\leqslant)$ и пусть в этом изоморфизме отображение $\varphi: A_1\to A_2$ есть 1–1–соответствие, сохраняющее порядок. Покажем, что отображение φ сохраняет операции $x\vee y=\sup(x,y)$ и $x\wedge y=\inf(x,y)$, то есть что

$$\begin{matrix} x\wedge_1 y=z \\ x\vee_1 y=z' \end{matrix} \quad\text{влечет}\quad \begin{matrix} \varphi(x)\wedge_2\varphi(y)=\varphi(z) \\ \varphi(x)\vee_2\varphi(y)=\varphi(z') \end{matrix}.$$

В самом деле, пусть $x\wedge_1 y=z$. Тогда $z=\inf(x,y)$, откуда

$$\begin{matrix} z\leqslant_1 x \\ z\leqslant_1 y \end{matrix} \;\&\; \forall u\in A_1 \left[\begin{matrix} u\leqslant_1 x \\ u\leqslant_1 y \end{matrix} \to u\leqslant_1 z\right].$$

Так как 1–1–отображение φ сохраняет порядок, то

$$\begin{matrix} \varphi(z)\leqslant_2\varphi(x) \\ \varphi(z)\leqslant_2\varphi(y) \end{matrix} \;\&\; \forall\varphi(u)\in A_2 \left[\begin{matrix} \varphi(u)\leqslant_2\varphi(x) \\ \varphi(u)\leqslant_2\varphi(y) \end{matrix} \to \varphi(u)\leqslant_2\varphi(z)\right],$$

то есть $\varphi(z)=\inf(\varphi(x),\varphi(y))$, откуда $\varphi(x)\wedge_2\varphi(y)=\varphi(z)$. Второе соотношение доказывается аналогично. Следовательно, $L_1 \cong L_2$.

Замечание. Для установления изоморфизма решеток достаточно показать изоморфизм соответствующих ЧУМ.

15.11. Булевы алгебры

Определение. Булева алгебра $\mathcal{A}=(A,\vee,\wedge,-,0,1)$ есть множество A, на котором определены три операции:

$$x\vee y\colon A\times A\to A,\quad x\wedge y\colon A\times A\to A,\quad \bar{x}\colon A\to A$$

и два отмеченных элемента 0 и 1 (универсальные границы), удовлетворяющие следующим аксиомам: $\forall x,y,z\in A$

*B*1. $x\wedge x = x$, $x\vee x = x$.
*B*2. $x\wedge y = y\wedge x$, $x\vee y = y\vee x$.
*B*3. $x\wedge(y\wedge z) = (x\wedge y)\wedge z$, $x\vee(y\vee z) = (x\vee y)\vee z$.
*B*4. $x\wedge(x\vee y) = x$, $x\vee x\wedge y = x$.
*B*5. $x\wedge(y\vee z) = x\wedge y\vee x\wedge z$, $x\vee y\wedge z = (x\vee y)\wedge(x\vee z)$.
*B*6. $\bar{\bar{x}} = x$.
*B*7. $x\wedge 1 = x$, $x\vee 0 = x$, $x\wedge 0 = 0$, $x\vee 1 = 1$.
*B*8. $x\vee\bar{x} = 1$, $x\wedge\bar{x} = 0$.
*B*9. $\overline{x\wedge y} = \bar{x}\vee\bar{y}$, $\overline{x\vee y} = \bar{\bar{x}}\wedge\bar{y}$.

Иногда вместо $x\wedge y$ пишут $x\&y$ или xy, а вместо $\bar{x}$ пишут $\neg x$.

Пример. 1. Если U есть конечное множество с n элементами и $\mathcal{P}(U)$ есть множество всех его 2^n подмножеств, то $(\mathcal{P}(U), \vee,\wedge,-,0,1)$ есть булева алгебра, в которой для подмножеств A и B в ней $A\vee B = A\cup B$, $A\wedge B = A\cap B$, $\bar{A} = U-A$, $0=\varnothing$, $1=U$.

2. Система $(\{0,1\},\vee,\&,\neg,0,1)$ с опреациями конъюнкция, дизъюнкция, отрицание есть булева алгебра.

Замечание. В качестве булевой алгебры можно взять систему $\mathcal{A} = (A,\vee,\wedge,\neg)$ с аксиомами *B*2–*B*5 и аксиомой

*B*10. $(x\wedge\bar{x})\vee y = y$, $(x\vee\bar{x})\wedge y = y$.

Константы $0 = x\wedge\bar{x}$, $1 = x\vee\bar{x}$. Остальные аксиомы могут быть выведены из *B*2–*B*5, *B*10.

Замечание (принцип двойственности). Если в булевой алгебре истинно некоторое равенство, построенное с помощью $\vee,\wedge,-$, 0,1, то истинно и равенство, полученное из исходного заменой в нем $\vee$ на $\wedge$, $\wedge$ на $\vee$, 0 на 1, 1 на 0.

Определение. Две булевых алгебры изоморфны: $\mathcal{A}_1=(A_1,\vee_1,\wedge_1, -_1,0_1,1_1) \cong (A_2,\vee_2,\wedge_2,-_2,0_2,1_2)=\mathcal{A}_2$, если существует взаимно однозначное соответствие $\varphi\colon A_1\to A_2$, сохраняющее операции и отмеченные элементы: $\forall x,y\in A_1$

$$\varphi(x\vee_1 y)=\varphi(x)\vee_2\varphi(y),\ \varphi(x\wedge_1 y)=\varphi(x)\wedge_2\varphi(y),\ \varphi(\neg_1 x)=\neg_2\varphi(x),$$
$$\varphi(0_1)=0_2,\ \varphi(1_1)=1_2.$$

Замечание. Всякая булева алгебра $(A,\vee,\wedge,-,0,1)$ есть решетка $(A,\vee,\wedge)$, в которой можно задать отношение частичного порядка, определяемое соотношением: $x\leqslant y \leftrightarrow xy=x$ (или $x\leqslant y \leftrightarrow x\vee y=y$). Тогда $(A,\leqslant)$ есть ЧУМ.

Теорема. Если $\mathcal{A}_i=(A_i,\vee_i,\wedge_i,-_i,0_i,1_i)$, $i=1,2$, есть две булевых алгебры, то $\mathcal{A}_1\cong\mathcal{A}_2 \leftrightarrow (A_1,\leqslant_1) \cong (A_2,\leqslant_2)$. Другими словами, булевы алгебры $\mathcal{A}_1$, $\mathcal{A}_2$ изоморфны $\leftrightarrow$ $\mathcal{A}_1$, $\mathcal{A}_2$ изоморфны как ЧУМ.

Доказательство аналогично доказательству соответствующей теоремы для решеток.

Включение множеств булевых алгебр, решеток, ЧУМ показано на рис.15.4.

Теорема (Стоуна о представлении). Каждая конечная булева алгебра изоморфна булевой алгебре всех подмножеств некоторого конечного множества.

Доказательство. Пусть $\mathcal{A}=(A,\vee,\wedge,-,0,1)$ есть конечная булева алгебра. Введем в $\mathcal{A}$ отношение частичного порядка, положив

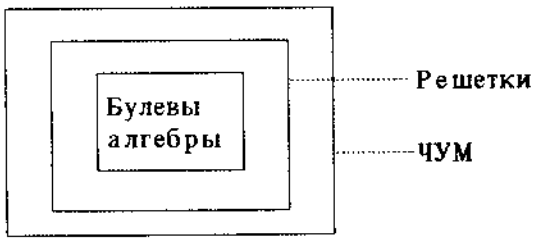


Рис.15.4

$x \leqslant y \longleftrightarrow xy=x$ (или $x \leqslant y \longleftrightarrow x \vee y=y$), и рассмотрим ЧУМ $(A, \leqslant)$. Элемент $a \in A$ есть атом, если a непосредственно следует за нулем $0<a$. Для двух атомов $a_1, a_2 \in A$ имеем

$$a_1 \wedge a_2 = \begin{cases} a_1, & \text{если } a_1=a_2, \\ 0, & \text{если } a_1 \neq a_2. \end{cases}$$

Верхнее равенство очевидно. Пусть $a_1 \neq a_2$. Тогда $a_1 \wedge a_2 \leqslant a_1$, $a_1 \wedge a_2 \leqslant a_2$. Так как $a_1 \neq a_2$, то одно из этих неравенств строгое, например, первое. Тогда $a_1 \wedge a_2 < a_1$, что возможно лишь при $a_1 \wedge a_2 = 0$.

Пусть $U=\{a_1, \ldots, a_n\}$ есть множество всех атомов булевой алгебры. Определим функцию $\varphi: \mathcal{P}(U) \to A$, положив

$$\varphi(\varnothing) = 0, \quad \varphi(\{a_{i_1}, \ldots, a_{i_k}\}) = a_{i_1} \vee \ldots \vee a_{i_k}.$$

Докажем, что отображение φ взаимно однозначно.

1. Покажем, что $\mathrm{Im}(\varphi)=A$. Пусть $a \neq 0$, $a \in A$. Так как $0<a$, то между 0 и a существует цепь $0<b_1<b_2<\ldots<b_m=a$, в которой элемент b_1 есть атом. Проведем все возможные цепи между 0 и a (рис.15.5) и пусть $\{c_1, \ldots, c_k\}$ есть множество всех атомов, составляющих эти цепи.

Очевидно, что $c_i \leqslant a$ $\forall i=1, \ldots, k$, и потому $b=c_1 \vee \ldots \vee c_k \leqslant a$. Покажем, что $b=a$. Это и будет означать, что $\varphi(\{c_1, \ldots, c_k\}) = c_1 \vee \ldots \vee c_k = b = a$, то есть $a \in \mathrm{Im}(\varphi)$. Допустим противное: $a \neq b$. Так как $b \leqslant a$ и $b \neq a$, то $b<a$. Пусть $c=a \wedge \bar{b}$. Тогда $c \wedge a = (a \wedge \bar{b}) \wedge a = (a \wedge a) \wedge \bar{b} = a \wedge \bar{b} = c$, $c \wedge a = c$, откуда $c \leqslant a$, $0 \leqslant c \leqslant a$.

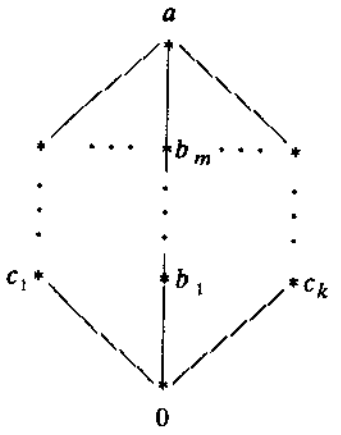


Рис.15.4

Покажем, что $c \neq 0$. Допустим противное: $c=0$. Тогда $c=a \wedge \bar{b}=0$, откуда $(a \wedge \bar{b}) \vee b = c \vee b = 0 \vee b = b$, $(a \wedge \bar{b}) \vee b = (a \vee \bar{b}) \wedge (\bar{b} \vee b) = a \vee b$ влечет $b=a \vee b$, то есть $a \leqslant b$. Противоречие с $b<a$.

Покажем, что $c \neq a$. Допустим противное: $c=a$, то есть $c = a \wedge \bar{b} = a$. Тогда $\underbrace{(a \wedge \bar{b})}_{=c=a} \wedge b = c \wedge b = a \wedge b = a \wedge (\bar{b} \wedge b) = a \wedge 0 = 0$, откуда $a \wedge b = 0$.

Так как $b<a$, то $a \neq 0$ и потому $b=0$. Противоречие с $b \neq 0$, ибо $b=c_1 \vee \ldots \vee c_k$ есть объединение ненулевых элементов.

Итак, $c \neq 0$, $c \neq a$, поэтому $0<c<a$. Соединим 0 с элементом a цепью, проходящей через c, и возьмем в этой цепи атом a^*. Тогда $a^*=c_p$ при некотором $p \in \{1, 2, \ldots, k\}$ (рис.15.6).

Далее, $a^* \leqslant c$ влечет $a^* \wedge c = a^*$. Но

$$a^* \wedge c = c_p \wedge (a \wedge \underbrace{\overbrace{\bar{c}_1 \vee \ldots \vee \bar{c}_p \vee \ldots \vee \bar{c}_k}^{\neg b}}_{c}) =$$

$$c_p \wedge \bar{c}_p \wedge (a \wedge (\bar{c}_1 \wedge \ldots \wedge \bar{c}_{p-1} \wedge \bar{c}_{p+1} \wedge \ldots \wedge \bar{c}_k) = 0.$$

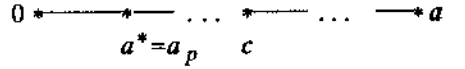


Рис.15.6

Итак, $a^* \wedge c = a^* \neq 0$ и $a^* \wedge c = 0$. Противоречие. Следовательно, $b=a$, откуда $\mathrm{Im}(\varphi)=A$.

2. Пусть $A'=\{a_{i_1}, \ldots, a_{i_k}\}$, $A''=\{a_{j_1}, \ldots, a_{j_l}\}$ есть подмножества атомов из U, причем $A' \neq A''$. Покажем, что $\varphi(A') \neq \varphi(A'')$. Допустим противное: $\varphi(A')=\varphi(A'')$. Тогда $a_{i_1} \vee \ldots \vee a_{i_k} = a_{j_1} \vee \ldots \vee a_{j_l}$. Умножим обе части на a_{i_p}. Тогда

$$(a_{i_1} \vee \ldots \vee a_{i_k}) a_{i_p} = (a_{j_1} \vee \ldots \vee a_{j_l}) a_{i_p},$$

$$\underbrace{a_{i_1} a_{i_p}}_{0} \vee \ldots \vee \underbrace{a_{i_p} a_{i_p}}_{a_{i_p}} \vee \ldots \vee \underbrace{a_{i_k} a_{i_p}}_{0} = a_{j_1} a_{i_p} \vee \ldots \vee$$

$$a_{j_1}a_{i_p},\quad a_{i_p} = a_{j_1}a_{i_p} \vee \ldots \vee a_{j_l}a_{i_p},$$

ибо произведение разных атомов равно нулю.

Правая часть равенства не 0, ибо левая часть (атом) не 0. Правая часть не 0, если только $a_i = a_{j_s}$ при некотором $s \in \{1, 2, \ldots, l\}$. Тогда $a_{i_{pp}} \in A''$ и $A' \subseteq^p A''_s$. Аналогично показываем $A'' \subseteq A'$. Отсюда $A' = A''$. Противоречие с $A' \neq A''$. Следовательно, $\varphi(A') \neq \varphi(A'')$. Итак,

$$\forall A', A'' \subseteq U \ (A' \neq A'' \rightarrow \varphi(A') \neq \varphi(A'')).$$

Таким образом, отображение φ взаимно однозначно.

Покажем, что отображение φ сохраняет булевы операции. Для этого достаточно показать, что отображение φ сохраняет отношение частичного порядка. Пусть $A' \subseteq A''$. Тогда

$$\{a_{i_1}, \ldots, a_{i_k}\} \subseteq \{a_{j_1}, \ldots, a_{j_l}\} \longleftrightarrow$$
$$(a_{i_1} \vee \ldots \vee a_{i_k}) \vee (a_{j_1} \vee \ldots, \vee a_{j_l}) = a_{j_1} \vee \ldots, \vee a_{j_l},$$

что для атомов очевидно.

Замечание. 1. Конечные булевы алгебры изоморфны, если и только если они равномощны.

2. Число элементов конечной булевой алгебры равно степени двойки.

3. Теорема Стоуна о представлении справедлива и для бесконечных булевых алгебр.

16. СИНТЕЗ СХЕМ ИЗ ФУНКЦИОНАЛЬНЫХ ЭЛЕМЕНТОВ

16.1. Схема из функциональных элементов

В современной технике среди управляющих устройств важное место занимают дискретные преобразователи (ДП) с несколькими входами и несколькими выходами (рис.16.1).

На входы ДП в дискретные моменты времени $0, 1, \ldots, t, \ldots$ поступают сигналы некоторого конечного входного алфавита X, ДП их обрабатывает и выдает на выходах сигналы некоторого

Рис.16.1

конечного выходного алфавита Y. Бывают дискретные преобразователи с памятью и без памяти. ДП с (конечной) памятью реализуют конечные автоматы. ДП без памяти реализуют функции. В дальнейшем мы будем рассматривать дискретные преобразователи без памяти. Если символы входного и выходного алфавита есть $E_k = \{0, 1, \ldots, k-1\}$, то ДП реализует функции k-значной логики алгебры логики. Их p для ДП с p выходами, и одна для ДП с одним выходом. Если символы входного и выходного алфавита есть $E_2 = \{0, 1\}$, то ДП реализует функции алгебры логики. Такие ДП называются функциональными элементами. Везде далее будем говорить о функциях алгебры логики

Пусть имеем конечное множество $F = \{F_1, \ldots, F_r\}$ из r функциональных элементов. Если входам и выходам элементов приписать некоторые переменные (рис.16.2), то элементы реализуют r функций алгебры логики $f_i(x_1, \ldots, x_{n_i})$, $i = 1, 2, \ldots, r$.

Определение (логической сети из функциональных элементов системы F).

1. Одна изолированная вершина (одновременно вход и выход) есть сеть.

2. *Объединение сетей*. Пусть S' и S'' есть две сети без общих элементов, имеющие n и m входов и p и q выходов соответственно, тогда *объединение* сетей S' и S'' есть сеть S, входы которой есть $n+m$ входов сетей S' и S'', а выходы есть $p+q$ выходов сетей S' и S'' (рис.16.3).

3. *Подключение элемента к сети*. Пусть S' есть логическая сеть с n входами и p выходами и пусть F_i есть элемент с числом входов $n_i \leqslant p$. Тогда результат *подключения* всех n_i входов элемента F_i к некоторым n_i выходам $j_1, \ldots, j_{n_i}$ сети S' есть сеть S, входы которой есть входы сети S'. Выходы сети S есть выходы сети S', кроме выходов $j_1, \ldots, j_{n_i}$, а также выход элемента F_i (рис.16.4).

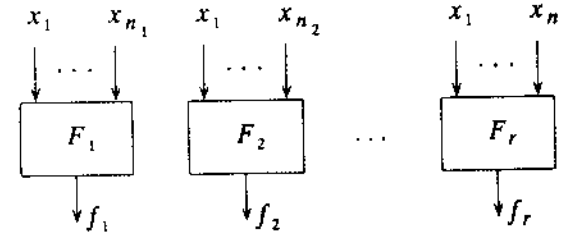


Рис.16.2

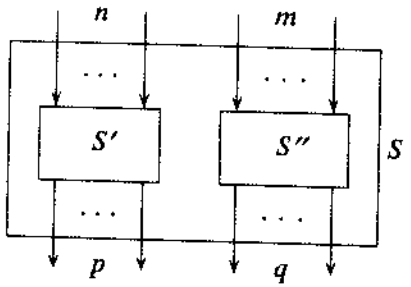


Рис.16.3

4. *Расщепление выходов сети.* Пусть S' есть сеть с выходами $1,2,\ldots,j,\ldots,p$. Тогда результат *расщепления* выхода j в сети S' есть сеть S с выходами $1,2,\ldots,j-1,j+1,\ldots,p$ сети S' и еще два выхода. Входы сети S это входы сети S' (рис.16.5).

Определение. *Схема из функциональных элементов* есть логическая сеть, входам и выходам которой приписаны различные буквы (переменные).

Обозначение. $S(x_1,\ldots x_n;y_1,..,y_p)$ есть схема с входами $x_1,\ldots x_n$ и выходами $y_1,y_2,\ldots,y_p$. Схема S реализует p функций $y_k=y_k(x_1,\ldots x_n)$, $k=1,\ldots,p$.

Определение. Множество $F=\{F_1,\ldots,F_r\}$ из r элементов функционально полно, если для любой функции f можно построить схему, которая реализует f.

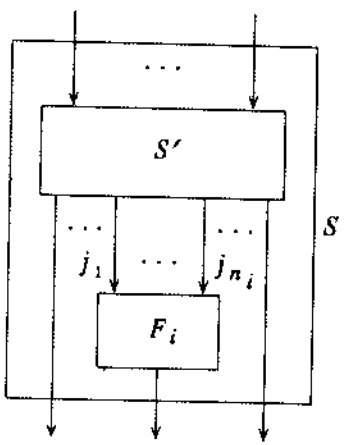


Рис.16.4

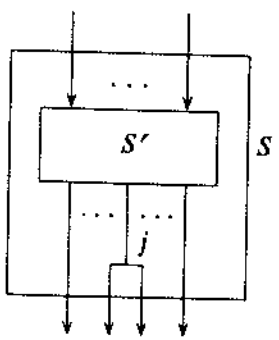


Рис.16.5

Замечание. Множество F функционально полно, если таковой является система реализуемых ими функций.

16.2. Функции Шеннона

Пусть $F=\{F_1,\ldots,F_r\}$ есть функционально полная система элементов.

Сложность $L(S)$ схемы S из функциональных элементов системы F есть число входящих в схему элементов.

Пусть $L(f)$ есть минимальная сложность схемы, реализующей функцию f, то есть $L(f)=\min L(S)$ по всем схемам S, реализующим функцию f.

Пусть $L(n)$ есть максимальная сложность минимальных реализаций всех функций n переменных, то есть

$L(n)=\max L(f)$ по всем функциям из $P_2^{(n)}$.

Пусть A есть алгоритм, с помощью которого можно построить любую функцию из функционально полного набора элементов F.

Пусть $L_A(f)$ есть сложность схемы, реализующей функцию f и построенной для f с помощью алгоритма A.

Пусть $L_A(n)=\max L_A(f)$ по всем функциям из $P_2^{(n)}$.

Функции $L(n)$ и $L_A(n)$ называются функциями Шеннона.

Очевидно, что $L_A(n)\geqslant L(n)$.

16.3. Элементарные методы синтеза схем

Наиболее полно исследованы вопросы синтеза и сложности схем, построенных из функционально полного набора элементов, реализующих функции $x\&y$, $x\vee y$, $\neg x$. Рассмотрим метод (алгоритм) синтеза, основанный на представлении функции с помощью СДНФ. Именно, пусть

$$f(x_1,\ldots,x_n)=\bigvee_{f(c_1,\ldots,c_n)=1} x_1^{c_1}\ldots x_n^{c_n}.$$

Пусть $K=x_1^{c_1}\ldots x_n^{c_n}$. Элементарную конъюнкцию K можно реализовать схемой S_k, изображенной на рис.16.6. Схема S_K имеет $n-1$ конъюнкторов и не более n отрицаний (инверторов). Всего не более $(n-1)+n$ элементов. Схема для элементарной конъюнкции K имеет сложность реализации $L(S_K)\leqslant 2n-1$.

Все элементарные конъюнкции $K_1,\ldots,K_u$ из представления функции f можно реализовать схемой G_K, изображенной на рис. 16.7. Схема G_K имеет сложность $L(G_k)\leqslant u(2n-1)$.

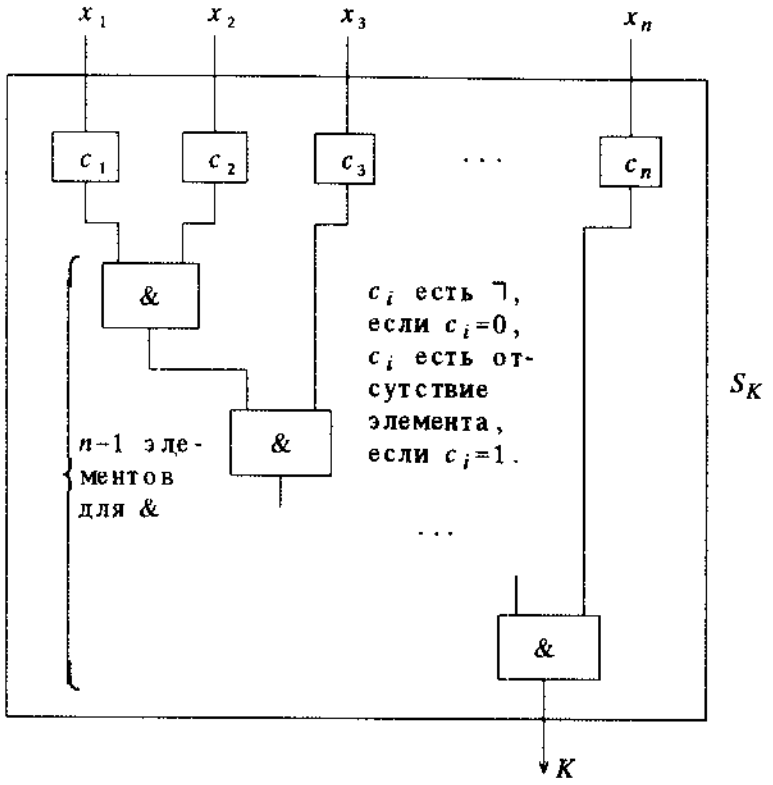


Рис.16.6

Далее устраиваем дизъюнкцию выходов схемы G_K, как это сделано на схеме S на рис.16.8.

Схема S реализует функцию f. Схема S имеет $u-1$ дизъюнкторов. Схема S имеет сложность $L(f) \leqslant u(2n-1)+(u-1) = u\cdot 2n-1$. В худшем случае $u=2^n$. Тогда сложность $L(n) \leqslant 2^n \cdot 2n-1 = n\,2^{n+1} - 1$. Для описанного алгоритма A функция Шеннона $L_A(n) \leqslant n\,2^{n+1}-1$.

Теорема (Лупанова). Для схем из функциональных элементов для $x\&y, x\vee y, \neg x$ функция Шеннона $L(n) \leqslant \frac{2^n}{n}\left\{1 + O\left(\frac{\log_2 n}{n}\right)\right\}$, $L(n) \gtrsim \frac{2^n}{n}$, где $a_n \gtrsim b_n$ означает, что $\overline{\lim\limits_{n\to\infty}} \frac{a_n}{b_n} \geqslant 1$. То есть асимптотически $L(n) \asymp 2^n/n$.

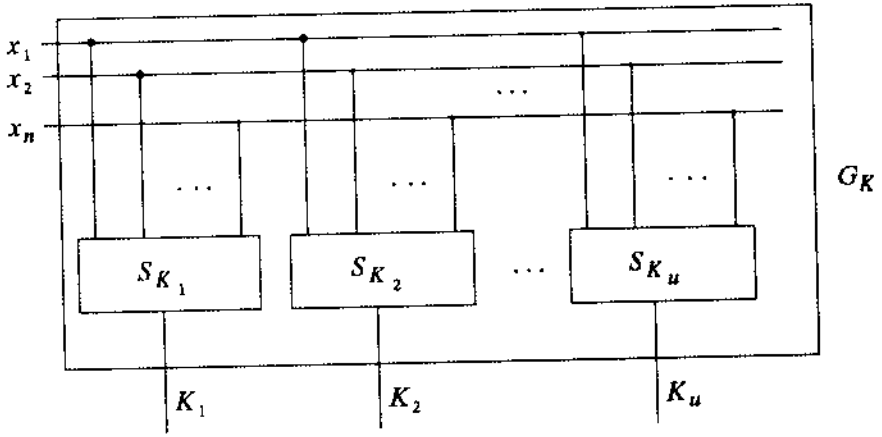


Рис.16.7

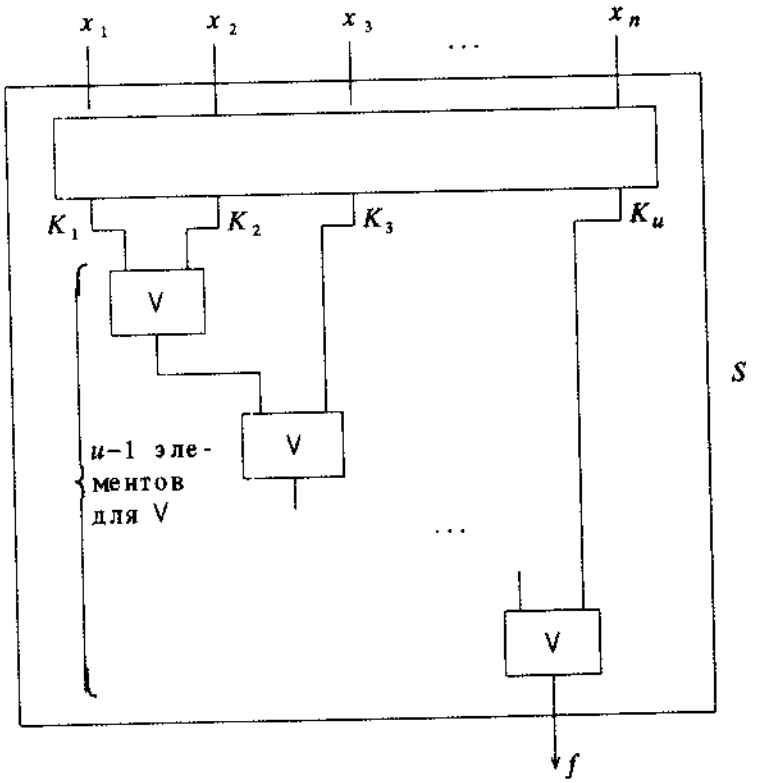


Рис.16.8

16.4. Синтез мультиплексоров

Определение. *Мультиплексор* MUX есть большая интегральная схема (БИС) $M(n)$, имеющая n управляющих входов, 2^n информационных входов и один выход. Для поданного на управляющие входы набора $(c_1,\ldots,c_n)$ из 0 и 1 схема делает проходимым (отпирает) и пропускает сигнал единственного информационного входа, помеченного набором $(c_1,\ldots,c_n)$; остальные информационные входы заперты и к выходу не проходимы (рис.16.9).

Теорема. Мультиплексор может быть реализован схемой, содержащей $3\cdot 2^n-3$ конъюнкций и дизъюнкций.

Доказательство. Индукция по n.

Базис. Схема $M(1)$ изображена на рис.16.10.

Схема $M(1)$ имеет $3\cdot 2^1-3 = 3$ конъюнкций и дизъюнкций.

Предположение индукции. Пусть мультиплексор $M(n-1)$ построен и имеет $3\cdot 2^{n-1}-3$ конъюнкций и дизъюнкций.

Шаг индукции. Мультиплексор $M(n)$ изображен на рис.16.11. Схема $M(n)$ имеет $2\cdot(3\cdot 2^{n-1}-3)+3 = 3\cdot 2^n-3$ конъюнкций и дизъюнкций. Шаг индукции установлен. Теорема доказана.

Утверждение. С помощью мультиплексора $M(n)$, инвертора, генератора нулей и единиц можно реализовать любую функцию алгебры логики от $n+1$ переменных.

Доказательство. По теореме о разложении функции по составляющим

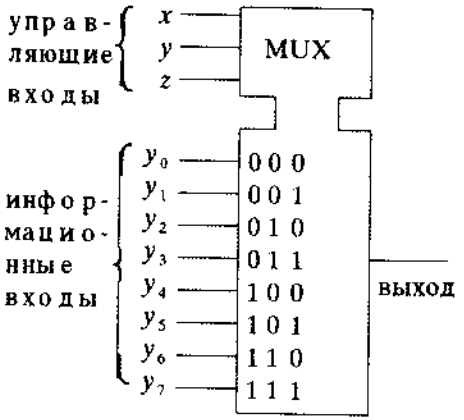


Рис.16.9

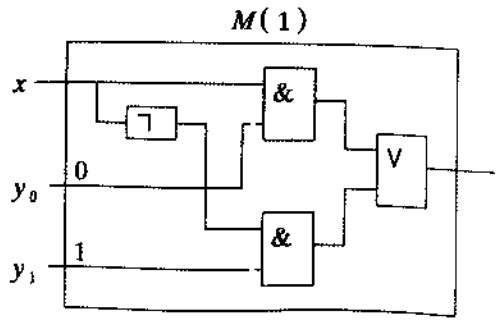


Рис.16.10

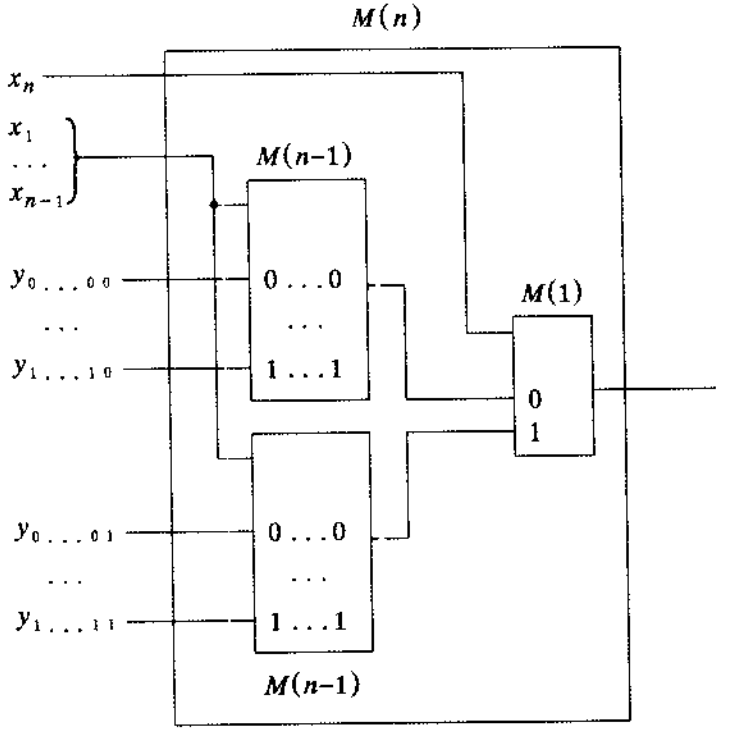


Рис.16.11

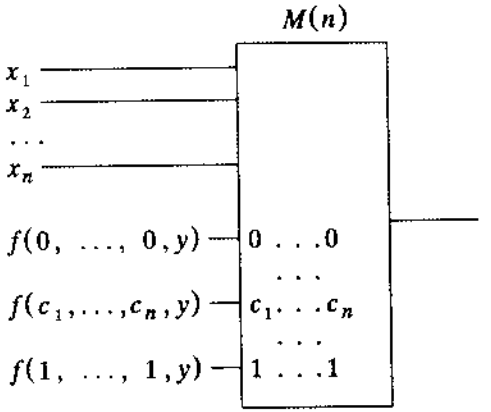


Рис.16.12

$$f(x_1,\ldots,x_n,y) = \bigvee_{(c_1,\ldots,c_n)\in E_2^n} x_1^{c_1}\ldots x_n^{c_n} f(c_1,\ldots,c_n,y).$$

Как функция одной переменной y функция $f(c_1,\ldots,c_n,y) \in \{0,1,y,\neg y\}$. Конъюнкция $x_1^{c_1}\ldots x_n^{c_n}$ равна 1 на единственном наборе $(c_1,\ldots,c_n)$. Тогда функция f может быть реализована с помощью мультиплексора так, как это изображено на рис.16.12.

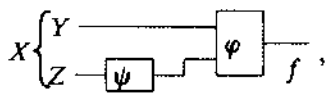
Пример. Реализовать функцию трех переменных $f(x,y,z) = 00101101$ с помощью мультиплексора $M(2)$.

В задаче используется разложение функции по переменным x,y. Требуемая реализация приведена на рис.16.13.

16.5. Элементы функциональной декомпозиции

Разобьем множество $X=\{x_1,\ldots,x_n\}$ из n переменных на два непересекающихся подмножества $Y=\{y_1,\ldots,y_m\}$, $Z=(z_{m+1},\ldots,z_n\}$ в сумме (в объединении) дающих все множество X.

Определение. *Простая непересекающаяся декомпозиция* функции $f(x_1,\ldots,x_n)$ есть ее представление в виде $f(X) = \varphi(Y,\psi(Z))$ при некоторых функциях φ и ψ.

Замечание. В случае декомпозиции функция $f(X)$ может быть реализована схемой 

, построенной из более простых функций φ и ψ.

В последующем для простоты будем считать, что $Y=\{x_1,\ldots,x_m\}$, $Z=\{x_{m+1},\ldots,x_n\}$.

x	y	z	f	$f(x,y,z) =$		z	$\bar z$	0	1	$M(2)$
0	0	0	0	$\bar x\bar y\cdot f(0,0,z)\vee$	$\bar x\bar y\cdot 0\vee$					x
0	0	1	0							y
0	1	0	1	$\bar x y\cdot f(0,1,z)\vee$	$\bar x y\cdot\bar z\vee$					00
0	1	1	0							01
1	0	0	1	$x\bar y\cdot f(1,0,z)\vee$	$x\bar y\cdot 1\vee$					10
1	0	1	1							11
1	1	0	0	$xy\cdot f(1,1,z) =$	$xy\cdot z.$					f
1	1	1	1							

Рис.16.13

Определение. Y-компонента функции $f(Y,Z)$ есть совокупность функций $\{f(c_1,\ldots,c_m,x_{m+1},\ldots,x_n) : (c_1,\ldots,c_m) \in E_2^m\}$. Z-компонента функции $f(Y,Z)$ есть совокупность функций $\{f(x_1,\ldots,x_m,c_{m+1},\ldots,c_n) : (c_{m+1},\ldots,c_n) \in E_2^{n-m}\}$.

Пример. $f(x_1,x_2,x_3,x_4)$, $Y=\{x_1,x_2\}$, $Z=\{x_3,x_4\}$ (рис.16.14).

Столбцы таблицы на рис.16.14 составляют Z-компоненту $\{f(0,0,x_3,x_4)=1001,\ f(0,1,x_3,x_4)=0110,\ f(1,1,x_3,x_4)=1101,\ f(1,0,x_3,x_4)=1011\}$ функции f.

Теорема 1. Простая непересекающаяся декомпозиция $f(X) = \varphi(Y,\psi(Z))$ для функции $f(X)$ существует $\longleftrightarrow$ всякая функция у ее Y-компоненты $f(c_1,\ldots,c_m,Z)\in\{0,1,\psi(Z),\neg\psi(Z)\}$.

Доказательство. *Достаточность*. Пусть $f(c_1,\ldots,c_m,Z)\in\{0,1,\psi(Z),\neg\psi(Z)\}\ \forall (c_1,\ldots,c_m)\in E_2^m$. Пусть $C=(c_1,\ldots,c_m)$, $Y^c = x_1^{c_1}\ldots x_m^{c_m}$. Построим функцию $\varphi(Y,u)$ следующим образом.

$$f(\underbrace{x_1,\ldots,x_m}_{Y},\ \underbrace{x_{m+1},\ldots,x_n}_{Z}) =$$

$$\bigvee_{(c_1,\ldots,c_m)\in E_2^m} \underbrace{x_1^{c_1}\ldots x_m^{c_m}}_{Y^C} f(\underbrace{c_1,\ldots,c_m}_{C},\underbrace{x_{m+1},\ldots,x_n}_{Z}),$$

где $f(C,Z)$ есть одна из Y-компонент в $\{0,1,\psi(Z),\neg\psi(Z)\}$. Группируем слагаемые относительно $0,1,\psi(Z),\neg\psi(Z)$.

$$f(Y,Z) = \left\{\bigvee_{f(C,Z)\equiv 0} Y^C\right\}\cdot 0 \vee \underbrace{\left\{\bigvee_{f(C,Z)\equiv 1} Y^C\right\}}_{h_1(Y)}\cdot 1 \vee$$

x_1x_2 \ x_3x_4	00	01	11	10	
00	1	0	1	1	$f(0,0,x_3,x_4)$
01	0	1	1	0	$f(0,1,x_3,x_4)$
11	0	1	0	1	$f(1,1,x_3,x_4)$
10	1	1	1	1	$f(1,0,x_3,x_4)$

Y-компонента

Рис.16.14

$$\Bigg(\underbrace{\bigvee_{f(C,Z)\equiv\psi(Z)} Y^C}_{h_2(Y)}\Bigg)\cdot\psi(Z)\ \vee\ \Bigg(\underbrace{\bigvee_{f(C,Z)\equiv\neg\psi(Z)} Y^C}_{h_3(Y)}\Bigg)\cdot\overline{\psi(Z)}\ =$$

$$h_1(Y)\ \vee\ h_2(Y)\cdot\psi(Z)\ \vee\ h_3(Y)\cdot\overline{\psi(Z)}.$$

Положим $\varphi(Y,u) = h_1(Y) \vee h_2(Y)\cdot u \vee h_3(Y)\cdot\bar{u}$. Тогда $f(Y,Z) = \varphi(Y,\psi(Z))$, то есть функция f допускает простую непересекающуюся декомпозицию.

Необходимость. Пусть функция $f(Y,Z)$ допускает простую непересекающуюся декомпозицию $f(Y,Z) = \varphi(Y,\psi(Z))$ при некоторых функциях φ и ψ. Разложим функцию f по переменным Y.

$$f(\underbrace{x_1,\ldots,x_m}_{X},\underbrace{x_{m+1},\ldots,x_n}_{Z}) = \varphi(\underbrace{x_1,\ldots,x_m}_{X},\psi(\underbrace{x_{m+1},\ldots,x_n}_{Z})) =$$

$$\bigvee_{C=(c_1,\ldots,c_m)\in E_2^m} \underbrace{x_1^{c_1}\ldots x_m^{c_m}}_{Y^C}\ \underbrace{\varphi(c_1,\ldots,c_m,\psi(x_{m+1},\ldots,x_n))}_{\varphi(C,\psi(Z))} =$$

$$\Bigg(\bigvee_{C\in E_2^m} Y^C\cdot\varphi(C,u)\Bigg)\Bigg|_{u=\psi(Z)} =$$

[Как функция одной переменной u, функция $\varphi(C,u) \in \{0,1,u,\bar{u}\}$. Группируем слагаемые относительно $0,1,u,\bar{u}$]

$$\Bigg(\bigvee_{\varphi(C,u)\equiv 0} Y^C\cdot\underbrace{\varphi(C,u)}_{\equiv 0} \vee \bigvee_{\varphi(C,u)\equiv 1} Y^C\cdot\underbrace{\varphi(C,u)}_{\equiv 1} \vee$$

$$\bigvee_{\varphi(C,u)\equiv u} Y^C\cdot\underbrace{\varphi(C,u)}_{\equiv u} \vee \bigvee_{\varphi(C,u)\equiv\neg u} Y^C\cdot\underbrace{\varphi(C,u)}_{\equiv\neg u}\Bigg)\Bigg|_{u=\psi(Z)} =$$

$$\bigvee_{\varphi(C,u)\equiv 0} Y^C\cdot\underbrace{\varphi(C,\psi(Z))}_{f(C,Z)\in Y\text{-комп}} \vee \bigvee_{\varphi(C,u)\equiv 1} Y^C\cdot\underbrace{\varphi(C,\psi(Z))}_{f(C,Z)\in Y\text{-комп}} \vee$$

$$\bigvee_{\varphi(C,u)\equiv u} Y^C\cdot\underbrace{\varphi(C,\psi(Z))}_{\equiv\psi(Z)} \vee \bigvee_{\varphi(C,u)\equiv\neg u} Y^C\cdot\underbrace{\varphi(C,\psi(Z))}_{\equiv\neg\psi(Z)} =$$

$f(C,Z)\in Y$-комп

$$\left[\begin{array}{c}\text{объединим все слагаемые}\\ \text{одной дизъюнкцией}\end{array}\right] = \bigvee_{C\in E_2^m} Y^C\cdot\underbrace{f(C,Z)}_{\in\{0,1,\psi(Z),\neg\psi(Z)\}} =$$

$$\bigvee_{(c_1,\ldots,c_m)\in E_2^m} x_1^{c_1}\ldots x_m^{c_m}\cdot\underbrace{f(c_1,\ldots,c_m,x_{m+1},\ldots,x_n)}_{\in\{0,1,\psi(Z),\neg\psi(Z)\}},$$

то есть всякая функция $f(c_1,\ldots,c_m,x_{m+1},\ldots,x_n)$ из Y-компоненты лежит в $\{0,1,\psi(Z),\neg\psi(Z)\}$. Теорема доказана.

Пример 1. Найти простую непересекающуюся декомпозицию функции $f(x_1,x_2,x_3,x_4)$, $Y=\{x_1,x_2\}$, $Z=\{x_3,x_4\}$ (рис.16.15).

Все функции Y-компоненты (по строкам) лежат в $\{0,1,\psi(x_3,x_4),\neg\psi(x_3,x_4)\}$. Функция f допускает простую непересекающуюся декомпозицию $f(x_1,x_2,.x_3,x_4) =$

$$f(0,0,x_3,x_4)\cdot\bar{x}_1\bar{x}_2 \vee f(0,1,x_3,x_4)\cdot\bar{x}_1x_2 \vee$$

$$f(1,0,x_3,x_4)\cdot x_1\bar{x}_2 \vee f(1,1,x_3,x_4)\cdot x_1x_2 =$$

$$\bar{x}_1\bar{x}_2\cdot\psi(x_3,x_4) \vee \bar{x}_1x_2\cdot 0 \vee x_1\bar{x}_2\cdot\overline{\psi(x_3,x_4)} \vee x_1x_2\cdot\psi(x_3,x_4) =$$

$$(\bar{x}_1x_2\cdot u \vee x_1\bar{x}_2\cdot\bar{u} \vee x_1x_2\cdot u)\big|_{u=\psi(x_3,x_4)} =$$

$\varphi(x_1,x_2,\psi(x_3,x_4))$, где $\varphi(x_1,x_2,u) = \bar{x}_1x_2u \vee x_1\bar{x}_2\bar{u} \vee x_1x_2u$,

$\psi(x_3,x_4) = 1011 = \bar{x}_3\bar{x}_4 \vee x_3\bar{x}_4 \vee x_3x_4$.

	x_3x_4 0 0	x_3x_4 0 1	x_3x_4 1 1	x_3x_4 1 0	Функции Y-компоненты
x_1x_2 0 0	1	0	1	1	$f(0,0,x_3,x_4)=\psi(x_3,x_4)$
0 1	0	0	0	0	$f(0,1,x_3,x_4)=0$
1 1	1	0	1	1	$f(1,1,x_3,x_4)=\psi(x_3,x_4)$
1 0	0	1	0	0	$f(1,0,x_3,x_4)=\neg\psi(x_3,x_4)$

Рис.16.15

Реализация функции

$f = \bar{x}_1\bar{x}_2 \cdot \psi(x_3,x_4) \vee \bar{x}_1 x_2 \cdot 0 \vee x_1\bar{x}_2 \cdot \overline{\psi(x_3,x_4)} \vee x_1 x_2 \cdot \psi(x_3,x_4)$ с помощью мультиплексора приведена на рис.16.16.

Пример 2. Найти простую непересекающуюся декомпозицию функции $f(x,x_2,x_3,x_4)$ (рис.16.17.).

Функция f простую непересекающуюся декомпозицию не допускает, ибо функции ее Y-компонент не лежат в $\{0,1,\psi(x_3,x_4), \neg\psi(x_3,x_4)\}$.

Замечание. Пусть непересекающиеся множества наборов $G_1 \subseteq E_2^n$, $G_2 \subseteq E_2^n$ в объединении дают E_2^n. Пусть функции

$$g_1(x_1,\ldots,x_n) = \bigvee_{(c_1,\ldots,c_n)\in G_1} x_1^{c_1}\ldots x_n^{c_n},$$

$$g_2(x_1,\ldots,x_n) = \bigvee_{(c_1,\ldots,c_n)\in G_2} x_1^{c_1}\ldots x_n^{c_n}.$$

Тогда $g_1(x_1,\ldots,x_n) = \neg g_2(x_1,\ldots,x_n)$.

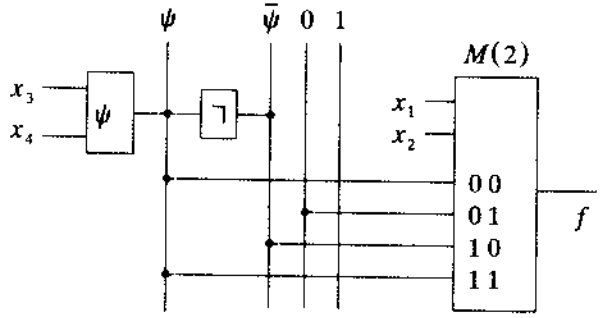


Рис.16.16

	x_3x_4	x_3x_4	x_3x_4	x_3x_4	Функции Y-компоненты
x_1x_2	0 0	0 1	1 1	1 0	
0 0	0	1	0	0	$f(0,0,x_3,x_4)=\psi(x_3,x_4)$
0 1	0	0	0	0	$f(0,1,x_3,x_4)=0$
1 1	0	1	1	0	$f(1,1,x_3,x_4)=\psi_1(x_3,x_4)$
1 0	1	1	1	1	$f(1,0,x_3,x_4)=1$

Рис.16.17

Пусть $X=\{x_1,\ldots,x_n\}$, $Y=\{x_1,\ldots,x_m\}$, $Z=\{x_{m+1},\ldots,x_n\}$.

Теорема 2. Функция $f(X)$ допускает простую непересекающуюся декомпозицию $f(Y,Z)=\varphi(Y,\psi(Z)) \leftrightarrow$ функция f имеет в Z-компоненте не более двух различных функций.

Доказательство. *Необходимость*. Пусть простая непересекающаяся декомпозиция $f(Y,Z)=\varphi(Y,\psi(Z))$. Разложим функцию f по переменным Z. Тогда $f(\underbrace{x_1,\ldots,x_m}_{Y},\underbrace{x_{m+1},\ldots,x_n}_{Z}) =$

$$\bigvee_{C=(c_{m+1},\ldots,c_n)\in E_2^{n-m}} x_{m+1}^{c_{m+1}}\ldots x_n^{c_n} f(x_1,\ldots,x_m,c_{m+1},\ldots,c_n).$$

Z-компонента $f(x_1,\ldots,x_m,c_{m+1},\ldots,c_n) = \varphi(x_1,\ldots,x_m,\psi(c_{m+1},\ldots,c_n)) = \varphi(Y,\psi(C))$ есть $\varphi(Y,0)$ или $\varphi(Y,1)$, то есть функция f имеет не более двух функций $\varphi(Y,0)$, $\varphi(Y,1)$ в своей Z-компоненте.

Достаточность. Пусть функция f имеет не более двух функций $h_1(Y)$, $h_2(Y)$ в своей Z-компоненте. Разложим функцию f по переменным Z: $f(\underbrace{x_1,\ldots,x_m}_{Y},\underbrace{x_{m+1},\ldots,x_n}_{Z}) =$

$$\bigvee_{C=(c_{m+1},\ldots,c_n)\in E_2^{n-m}} \underbrace{x_{m+1}^{c_{m+1}}\ldots x_n^{c_n}}_{Z^c} \underbrace{f(x_1,\ldots,x_m,c_{m+1},\ldots,c_n)}_{h_1(Y)\text{ или }h_2(Y)} =$$

[группируем слагаемые относительно $h_1(Y)$ и $h_2(Y)$]

$$\underbrace{\left\{\bigvee_{f(Y,C)=h_1(Y)} Z^c\right\}}_{g_1(Z)} \cdot h_1(Y) \vee \underbrace{\left\{\bigvee_{f(Y,C)=h_2(Y)} Z^c\right\}}_{g_2(Z)} \cdot h_2(Y) =$$

$$g_1(Z)\cdot h_1(Y) \vee g_2(Z)\cdot h_2(Y) =$$

[По выше приведенному замечанию $g_2(Z) = \overline{g_1(Z)}$. Положим $\psi(Z) = g_1(Z)$.]

$$g_1(Z)\cdot h_1(Y) \vee \overline{g_1(Z)}\cdot h_2(Y).$$

Положим $\varphi(Y,u) = u\cdot h_1(Y) \vee \bar{u}\cdot h_2(Y)$. Тогда $\varphi(Y,\psi(Z)) = \psi(Z)\cdot h_1(Y) \vee \overline{\psi(Z)}\cdot h_2(Y) = g_1(Z)\cdot h_1(Y) \vee \overline{g_1(Z)}\cdot h_2(Y) =$

$f(Y,Z)$, то есть функция f допускает простую непересекающуюся декомпозицию. Теорема доказана.

Пример 1. Найти простую непересекающуюся декомпозицию функции $f(x,x_2,x_3,x_4)$, $Y=\{x_1,x_2\}$, $Z=\{x_3,x_4\}$ (рис.16.18).

$$h_1(x_1,x_2)\bar{x}_3\bar{x}_4 \vee h_2(x_1,x_2)\bar{x}_3x_4 \vee h_1(x_1,x_2)x_3\bar{x}_4 \vee h_1(x_1,x_2)x_3x_4 =$$

$$h_1(x_1,x_2)\cdot(\underbrace{\bar{x}_3\bar{x}_4 \vee x_3\bar{x}_4 \vee x_3x_4}_{\psi(x_3,x_4)}) \vee h_2(x_1,x_2)\cdot \underbrace{\bar{x}_3x_4}_{\overline{\psi(x_3,x_4)}} =$$

$$h_1(x_1,x_2)\cdot\psi(x_3,x_4) \vee h_2(x_1,x_2)\cdot\overline{\psi(x_3,x_4)} =$$

$$(\underbrace{h_1(x_1,x_2)\cdot u \vee h_2(x_1,x_2)\cdot\bar{u}}_{\varphi(Y,u)})\Big|_{u=\psi(x_3,x_4)} = \varphi(Y,\psi(x_3,x_4)).$$

Функция f допускает простую непересекающуюся декомпозицию $f(Y,Z) = \varphi(Y,\psi(Z))$. Реализация функции

$$f = h_1(x_1,x_2)\bar{x}_3\bar{x}_4 \vee h_2(x_1,x_2)\bar{x}_3x_4 \vee h_1(x_1,x_2)x_3\bar{x}_4 \vee h_1(x_1,x_2)x_3x_4$$

с помощью мультиплексора приведена на рис.16.19.

16.6. Обнаружение неисправностей в схемах

Пусть мы конструируем схему из функциональных элементов для данной функции. В результате брака схема может реализовать другую функцию. Задача контроля состоит в том, чтобы определить

$$f(Y,Z) = f(x,x_2,x_3,x_4) =$$

$$f(x_1,x_2,0,0)\cdot\bar{x}_3\bar{x}_4 \vee f(x_1,x_2,0,1)\cdot\bar{x}_3x_4 \vee$$

$$f(x_1,x_2,1,0)\cdot x_3\bar{x}_4 \vee f(x_1,x_2,1,1)\cdot x_3x_4 =$$

x_1x_2	x_3x_4 0 0	x_3x_4 0 1	x_3x_4 1 1	x_3x_4 1 0	Функции Z-компоненты
0 0	0	1	0	0	$f(x_1,x_2,0,0)=h_1(x_1,x_2)$
0 1	1	0	1	1	$f(x_1,x_2,0,1)=h_2(x_1,x_2)$
1 1	0	0	0	0	$f(x_1,x_2,1,1)=h_1(x_1,x_2)$
1 0	0	0	0	0	$f(x_1,x_2,1,0)=h_1(x_1,x_2)$

Рис.16.18

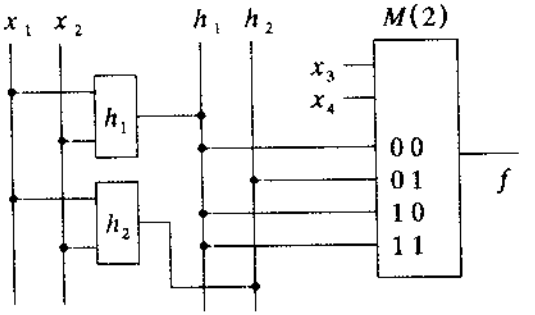


Рис.16.19

1) исправна ли схема, то есть реализует ли схема данную функцию,

2) какую функцию реализует схема в случае неисправности.

Обнаружение неисправностей в схеме (тестирование схемы) есть важная и непростая задача. Для знакомства с ней мы рассмотрим следующий простой случай.

Пусть мы конструируем схему S из функциональных элементов для функции $f_0(x_1,\ldots,x_n)$. В результате брака схема S может реализовать другую функцию из числа функций $f_1,f_2,\ldots,f_r$ от n переменных. Задача контроля состоит в том, чтобы определить 1) исправна ли схема, то есть реализует ли схема функцию f_0, 2) какую функцию из $f_1,f_2,\ldots,f_r$ реализует схема в случае неисправности (Табл.16.1).

Определение. *Тест* (тестовый набор) для таблично заданной функции $f_0(x_1,\ldots,x_n)$ есть совокупность наборов длины n из 0 и 1 (совокупность строк в таблице функции f), которая высекает из столбцов значения функций $f_0,f_1,\ldots,f_r$ столбцы со следующими свойствами.

1) все высекаемые столбцы различны и тогда тест полный;

Таблица 16.1

$x_1,\ldots,x_{n-1},x_n$	f_0	f_1	f_2	$\ldots$	f_r
0 ... 0 0	α_0	β_0	γ_0	$\ldots$	δ_0
0 ... 0 1	α_1	β_1	γ_1	$\ldots$	δ_1
...				...	
a_1 ... $a_{n-1}a_n$	α_p	β_p	γ_p	$\ldots$	δ_p
...				...	
1 ... 1 1	α_z	β_z	γ_z	$\ldots$	δ_z

где $z = 2^n-1$.

2) все высекаемые столбцы отличны от столбца для f_0 и тогда тест проверяющий.

Замечание. 1. Полный тест есть проверяющий тест.

2. Множество всех наборов длины n есть полный тест.

Определение. *Сложность теста* есть число входящих в него наборов.

Определение. Тест *минимальный* (*тупиковый*), если удаление из него любого набора приводит к совокупности наборов, которая тестом уже не является. Тест *наименьший*, если он имеет наименьшую сложность.

Замечание. Все наименьшие тесты находятся среди тупиковых тестов.

Пример построения всех тупиковых и наименьших тестов

Пусть строится схема для функции $f(x_1,\ldots,x_n)$. Примем список следующих неисправностей.

s_{ij} – слипание входов x_i и x_j.

Тогда реализуется функция

$$g(\ldots,x_i,\ldots,x_j,\ldots) = f(\ldots,x_i \vee x_j,\ldots,x_i \vee x_j,\ldots).$$

0_i – обрыв входа x_i.

Тогда реализуется функция $g(\ldots,x_i,\ldots) = f(\ldots,0,\ldots)$.

1_i – замыкание входа x_i.

Тогда реализуется функция $g(\ldots,x_i,\ldots) = f(\ldots,1,\ldots)$.

Построим тесты для функции f_0=10110110 для группы неисправностей $\{0_3, s_{12}\}$. Возможны только указанные неисправности, причем каждая схема может иметь только одну неисправность.

Для неисправности 0_3 функция $f_1(x_1,x_2,x_3) = f_0(x_1,x_2,0)$.

Для неисправности s_{12} функция $f_2(x_1,x_2,x_3) = f_0(x_1 \vee x_2, x_1 \vee x_2, x_3)$.

Пусть $y_0,\ldots,y_7$ есть восемь наборов длины 3 из 0 и 1 (табл.16.2). Каждый тест имеет вид $y_{i_1} \& y_{i_2} \& \ldots \& y_{i_p}$.

Построим все тупиковые проверяющие тесты.

f_0 от f_1 отличают строки y_1, y_5, y_7.

Положим $D_{f_0,f_1} = y_1 \vee y_5 \vee y_7$.

f_0 от f_2 отличают строки y_3, y_4, y_5.

Положим $D_{f_0,f_2} = y_3 \vee y_4 \vee y_5$.

Построим все тупиковые проверяющие тесты.

$D_{f_0,f_1} \& D_{f_0,f_2} = (y_1 \vee y_5 \vee y_7)(y_3 \vee y_4 \vee y_5) =$

Таблица 16.2

	x_1	x_2	x_3	f_0	f_1	f_2
y_0	0	0	0	1	1	1
y_1	0	0	1	0	1	0
y_2	0	1	0	1	1	1
y_3	0	1	1	1	1	0
y_4	1	0	0	0	0	1
y_5	1	0	1	1	0	0
y_6	1	1	0	1	1	1
y_7	1	1	1	0	1	0

$y_1y_3 \vee y_1y_4 \vee y_1y_5 \vee y_3y_5 \vee y_4y_5 \vee y_5 \vee y_3y_7 \vee y_4y_7 \vee y_5y_7 =$
$y_1y_3 \vee y_1y_4 \vee y_5 \vee y_3y_7 \vee y_4y_7.$

Получили пять тупиковых проверяющих тестов. Наименьший проверяющий тест есть y_5:

	x_1	x_2	x_3	f_0	f_1	f_2
y_5	1	0	1	1	0	0

Подаем на вход схемы набор 101. Если на выходе 1, то схема реализует функцию f_0. Если на выходе 0, то схема неисправна и функцию f_0 не реализует.

Построим все полные тупиковые тесты.

f_1 от f_2 отличают строки y_1, y_3, y_4, y_7.

Положим $D_{f_1,f_2} = y_1 \vee y_3 \vee y_4 \vee y_7$. Тогда

$$D_{f_0,f_1} \& D_{f_0,f_2} \& D_{f_1,f_2} =$$

$$(y_1 \vee y_5 \vee y_7)(y_3 \vee y_4 \vee y_5)(y_1 \vee y_3 \vee y_4 \vee y_7) =$$

$$(y_1y_3 \vee y_1y_4 \vee y_5 \vee y_3y_7 \vee y_4y_7)(y_1 \vee y_3 \vee y_4 \vee y_7) =$$

$$y_1y_3 \vee y_1y_3 \vee y_1y_5 \vee y_1y_3y_7 \vee y_1y_4y_7 \vee y_1y_3 \vee y_1y_3y_4 \vee y_3y_5 \vee y_3y_7 \vee$$

$$y_3y_4y_7 \vee y_1y_3y_4 \vee y_1y_4 \vee y_4y_5 \vee y_3y_4y_7 \vee y_4y_7 \vee y_1y_3y_7 \vee y_1y_4y_7 \vee$$

$$y_5y_7 \vee y_3y_7 \vee y_4y_7 = y_1y_3 \vee y_1y_4 \vee y_1y_5 \vee y_3y_5 \vee y_3y_7 \vee y_4y_5 \vee y_4y_7 \vee y_5y_7.$$

Получили 8 полных тупиковых тестов. Они все наименьшие. Например, y_1y_3:

	x_1	x_2	x_3	f_0	f_1	f_2
y_1	0	0	1	0	1	0
y_3	0	1	1	1	1	0

Подаем на вход схемы последовательно наборы y_1, y_3, то есть наборы $\begin{matrix}0&0&1\\0&1&1\end{matrix}$. Пара $\begin{matrix}0\\1\end{matrix}$ на выходе укажет, что схема реализует функцию f_0, пара $\begin{matrix}1\\1\end{matrix}$ – функцию f_1, пара $\begin{matrix}0\\0\end{matrix}$ – функцию f_2.

17. ЛОГИКА ПРЕДИКАТОВ

17.1. Предикаты, кванторы

Предикат есть функция, определенная на некотором множестве и принимающая одно из двух значений: истина (И) или ложь (Л). Все переменные предиката свободные.

Пусть $P(x,x_1,\ldots,x_k)$ есть $(k+1)$-местный предикат, определенный на множестве D. Запись $(\exists x)P(x,x_1,\ldots,x_k)$ будем понимать как "существует такой элемент x из D, для которого $P(x,x_1,\ldots,x_k)$ истинно"; $(\forall x)P(x,x_1,\ldots,x_k)$ – "для всех x из D $P(x,x_1,\ldots,x_k)$ истинно". Скажем, что предикаты

$$Q(x_1,\ldots,x_k) = (\exists x)P(x,x_1,\ldots,x_k),$$
$$R(x_1,\ldots,x_k) = (\forall x)P(x,x_1,\ldots,x_k)$$

получены из предиката $P(x,x_1,\ldots,x_k)$ навешиванием *квантора* соответственно существования и общности на предикат $P(x,x_1,\ldots,x_k)$ по переменной x. Переменная x в выражении $(Qx)P(x,x_1,\ldots,x_k)$ находится в области действия квантора (Qx). Навешивание квантора существования или квантора общности по некоторой переменной на $(k+1)$-местный предикат, содержащий эту переменную, приводит к другому предикату, местность которого на единицу меньше. Переменная x, находящаяся в области действия квантора (Qx), называется связанной. В противном случае переменная x свободна.

Из простых предикатов с помощью логических операций можно строить и более сложные предикаты. В случае сложного предиката $R(x)$ со свободной переменной x и возможно другими свободными переменными в предикате $(Qx)R(x)$ все свободные вхождения переменной x в R находятся в области действия квантора (Qx). Например, в выражении $(\forall x)(P(x) \vee (\exists x)R(x))$ переменная x в $P(x)$ находится в области действия квантора $(\forall x)$, а переменная x в $R(x)$ – в области действия квантора $(\exists x)$.

Пусть $P(x_1,\ldots,x_k)$ – предикат, определенный на множестве D; $\lceil P \rceil$ означает множество истинности предиката P, т.е. множество всех наборов $(a_1,\ldots,a_k)$ длины k элементов множества D, на которых предикат P принимает значение И.

Пусть $P(x_1,\ldots,x_k)$ и $R(x_1,\ldots,x_k)$ – два предиката, определенных на множестве D. Тогда $\lceil P \rceil \cup \lceil R \rceil$, $\lceil P \rceil \cap \lceil R \rceil$, $D^k - \lceil P \rceil$ есть множества истинности предикатов $P \vee R$, $P \,\&\, R$, $\neg P$ соответственно. Легкой модификацией это определение можно распространить на случай, когда местности предикатов P и Q и их переменные могут быть различными.

Пример. Пусть $D = \{0,1,2,\ldots\}$ есть множество натуральных чисел. Определим на множестве D следующие предикаты:

$Pr(x)$ – x есть простое число;
$Ev(x)$ – число x четно;
$Div(x,y)$ – число x делит число y;
$x = y$ – число x равно числу y;
$x \leqslant y$ – число x меньше или равно числу y;
$x = 2$ – число x совпадает с двойкой;
$Q(x,y)$ есть $\neg(x = 1) \,\&\, \neg(x = y) \,\&\, Div(x,y)$;
$P(y)$ есть $(\exists x)Q(x,y)$.

Тогда $(\exists x)Pr(x)$ есть истинное, $(\forall x)Pr(x)$ ложное, $P(2)$ ложное, $P(4)$ истинное высказывания.

Пусть $D = \{a_1,a_2,\ldots,a_k\}$ – конечное множество из k элементов и $P(x)$ – предикат, определенный на D.

Утверждение. Справедливы следующие равенства:

$$(\exists x)P(x) = P(a_1) \vee P(a_2) \vee \ldots \vee P(a_k).$$
$$(\forall x)P(x) = P(a_1) \,\&\, P(a_2) \,\&\, \ldots \,\&\, P(a_k).$$

Доказательство. Установим первое равенство. Пусть высказывание $(\exists x)P(x)$ истинно. Тогда существует $x = a_i \in D$, для которого $P(a_i)$ истинно. Отсюда $P(a_1) \vee \ldots \vee P(a_k)$ истинно.

Пусть $P(a_1) \vee \ldots \vee P(a_k)$ истинно. Тогда для некоторого i, $1 \leqslant i \leqslant k$, $P(a_i)$ истинно; поэтому существует $x = a_i \in D$, для которого $P(a_i)$ истинно, отсюда $(\exists x)P(x)$ истинно.

Установим второе равенство. Пусть $(\forall x)P(x)$ истинно. Тогда для всякого x из D $P(x)$ истинно, т.е. истинны $P(a_1),\ldots,P(a_k)$, откуда истинно $P(a_1) \,\&\ldots\&\, P(a_k)$. Пусть теперь $P(a_1) \,\&\, \ldots \,\&\, P(a_k)$ истинно, т.е. истинны $P(a_1),\ldots,P(a_k)$. Отсюда получаем, что для всякого x из D $P(x)$ истинно, т.е. $(\forall x)P(x)$ истинно. Утверждение доказано.

Аналогично доказываются следующие равенства.

$$(\exists x)P(x,x_1,\ldots,x_n) = \bigvee_{i=1}^{k} P(a_i,x_1,\ldots,x_n).$$

$$(\forall x)P(x,x_1,\ldots,x_n) = \mathop{\&}_{i=1}^{k} P(a_i,x_1,\ldots,x_n).$$

Переменная x может стоять на любом аргументном месте предиката P. Если кванторов несколько, то они элиминируются (устраняются) постепенно. Например,

$$(\forall x)(\exists y)P(x,y) = (\forall x)(\mathop{\vee}_{j=1}^{k} P(x,a_j)) = \mathop{\&}_{i=1}^{k} (\mathop{\vee}_{j=1}^{k} P(a_i,a_j)).$$

Замечание. Пусть D есть некоторое множество и множество $K \subseteq D^n$. Проекция K на оси $i_1,\ldots,i_p$ есть множество

$$\mathrm{Пр}_{i_1,\ldots,i_p} K = \{(x_{i_1},\ldots,x_{i_p}) \in D^p :$$
$$\exists(x_1,\ldots,x_{i_1},\ldots,x_{i_p},\ldots,x_n) \in K\}.$$

Проекция множества K на оси $i_1,\ldots,i_p$ может быть получена вычеркиванием (стиранием) во всех наборах $(x_1,\ldots,x_{i_1},\ldots,x_{i_p},\ldots,x_n)$ из K всех координат, кроме $x_{i_1},\ldots,x_{i_p}$.

Пусть $P(x_1,\ldots,x_k,x)$ – $(k+1)$-местный предикат, определенный на множестве D. Пусть k-местный предикат $Q(x_1,\ldots,x_k) = (\exists x)P(x_1,\ldots,x_k,x)$. Пусть

$$\lceil P \rceil = \{(x_1,\ldots,x_k,x) \in D^{k+1} : P(x_1,\ldots,x_k,x) \text{ истинно}\} \text{ и}$$
$$\lceil Q \rceil = \{(x_1,\ldots,x_k) \in D^k : Q(x_1,\ldots,x_k) \text{ истинно}\}$$

есть множества истинности предикатов P и Q. Тогда

$$\lceil Q \rceil = \mathrm{Пр}_{1,2,\ldots,k} \lceil P \rceil.$$

Пусть теперь k-местный предикат $R(x_1,\ldots,x_k) = (\forall x)P(x_1,\ldots,x_k,x) = \neg(\exists x)\neg P(x_1,\ldots,x_k,x)$, то

$$\lceil R \rceil = D^k - (\mathrm{Пр}_{1,2,\ldots,k} (D^{k+1} - \lceil R \rceil)).$$

Таков теоретико-множественный смысл кванторов существования и общности.

17.2. Выполнимость, невыполнимость, общезначимость, опровержимость формул логики предикатов

Определим понятие формулы в логике предикатов (ЛП).

Алфавит.

1. $x,y,z,x_1,y_1,z_1,\ldots$ – символы предметных переменных.
2. $a,b,c,a_1,b_1,c_1,\ldots$ – символы предметных констант.
3. $f,g,h,f_1,h_1,g_1,\ldots$ – символы функциональных переменных с указанными местностями.
4. $P,Q,R,P_1,Q_1,R_1,\ldots$ – символы предикатных переменных с

указанными местностями.

5. $\&$, $\vee$, $\rightarrow$, $\neg$, $\exists$, $\forall$ – логические символы (связки).
6. (,) – скобка левая, запятая, скобка правая.

Термы.

1. Всякая предметная переменная и константа есть терм.
2. Если $t_1,\ldots,t_k$ – термы, а f – функциональная переменная местности k, то $f(t_1,\ldots,t_k)$ есть терм.

Если $\{x_1,\ldots,x_s\}$ есть множество всех предметных переменных входящих в терм t, то терм t имеет местность s и тогда пишем $t(x_1,\ldots,x_s)$. Все предметные переменные терма свободны.

Формулы.

1. Если $t_1,\ldots,t_k$ – термы; $\{x_1,\ldots,x_m\}$ – множество всех предметных переменных, входящих в термы $t_1,\ldots,t_k$; P – предикатная переменная местности k, то $P(t_1,\ldots,t_k)$ есть элементарная формула (атом), а $x_1,\ldots,x_m$ – все ее *свободные* (предметные) переменные.
2. Если A и B – формулы, то (A & B), (A $\vee$ B), (A $\rightarrow$ B) – формулы. Их свободные переменные есть свободные переменные формул A и B. Выражение ($\neg$A) тоже формула; ее свободные переменные есть свободные переменные формулы A.
3. Если A есть формула, то выражения ($\exists$x)A и ($\forall$x)A есть формулы. Если формула A свободной переменной x не имеет, то свободные переменные формул ($\exists$x)A и ($\forall$x)A есть свободные переменные формулы A. Если формула A(x) имеет свободную переменную x (и возможно другие свободные переменные), то в формулах ($\exists$x)A(x), ($\forall$x)A(x) переменная x *связана* (квантором), а свободные переменные формул ($\exists$x)A(x), ($\forall$x)A(x) есть свободные переменные формулы A без x.

Формула без свободных предметных переменных называется *замкнутой*.

Здесь действуют те же правила опускания скобок, что и в логике высказываний. Кванторы имеют высший приоритет. В формуле (Qx)A формула A есть область действия квантора (Qx).

Подформулы.

1. Подформулой элементарной формулы является она сама.
2. Подформулами формулы $\neg$A являются она сама и все подформулы формулы A.
3. Подформулами формулы A * B, где знак * $\in \{\&,\vee,\rightarrow\}$, являются она сама и все подформулы формул A и B.
4. Подформулами формулы (Qx)A, где Q $\in \{\exists,\forall\}$, являются она сама и все подформулы формулы A.

Пусть A(B) означает, что формула B есть подформула формулы A.

Как и в исчислении высказываний, будем интерпретировать логические связки &, ∨, →, ¬ как соответствующие логические операции.

Если предикатные, функциональные, свободные предметные переменные и предметные константы формулы A содержатся в списке $P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r$, где $P_1,\ldots,P_n$ – предикатные переменные; $f_1,\ldots,f_m$ – функциональные переменные; $x_1,\ldots,x_k$ – предметные переменные; $a_1,\ldots,a_r$ – предметные константы формулы A, то пишем

$$A(P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r).$$

Введем понятие выполнимости и общезначимости формул. Пусть формула A есть

$$A(P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r).$$ Набор
$$(P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$$

называется *сигнатурой* формулы A. Зафиксируем некоторое множество D, зададим на D предикаты $P_1,\ldots,P_n$; функции $f_1,\ldots,f_m$; предметы $x_1,\ldots,x_k$, $a_1,\ldots,a_r$. Набор

$$I = (D,P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$$

называется *интерпретацией* формулы A. Множество D называется *предметной областью* интерпретации I. Формула A при таком выборе переменных (предикатных, функциональных, предметных) и констант превращается в высказывание

$$A = A(P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1\ldots,a_r),$$

истинное или ложное.

Пример. Пусть формула

$$A(P,Q,R,y,z,a) \text{ есть}$$
$$(\exists x)(\neg P(x,a)\ \&\ \neg P(x,z)\ \&\ Q(x,y)\ \&\ R(x,y));$$

$D = \{0,1,2,\ldots\}$ – множество натуральных чисел;
$P(x,y)$ есть $x = y$;
$Q(x,y)$ есть $x \leqslant y$;
$R(x,y)$ есть $Div(x,y)$ (x делит y);
$y = 8$, $z = 4$, $a = 1$.

Тогда высказывание $A = A(P,Q,R,y,z,a) =$
$(\exists x)(\neg P(x,a)\ \&\ \neg P(x,z)\ \&\ Q(x,y)\ \&\ R(x,y)) =$
$(\exists x)(x \neq 1\ \&\ x \neq 4\ \&\ x \leqslant 8\ \&\ Div(x,8))$
истинно.

Определение (значение $t[x_1,\ldots,x_s]$ терма $t(x_1,\ldots,x_s)$ на интерпретации I).

1. Если $t = x_i$, то $t[x_1,\ldots,x_s] = x_i$.
2. Если $t = a_i$, то $t[x_1,\ldots,x_s] = a_i$.
3. Если $t = f(t_1,\ldots,t_u)$, то $t[x_1,\ldots,x_s] =$ $f(t_1[x_1,\ldots,x_s],\ldots,t_u[x_1,\ldots,x_s])$.

Определение. Интерпретация

$$I = (D,P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$$

выполняет формулу

$$A(P_1,\ldots,P_n,\ f_1,\ldots,f_m,\ x_1,\ldots,x_k,\ a_1,\ldots,a_r)$$

(обозначения: $I \models A$, $A(I)$=И, $A(I)$=1), если высказывание

$$A = A(P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$$

истинно. Интерпретация I *опровергает* формулу A (обозначения: $I \not\models A$, $A(I)$=И, $A(I)$=1), если высказывание A ложно.

Будем вычислять истинностное значение высказывания $I \models A$ индукцией по построению формулы следующим образом.

1. Если A есть атом $P(t_1,\ldots,t_v)$, то
$$I \models A \longleftrightarrow P(t_1[x_1,\ldots,x_k],\ldots,t_v[x_1,\ldots,x_k]) = \text{И}.$$
2. Если A есть $B * C$, где знак $* \in \{\&,\vee,\rightarrow\}$, то
$$I \models A \longleftrightarrow I \models B * I \models C.$$
3. Если A есть $\neg B$, то $I \models B \longleftrightarrow I \not\models B$.
4. Пусть A есть $(\exists x)B(x)$ и пусть A и $B(x)$ есть
$$A(P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r),$$
$$B(P_1,\ldots,P_n,f_1,\ldots,f_m,x,x_1,\ldots,x_k,a_1,\ldots,a_r)$$
соответственно. Пусть
$$I_A = (D,P_1,\ldots,P_n,f_1,\ldots,f_m,x_1,\ldots,x_k,a_1,\ldots,a_r),$$
$$I_B = (D,P_1,\ldots,P_n,f_1,\ldots,f_m,x,x_1,\ldots,x_k,a_1,\ldots,a_r)$$
есть интерпретации формул A и $B(x)$ соответственно. Тогда
$$I_A \models (\exists x)B(x) \longleftrightarrow \exists x \in D\ I_B \models B(x).$$
Аналогично, если A есть $(\forall x)B(x)$, то
$$I_A \models (\forall x)B(x) \longleftrightarrow \forall x \in D\ I_B \models B(x).$$

Определение. Формула A *общезначима на множестве* D, если всякая интерпретация $I = (D,\ldots)$ с предметной областью D выполняет A. Формула A *выполнима на множестве* D, если существует интерпретация $I = (D,\ldots)$ с предметной областью D, которая выполняет A. Формула A *опровержима на множестве* D, ес-

ли существует интерпретация $I = (D,\ldots)$ с предметной областью D, которая опровергает А. Формула А *невыполнима на множестве* D, если всякая интерпретация $I = (D,\ldots)$ с предметной областью D опровергает А.

Пример. Пусть формула $A = (\forall x)(\exists y)P(x,y)\to(\exists y)(\forall x)P(x,y)$.

1. Положим интерпретацию $I = (D,\ P(x,y)) = (\{1,2\},\ x\leqslant y)$. Высказывание $(\forall x)(\exists y)\ x \leqslant y \to (\exists y)(\forall x)\ x \leqslant y$ истинно из-за истинности заключения. Поэтому $I \vDash A$, т.е. интерпретация I превращает формулу А в истинное высказывание.

2. Пусть интерпретация $I = (\{1,2\},\ x = y)$. Высказывание $(\forall x)(\exists y)\ x = y \to (\exists y)(\forall x)\ x = y$ ложно, ибо его посылка истинна, а его заключение ложно. Поэтому $I \nvDash A$, т.е. I опровергает А.

Определение. Формула А *общезначима*, если всякая интерпретация I выполняет А. Формула А *выполнима*, если существует интерпретация I, которая выполняет А. Формула А *невыполнима*, если всякая интерпретация I опровергает А. Формула А *опровержима*, если существует интерпретация I, которая опровергает А.

Заметим, что формулы $A(x_1,\ldots,x_k)$ и $(\forall x_1)\ldots(\forall x_k)A(x_1,\ldots,x_k)$ общезначимы или не общезначимы одновременно. Поэтому при вычислении общезначимости формул можно ограничиться замкнутыми формулами (т.е. формулами без свободных предметных переменных).

Теорема. Чтобы формула логики предикатов была общезначима, необходимо и достаточно, чтобы ее отрицание было невыполнимо.

Доказательство. Пусть формула А общезначима. Допустим, что формула $\neg A$ выполнима. Тогда существует интерпретация I, для которой $I \vDash \neg A$. Поэтому $I \nvDash A$. Противоречие с общезначимостью А.

Пусть теперь формула $\neg A$ невыполнима. Допустим, что формула А общезначимой не является. Тогда существует интерпретация I, которая опровергает А, т.е. $I \nvDash A$. Поэтому $I \vDash \neg A$. Противоречие с невыполнимостью формулы $\neg A$.

Определение. Формула В есть *логическое следствие* формул $A_1,A_2,\ldots,A_k$, если формула $A_1\& A_2\&\ldots\ \&\ A_k\to B$ общезначима.

Заметим, что если формула логики предикатов не содержит кванторов, то ее можно рассматривать как подстановковый случай в формулу алгебры логики, то есть как подстановку

$$\Pi\begin{matrix}B_1 & B_2 & \ldots & B_k\\ p_1 & p_2 & \ldots & p_k\end{matrix}(A(p_1,p_2,\ldots,p_k))$$

формул логики предикатов $B_1,B_2,\ldots,B_k$ вместо всех пропозициональных переменных формулы логики высказываний А.

Если предикатные, функциональные, свободные предметные переменные и предметные константы множества формул S содержатся в списке $Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1,\ldots,a_r$, где $Q_1,\ldots,Q_n$ – предикатные переменные; $g_1,\ldots,g_m$ – функциональные переменные; $x_1,\ldots,x_k$ – предметные переменные; $a_1,\ldots,a_r$ – предметные константы множества формул S, то пишем

$$S(Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1,\ldots,a_r).$$

Введем понятие выполнимости и общезначимости множества формул S. Пусть множество S есть

$S(Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$. Набор
$(Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$

называется *сигнатурой* множества формул S. Зафиксируем некоое множество D, зададим на D предикаты $Q_1,\ldots,Q_n$; функции $g_1,\ldots,g_m$; предметы $x_1,\ldots,x_k$, $a_1,\ldots,a_r$. Набор

$$I = (D,Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$$

называется *интерпретацией* для S. Множество D называется *предметной областью* интерпретации I. Конъюнкция K_S множества Формул S при таком выборе переменных (предикатных, функциональных, предметных) и констант превращается в высказывание

$$K_S = K_S(Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1\ldots,a_r),$$

истинное или ложное. Интерпретация $I = (D,Q_1,\ldots,Q_n,g_1,\ldots,g_m,x_1,\ldots,x_k,a_1,\ldots,a_r)$ *выполняет* множество формул S (обозначения: $I \vDash S$, $S(I)$=И, $S(I)$=1), если I выполняет любую формулу из S (то есть $I \vDash K_S$). Интерпретация I *опровергает* множество формул S (обозначения $I \nvDash S$, $S(I)$=Л, $S(I)$=0), если I опровергает хотя бы одну формулу из S (то есть $I \nvDash K_S$).

Интерпретация I есть *модель* для формулы А (для множества формул S), если $I \vDash A$ (I выполняет всякую формулу из S).

17.3. Равносильность формул

Две формулы А и В логики предикатов называются *равносильными*, или эквивалентными, (обозначение $A \Leftrightarrow B$), если на всякой интерпретации множества $S = \{A,B\}$ эти формулы одновременно истинны или ложны.

Справедливо *правило замены*: если $A \Leftrightarrow B$ и А есть подформу-

ла формулы C, то C(A) ⇔ C(B).

Приведем примеры равносильных формул.

1. $(\exists x)A(x) \Leftrightarrow (\exists y)A(y)$. 2. $(\forall x)A(x) \Leftrightarrow (\forall y)A(y)$, т.е. возможны переименования связанных предметных переменных.

3. $(\exists x)(\exists y)A(x,y) \Leftrightarrow (\exists y)(\exists x)A(x,y)$.

4. $(\forall x)(\forall y)A(x,y) \Leftrightarrow (\forall y)(\forall x)A(x,y)$,

т.е. соседние одноименные кванторы можно менять местами.

5. $\neg(\forall x)A(x) \Leftrightarrow (\exists x)\neg A(x)$.

В самом деле, пусть D – произвольное множество, на котором определены все свободные переменные и константы формулы A, кроме x. Тогда на указанном наборе переменных высказывание $\neg(\forall x)A(x)$ истинно $\longleftrightarrow$ не для всех x из D $A(x)$ истинно $\longleftrightarrow$ существует $x = x_0 \in D$, для которого $A(x_0)$ ложно $\longleftrightarrow$ существует $x = x_0 \in D$, для которого $\neg A(x_0)$ истинно $\longleftrightarrow$ $(\exists x)\neg A(x)$ истинно.

6. $\neg(\exists x)A(x) \Leftrightarrow (\forall x)\neg A(x)$.

В самом деле, $(\forall x)\neg A(x) \Leftrightarrow \neg\neg(\forall x)\neg A(x) \Leftrightarrow \neg(\exists x)\neg\neg A(x) \Leftrightarrow \neg(\exists x)A(x)$.

7. $(\forall x)A(x) \Leftrightarrow \neg(\exists x)\neg A(x)$.

В самом деле, $(\forall x)A(x) \Leftrightarrow \neg\neg(\forall x)A(x) \Leftrightarrow \neg(\exists x)\neg A(x)$.

8. $(\exists x)A(x) \Leftrightarrow \neg(\forall x)\neg A(x)$.

В самом деле, $(\exists x)A(x) \Leftrightarrow \neg\neg(\exists x)A(x) \Leftrightarrow \neg(\forall x)\neg A(x)$.

9. $(\exists x)(A(x) \,\&\, H) \Leftrightarrow (\exists x)A(x) \,\&\, H$, где формула H не содержит переменную x свободно.

В самом деле, $(\exists x)(A(x) \,\&\, H)$ истинно $\longleftrightarrow$ существует $x = x_0 \in D$, для которого $A(x_0)$ и H истинны $\longleftrightarrow$ $(\exists x)A(x) \,\&\, H$ истинно.

10. $(\forall x)(A(x) \vee H) \Leftrightarrow (\forall x)A(x) \vee H$, где формула H не содержит переменную x свободно.

В самом деле, $(\forall x)(A(x) \vee H) \Leftrightarrow \neg\neg(\forall x)(A(x) \vee H) \Leftrightarrow \neg(\exists x)\neg(A(x) \vee H) \Leftrightarrow \neg(\exists x)(\neg A(x) \,\&\, \neg H) \Leftrightarrow \neg(\exists x)\neg A(x) \vee \neg\neg H \Leftrightarrow (\forall x)\neg\neg A(x) \vee H \Leftrightarrow (\forall x)A(x) \vee H$.

11. $(\exists x)(A(x) \vee B(x)) \Leftrightarrow (\exists x)A(x) \vee (\exists x)B(x)$.

В самом деле, $(\exists x)(A(x) \vee B(x))$ истинно $\longleftrightarrow$ существует $x = x_0$, для которого $A(x_0)$ или $B(x_0)$ истинно $\longleftrightarrow$ $(\exists x)A(x_0) \vee B(x_0)$ истинно.

12. $(\forall x)(A(x) \,\&\, B(x)) \Leftrightarrow (\forall x)A(x) \,\&\, (\forall x)B(x)$.

В самом деле, пусть $C(x) \Leftrightarrow A(x) \,\&\, B(x)$. Тогда $(\forall x)C(x) \Leftrightarrow \neg\neg(Ax)C(x) \Leftrightarrow \neg(\exists x)\neg(A(x) \,\&\, B(x)) \Leftrightarrow \neg(\exists x)(\neg A(x) \vee \neg B(x)) \Leftrightarrow ((\exists x)\neg A(x) \vee (\exists x)\neg B(x)) \Leftrightarrow \neg(\exists x)\neg A(x) \,\&\, \neg(\exists x)\neg(B(x) \Leftrightarrow (\forall x)A(x) \,\&\, (\forall x)B(x)$.

Формулы 11 и 12 утверждают, что квантор существования можно распределять по дизъюнкции, а квантор общности по конъюнкции. Что касается распределения квантора существования по конъюнкции, а квантора общности по дизъюнкции, то здесь общезначимы формулы

$(\exists x)(A(x) \,\&\, B(x)) \to (\exists x)A(x) \,\&\, (\exists x)B(x)$;

$(\forall x)A(x) \vee (\forall x)B(x) \to (\forall x)(A(x) \vee B(x))$.

Обратное следование неверно.

13. $(\exists x)A(x) \,\&\, (\exists x)B(x) \Leftrightarrow (\exists x)(\exists y)(A(x) \,\&\, B(y))$.

В самом деле, высказывание $(\exists x)A(x) \,\&\, (\exists x)B(x)$ истинно $\longleftrightarrow$ существуют элементы x_0 и y_0, для которых $A(x_0)$ и $B(y_0)$ истинны $\longleftrightarrow$ $(\exists x)(\forall y)(A(x) \,\&\, B(y))$ истинно.

14. $(\forall x)A(x) \vee (\forall x)B(x) \Leftrightarrow (\forall x)(\forall y)(A(x) \vee B(y))$.

В самом деле, $(\forall x)A(x) \vee (\forall x)B(x) \Leftrightarrow (\forall x)A(x) \vee (\forall y)B(y) \Leftrightarrow (\forall x)(A(x) \vee (\forall y)B(y)) \Leftrightarrow (\forall x)(\forall y)(A(x) \vee B(y))$.

15. $(\forall x)(C \to B(x)) \Leftrightarrow C \to (\forall x)B(x)$, где формула C не содержит переменной x свободно.

В самом деле, $(\forall x)(C \to B(x)) \Leftrightarrow (Ax)(\neg C \vee B(x)) \Leftrightarrow \neg C \vee (\forall x)B(x) \Leftrightarrow C \to (\forall x)B(x)$.

16. $(\forall x)(B(x) \to C) \Leftrightarrow (\exists x)B(x) \to C$, где формула C не содержит переменную x свободно.

В самом деле, $(\forall x)(B(x) \to C) \Leftrightarrow (\forall x)(\neg B(x) \vee C) \Leftrightarrow (\forall x)\neg B(x) \vee C \Leftrightarrow \neg(\exists x)B(x) \vee C \Leftrightarrow (\exists x)B(x) \to C$.

17. $(\exists x)(C \to B(x)) \Leftrightarrow C \to (\exists x)B(x)$,

$(\exists x)(B(x) \to C) \Leftrightarrow (\forall x)B(x) \to C$,

где формула C не содержит переменной x свободно.

Замечание. Если в формулах с 1 по 17 равносильность двух формул $A \Leftrightarrow B$ понимать как $(A \to B) \,\&\, (B \to A)$, то формулы с 1 по 17 станут общезначимыми.

17.3.1. Релятивизованные кванторы

Пусть A(x) и B(x) – формулы логики предикатов, имеющие свободную переменную x и возможно другие свободные переменные. Пусть

$(\exists x)_{A(x)}$ B(x) означает $(\exists x)(A(x) \,\&\, B(x))$,

$(\forall x)_{A(x)}$ B(x) означает $(\forall x)(A(x) \to B(x))$.

Кванторы $(\exists x)_{A(x)}$ и $(\forall x)_{A(x)}$ называются релятивизованными.

Справедливы следующие равносильности:

$\neg(\exists x)_{A(x)}$ B(x) ⇔ $(\forall x)_{A(x)}$ $\neg$B(x);

$\neg(\forall x)_{A(x)}$ B(x) ⇔ $(\forall x)_{A(x)}$ $\neg$B(x).

В самом деле, $\neg(\exists x)_{A(x)}\, B(x) \Leftrightarrow \neg(\exists x)(A(x) \,\&\, B(x)) \Leftrightarrow$
$(\forall x)\neg(A(x) \,\&\, B(x)) \Leftrightarrow (\forall x)(\neg A(x) \vee \neg B(x)) \Leftrightarrow$
$(\forall x)(A(x) \to \neg B(x)) \Leftrightarrow (\forall x)_{A(x)} \neg B(x)$;

$\neg(\forall x)_{A(x)}\, B(x) \Leftrightarrow \neg(\forall x)(A(x) \to B(x)) \Leftrightarrow$
$(\exists x)\neg(\neg A(x) \vee B(x)) \Leftrightarrow (\exists x)(A(x) \,\&\, \neg B(x)) \Leftrightarrow (\exists x)_{A(x)} \neg B(x)$.

17.4. Префиксная нормальная форма

Определение. Формула A логики предикатов задана в *префиксной* нормальной форме, если она имеет вид

$(Q_1x_1)\ldots(Q_kx_k)B(x_1,\ldots,x_k)$, где $Q_i \in \{\exists,\forall\}$, $i=1,2,\ldots,k$,

а B есть бескванторная формула.

Теорема. Для всякой формулы A в ЛП существует равносильная ей формула в префиксной нормальной форме.

Доказательство. Индукция по построению формулы A. Пусть число логических связок в формуле A равно k.

Базис. $k = 0$. Тогда A есть элементарная формула без кванторов; поэтому формула A находится в префиксной нормальной форме.

Предположение индукции. Допустим, что теорема верна для всякой формулы с числом логических связок меньше k.

Шаг индукции. Покажем, что теорема справедлива для всякой формулы с числом логических связок k. Пусть формула A имеет k логических связок. Возможны следующие случаи.

1. A есть $(Qx)B(x)$, где $Q \in \{\exists,\forall\}$. Так как глубина построения формулы $B(x)$ меньше k, то по предположению индукции формула $B(x)$ представима в префиксной нормальной форме $C(x)$. Тогда $(Qx)C(x)$ есть префиксная нормальная форма формулы A.

2. A есть $B * C$, где знак $* \in \{\&, \vee, \to\}$. Так как число логических связок в формулах B и C меньше k, то по предположению индукции существуют префиксные нормальные формы E и F для формул B и C соответственно. Пользуясь доказанными в п.3.3 равносильностями 1–17, выносим в формуле $E * F$ кванторы за скобки и получаем искомую префиксную нормальную форму для формулы A.

3. A есть $\neg B$. Так как число логических связок в формуле B меньше k, то по предположению индукции формула B имеет префиксную нормальную форму $(Q_1x_1)\ldots(Q_kx_k)C$, где C есть бескванторная формула. Тогда формула $\neg B$ имеет префиксную нормальную форму $(Q_1'x_1)\ldots(Q_k'x_k)\neg C$, где Q_i' есть $\exists$, если Q_i есть $\forall$, и Q_i' есть $\forall$, если Q_i есть $\exists$, $i = 1,2,\ldots,k$.

Пример. $(\forall y)(\neg P(x) \to Q(y)) \to (\exists y)(\neg(\forall z)(P(y) \vee Q(z))) \sim$
$\neg(\forall y)(\neg P(x) \to Q(y)) \vee (\exists y)(\exists z)\neg(P(y) \vee Q(z)) \sim$
$(\exists y)\neg(P(x) \vee Q(y)) \vee (\exists y)(\exists z)(\neg P(y) \,\&\, \neg Q(z)) \sim$
$(\exists y)(\neg P(x) \,\&\, \neg Q(y) \vee (\exists z)(\neg P(y) \,\&\, \neg Q(z)) \sim$
$(\exists y)(\exists z)(\neg P(x) \,\&\, \neg Q(y) \vee \neg P(y) \,\&\, \neg Q(z))$.

17.5. Проблема разрешимости в логике предикатов

Основной вопрос математической логики есть проблема разрешимости ее формул, т.е. вопрос об установлении общезначимости (или формальной доказуемости) формул этой логики. *Проблема разрешимости* в ЛП состоит в получении ответа на вопрос: "существует ли алгоритм, который по любой формуле логики предикатов устанавливал бы, является эта формула общезначимой или не является".

Теорема (А.Черча). Проблема распознавания общезначимости формул логики предикатов алгоритмически неразрешима.

Б.А.Трахтенброт установил алгоритмическую неразрешимость ЛП на конечных классах: проблема разрешимости в ЛП остается неразрешимой, если при установлении общезначимости формул ЛП ограничиться конечными множествами.

До конца этого параграфа речь будет идти о формулах в ЛП без функциональных символов.

Хотя и не существует алгоритма, который по любой предъявленной формуле в ЛП устанавливал бы, является эта формула общезначимой или не является, возможны попытки отыскания таких алгоритмов для некоторых классов формул логики предикатов. Были найдены, например, алгоритмы, устанавливающие общезначимость класса замкнутых формул, префиксные нормальные формы которых содержат только кванторы существования. Существует алгоритм распознавания общезначимости класса замкнутых формул, префиксные нормальные формы которых содержат только кванторы общности. Алгоритмически разрешим класс формул, построенных только из одноместных предикатов (монадическая логика). Монадическая логика остается разрешимой, даже если допустить кванторы по предикатным переменным. Оказался алгоритмически разрешимым ряд других классов формул логики предикатов. Исследования по установлению разрешающих алгоритмов в ЛП показали, что относительно мало классов формул имеют разрешающие алгоритмы. Многие классы формул оказались алгоритмически неразрешимыми. Работа по отысканию разрешающих алгоритмов продолжается и поныне.

Покажем разрешимость выше упомянутых примеров классов формул логики предикатов (без функциональных символов).

Формулы $A(x_1,\ldots,x_n)$ и $(\forall x_1)\ldots(\forall x_n)A(x_1,\ldots,x_n)$ общезначимы или не общезначимы одновременно. Кроме того, для всякой формулы логики предикатов существует ей эквивалентная формула в префиксной нормальной форме. Поэтому проблему разрешимости в логике предикатов достаточно решать для замкнутых формул в префиксной нормальной форме.

17.5.1. Проблема разрешимости ∃-формул

Определение. ∃-формула есть замкнутая формула логики предикатов в префиксной нормальной форме, содержащая только кванторы существования.

Класс ∃-формул алгоритмически разрешим, т.е. существует алгоритм, который для всякой ∃-формулы A устанавливает, является формула A общезначимой или не является. Разрешающий алгоритм вытекает из следующей теоремы.

Теорема. ∃-формула общезначима тогда и только тогда, когда она тождественно истинна на одноэлементном множестве.

Доказательство. Если формула A общезначима, она общезначима на всяком множестве, в том числе на одноэлементном множестве.

Пусть теперь ∃-формула

$$B = (\exists x_1)\ldots(\exists x_m)A(P_1,\ldots,P_n,x_1,\ldots,x_m),$$

где бескванторная формула A со свободными предикатными и предметными переменными $P_1,\ldots,P_n,x_1,\ldots,x_m$ тождественно истинна на одноэлементном множестве. Покажем, что формула B общезначима. Допустим противное: формула B общезначимой не является, т.е. существует множество S, предикаты $P_1,\ldots,P_n$, определенные на S, для которых высказывание

$$B = (\exists x_1)\ldots(\exists x_m)\,A(P_1,\ldots,P_n,x_1,\ldots,x_m) \text{ ложно.}$$

Тогда отрицание этого высказывания

$$B = (\forall x_1)\ldots(\forall x_m)\,\neg A(P_1,\ldots,P_n,x_1,\ldots,x_m) \text{ истинно.}$$

Пусть элемент $a \in S$. Тогда высказывание

$$\neg A(P_1,\ldots,P_n,a,\ldots,a) \text{ истинно.}$$

Пусть $M = \{a\}$ – одноэлементное множество. Определим на множестве M предикаты $Q_i(x_1,\ldots x_{n_i})$ равенствами $Q_i(a,\ldots,a) = P_i(a,\ldots,a)$, $i=1,2,\ldots,n$. Тогда

$$\neg A(Q_1,\ldots,Q_n,a,\ldots,a) \text{ истинно,}$$

$$A(Q_1,\ldots,Q_n,a,\ldots,a) \text{ ложно,}$$

$$(\exists x_1)\ldots(\exists x_m)\,A(Q_1,\ldots,Q_n,x_1,\ldots,x_m) \text{ ложно на } M,$$

т.е. формула $(\exists x_1)\ldots(\exists x_m)A(P_1,\ldots,P_n,x_1,\ldots,x_m)$ не является тождественно истинной на одноэлементном множестве. Противоречие. Следовательно, формула B общезначима.

17.5.2. Проблема разрешимости ∀-формул

Определение. ∀-формула есть замкнутая формула логики предикатов в префиксной нормальной форме, содержащая только кванторы общности.

Класс ∀-формул алгоритмически разрешим, т.е. существует алгоритм, который устанавливает для всякой ∀-формулы A, является формула A общезначимой или не является. Разрешающий алгоритм вытекает из следующей теоремы.

Теорема. ∀-формула с m кванторами общности общезначима тогда и только тогда, когда она тождественно истинна на всяком множестве, состоящем из не более чем m элементов.

Доказательство. Если формула A общезначима, то она общезначима на всяком множестве, в том числе на всяком множестве, состоящем из не более чем m элементов.

Пусть теперь ∀-формула

$$A = (\forall x_1)\ldots(\forall x_m)B(P_1,\ldots,P_n,x_1,\ldots,x_m),$$

где бескванторная формула B со свободными предикатными и предметными переменными $P_1,\ldots,P_n,x_1,\ldots,x_m$ тождественно истинна на всяком множестве, состоящем из не более чем m элементов. Покажем, что формула A общезначима. Допустим противное: формула A общезначимой не является, т.е. существует множество S, предикаты $P_1,\ldots,P_n$, определенные на S, для которых высказывание

$$A = (\forall x_1)\ldots(\forall x_m)\,B(P_1,\ldots,P_n,x_1,\ldots,x_m) \text{ ложно.}$$

Отрицание этого высказывания

$$A = (\exists x_1)\ldots(\exists x_m)\,\neg B(P_1,\ldots,P_n,x_1,\ldots,x_m) \text{ истинно.}$$

Тогда существуют элементы $a_1,\ldots a_m$, для которых высказывание

$$\neg B(P_1,\ldots,P_n,a_1,\ldots,a_m) \text{ истинно.}$$

Пусть $M = \{a_1,\ldots,a_m\}$ есть множество, содержащее не более чем m элементов. Определим на множестве M предикаты $Q_i(x_1,\ldots,x_{n_i})$ равенствами $Q_i(t_1,\ldots,t_{n_i})$, $i=1,2,\ldots,n$; $t_1,\ldots,t_{n_i} \in M$. Тогда

$$\neg B(Q_1,\ldots,Q_n,t_1,\ldots,t_{n_i}) \text{ истинно,}$$

$(\exists x_1)\ldots(\exists x_m)\,\neg B(Q_1,\ldots,Q_n,x_1,\ldots,x_m)$ истинно on M,
$\neg(\exists x_1)\ldots(\exists x_m)\,B(Q_1,\ldots,Q_n,x_1,\ldots,x_m)$ ложно on M,
$(\forall x_1)\ldots(\forall x_m)\,B(Q_1,\ldots,Q_n,x_1,\ldots,x_m)$ ложно on M,

т.е. формула A не является тождественно истинной на множестве, состоящем из не более чем m элементов. Противоречие. Поэтому формула A общезначима.

17.5.3. Проблема разрешимости логики одноместных предикатов

Определение. Монадическая формула есть формула логики предикатов, построенная только из одноместных предикатных переменных.

Класс монадических формул алгоритмически разрешим, т.е. существует алгоритм, который устанавливает для всякой монадической формулы A, общезначима A или не общезначима. Разрешающий алгоритм вытекает из следующей теоремы.

Теорема. Пусть A есть замкнутая монадическая формула в префиксной нормальной форме, построенная из n одноместных предикатных переменных. Тогда A общезначима тогда и только тогда, когда A тождественно истинна на всяком множестве, состоящем из не более чем 2^n элементов.

Доказательство. Если формула A общезначима, то она общезначима на всяком множестве, в том числе на всяком множестве, состоящем из не более чем 2^n элементов.

Пусть теперь монадическая формула A тождественно истинна на всяком множестве, состоящем из не более чем 2^n элементов. Пусть $A = (Q_1x_1)\ldots(Q_mx_m)\,B(P_1,\ldots,P_n,x_1,\ldots,x_m)$, где $Q_i \in \{\exists,\forall\}$; B – бескванторная формула; $P_1,\ldots,P_n$ есть полный список одноместных предикатных переменных формулы A. Покажем, что формула A общезначима. Допустим противное: формула A общезначимой не является: существует множество M, одноместные предикаты $P_1,\ldots,P_n$, определенные на M, для которых высказывание

$(Q_1x_1)\ldots(Q_mx_m)\,B(P_1,\ldots,P_n,x_1,\ldots,x_m)$ ложно на M.

Тогда его отрицание

$(Q_1'x_1)\ldots(Q_m'x_m)\,\neg B(P_1,\ldots,P_n,x_1,\ldots,x_m)$ истинно на M.

Здесь Q_i' есть $\exists$, если Q_i есть $\forall$, и Q_i' есть $\forall$, если Q_i есть $\exists$, $i=1,\ldots,m$.

Пусть x_0 – произвольный элемент из S. Пусть набор $(a_1,\ldots,a_n) = (P_1(x_0),\ldots,P_n(x_0))$ есть набор истинностных значений предикатов $P_i(x_0)$). Определим на множестве M отношение

эквивалентности $x \sim y$, положив для элементов x и y из M

$$x \sim y \leftrightarrow (P_1(x),\ldots,P_n(x)) = (P_1(y),\ldots,P_n(y)).$$

Множество M тогда разобьется на не более чем 2^n попарно непересекающихся классов эквивалентности

$$C_{a_1\ldots a_n} = \{x \in M:\ P_i(x)=a_i,\ i=1,\ldots,n\}.$$

Пусть $E = \{e_1,\ldots,e_k\}$ есть множество построенных классов эквивалентности. Определим предикаты $R_i(x)$ на M, положив $R_i(e_j) = P_i(a_j)$, $i=1,2,\ldots,n$; $j=1,2,\ldots,k \leqslant 2^n$; a_j – произвольный элемент из e_j. Так как множество E состоит из не более чем 2^n элементов, то высказывание

$$\underbrace{(Q_1x_1)(Q_2x_2)\ldots(Q_mx_m)\,B(R_1,\ldots,R_n,x_1,\ldots,x_m)}_{T(x_1)} \qquad (*)$$

истинно на множестве E, состоящем из не более чем 2^n элементов. Высказывание

$$\underbrace{(Q_1'x_1)(Q_2'x_2)\ldots(Q_m'x_m)\,\neg B(P_1,\ldots,P_n,x_1,\ldots,x_m)}_{S(x_1)} \qquad (**)$$

истинно на M. То есть

$(Q_1x_1)\,T(x_1)$ истинно на E, и $(Q_1'x_1)\,S(x_1)$ истинно на M.

Утверждение. Если $(Qx)\,T(x)$ истинно на E и $(Q'x)\,S(x)$ истинно на M, то существуют элементы $e_j \in E$ и $a_j \in e_j$, для которых $T(e_j)$ истинно на E и $S(a_j)$ истинно на M.

Доказательство. Возможны следующие два случая.

1. (Qx) есть $(\exists x)$. Тогда $(Q'x)$ есть $(\forall x)$, откуда $(\exists x)T(x)$ истинно на E и $(\forall x)S(x)$ истинно на M. По смыслу квантора существования существует элемент $x = e_j \in E$ для которых $T(e_j)$ истинно на E. По смыслу квантора общности $S(x)$ истинно для всякого $x \in M$. Выберем это $x = a_j \in e_j$. Тогда $S(a_j)$ истинно на M.

2. (Qx) есть $(\forall x)$. Тогда $(Q'x)$ есть $(\exists x)$, откуда $(\forall x)T(x)$ истинно на E и $(\exists x)\,S(x)$ истинно на M. По смыслу квантора существования существует элемент $x = a_j \in M$, для которого $S(a_j)$ истинно на M. По смыслу квантора общности $T(x)$ истинно for any $x \in E$, в том числе для элемента e_j, содержащего a_j. Тогда $T(e_j)$ истинно на M.

В обоих случаях существуют элементы $e_j \in E$ и $a_j \in e_j$ для которых $T(e_j)$ истинно на E и $S(a_j)$ истинно на M. Утверждение доказано.

Продолжим доказательство теоремы. Применим доказанное утверждение к формулам (*) и (**) и найдем элементы $e_1 \in E$ и $a_1 \in e_1$, для которых высказывание

$(Q_2x_2)\ldots(Q_mx_m)\ B(R_1,\ldots,R_n,e_1,x_2,\ldots,x_m)$ истинно на E,
$(Q_2'x_2)\ldots(Q_m'x_m)\ \neg B(P_1,\ldots,P_n,a_1,x_2,\ldots,x_m)$ истинно на M.

Аналогично, применяя доказанное утверждение, последовательно находим элементы $e_1,\ldots,e_m$ из E и элементы $a_1,\ldots,a_m$ из M, $a_j \in e_j$, j=1,2,...,m, для которых высказывание

$B(R_1,\ldots,R_n,e_1,\ldots,e_m)$ истинно на E,
$\neg B(P_1,\ldots,P_n,a_1,\ldots,x_m)$ истинно на M.

Эти формулы имеют одну и ту же алгебраическую структуру: они построены из одной и той же булевой формулы $F(q_1,\ldots,q_r)$, в которой пропозициональные переменные q_i замещаются на $R_j(e_k)$ в первой формуле и на $P_j(a_k)$ во второй. Так как $R_j(e_k) = P_j(a_k)$, то обе формулы эквивалентны. Противоречие, ибо одна из них есть отрицание другой. Поэтому формула А общезначима.

Теорема доказана.

17.6. Отношения

Пусть $A_1,A_2,..,A_n$ – произвольные множества, вообще говоря, разнородные.

Определение. *n–арное отношение* ρ^n на множествах $A_1,A_2,..,A_n$ есть подмножество декартова произведения $A_1\times A_2\times\ldots\times A_n$.

Замечание. n–арное отношение ρ^n на множестве A есть подмножество декартова произведения $A\times A\times\ \ldots\times A$ (n раз). Индекс n арности (местности) отношения иногда опускается.

Иногда отношение определяют на множестве $A_1\times A_2\times\ldots\times A_n$.

Возможна предикатная $\rho(x_1,\ldots,x_n)$ и множественная $(x_1,\ldots,x_n)\in\rho$ формы записи отношений. Отношение ρ называют также предикатом. Для бинарного отношения используюся записи $x\rho y$ и $\rho(x,y)$. Унарное отношение $\rho\subseteq E$ есть подмножество из E.

Набор $\mathbf{a}=(a_1,a_2,\ldots,a_n)\in\rho$ (допустима запись $\rho(a_1,a_2,\ldots,a_n)$) называется элементом отношения.

Замечание. График функции $f: A\to B$ есть некоторое отношение, определенное на $A\times B$.

Определение. Отношение конечно, если оно состоит из конечного числа элементов.

Конечное отношение удобно задавать матрицей, строки которой есть элементы отношения:

$$\rho = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{r1} & a_{r1} & \ldots & a_{r1} \end{bmatrix}$$

Так как отношения есть множества, то над отношениями возможны операции объединения, пересечения, дополнения.

Утверждение. Семейство всех отношений, определенных на некотором множестве $A_1\times A_2\times\ldots\times A_n$, относительно операций $\cup,\cap$, $\neg$ образует булеву алгебру.

Декартово произведение отношений ρ^n и σ^m на множествах $A_1,\ldots,A_n$ и $A_{n+1},\ldots,A_{n+m}$ соответственно, есть отношение

$\theta^{n+m}=\rho^n\times\sigma^m =$
$\{(a_1,\ldots,a_n,a_{n+1},\ldots,a_{n+m}) \in A_1\times\ldots\times A_n\times A_{n+1}\times\ldots\times A_{n+m}$:
$(a_1,\ldots,a_n) \in \rho^n,\ (a_{n+1},\ldots,a_{n+m}) \in \sigma^m\}$.

Пример. $\rho = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$, $\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix}$, $\rho\times\sigma = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 & 1 \\ 0 & 1 & 2 & 2 & 1 & 2 \\ 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 2 & 1 \\ 1 & 2 & 2 & 2 & 1 & 2 \end{bmatrix}$.

Проекция $Pr_{j_1,j_2,\ldots,j_k}P$ отношения $P \subseteq A_1\times A_2\times\ldots\times A_n$ на оси $j_1,j_2,\ldots,j_k$ $(1 \leqslant j_1<j_2<\ldots<j_k \leqslant n)$ есть все те наборы длины k, которые получаются из наборов отношения P, где стерты все компоненты, кроме компонент на местах $j_1,\ldots,j_k$.

Функция $f: A_1\times\ldots\times A_{n-1} \to A_n$ униформизует отношение $P \subseteq A_1\times A_2\times\ldots\times A_n$ (по оси n), если $\forall(a_1,\ldots,a_{n-1}) \in Pr_{1,\ldots,n-1}P$ набор $(a_1,\ldots,a_{n-1},f(a_1,\ldots,a_{n-1})) \in P$. Можно говорить об униформизации отношения P функцией f по любой другой оси. Для двумерного случая функция f: $A \to B$ униформизует (по второй оси) бинарное отношение P из $A\times B$, если $\forall a \in Pr_1P$ пара $(a,f(a)) \in P$. Функция f: $B \to A$ униформизует (по первой оси) бинарное отношение P из $A\times B$, если $\forall b \in Pr_2P$ пара $(f(b),b)\in P$. Например, если A = {0,1,2,3}, B = {0,1,2,3,4}, отношение P = {(0,1),(0,3),(0,4),(1,0),(1,4), (2,1),(2,3),(3,0),(3,4)}, то функция f: $A \to B$, для которой $f(0)$=3, $f(1)$=0, $f(2)$=1, $f(3)$=4, униформизует отношение P по второй оси (рис.17.1).

На множестве отношений можно определить другие операции, обогащающие алгебру отношений и облегчающие их обработку.

Пусть $p=\begin{pmatrix}1 & 2 & \dots & n\\ i_1 & i_2 & \dots & i_n\end{pmatrix}$ есть подстановка элементов множества $M=\{1,2,\dots,n\}$, то есть взаимно однозначная функция $p: M \to M$. Операция перестановки координат в отношении ρ^n порождает отношение $p(\rho^n)$ на $A_{p(1)}\times A_{p(2)}\times\dots\times A_{p(n)}$, для которого $(a_1,\dots,a_n)\in\rho \longleftrightarrow (a_{p(1)},\dots,a_{p(n)})\in p(\rho^n)$.

Например, если

$$\rho=\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ & \dots & & \\ a_{r1} & a_{r2} & \dots & a_{rn}\end{bmatrix},\ \text{то}\ p(\rho)=\begin{bmatrix}a_{1,p(1)} & a_{1,p(2)} & \dots & a_{1,p(n)}\\ a_{2,p(1)} & a_{2,p(2)} & \dots & a_{2,p(n)}\\ & & \dots & \\ a_{r,p(1)} & a_{r,p(2)} & \dots & a_{r,p(n)}\end{bmatrix}.$$

Если $p=\begin{pmatrix}1 & 2 & \dots & n-1 & n\\ 2 & 3 & \dots & n & 1\end{pmatrix}$, то отношение $p(\rho^n)$ называется обратным к ρ и обозначается ρ^{-1}.

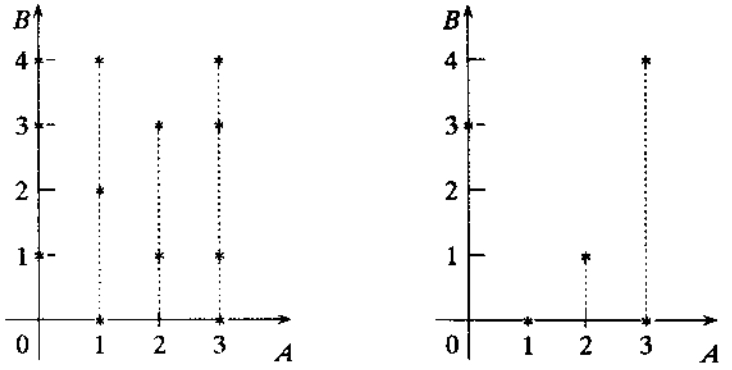
Очевидно, что $(a_1,a_2,\dots,a_n)\in\rho \longleftrightarrow (a_2,\dots,a_n,a_1)\in\rho^{-1}$.

Для обратных отношений справедливы следующие равенства.

1. $(\rho^{-1})^{-1}=\rho$.
2. $\rho\subseteq\sigma \to \rho^{-1}\subseteq\sigma^{-1}$.
3. $(\bigcap\limits_{i\in I}\rho)^{-1}=\bigcap\limits_{i\in I}\rho^{-1}$.
4. $(\overline{\rho})^{-1}=\overline{\rho^{-1}}$.

В частности, для бинарных отношений $x\leqslant y$, $x<y$, $x\geqslant y$, $x>y$ справедливы соотношения:

$(\leqslant)^{-1}=\geqslant$, $(<)^{-1}=>$, $(\geqslant)^{-1}=\leqslant$, $(>)^{-1}=<$.



Отношение P Функция f

Рис.17.1

17.7. Суперпозиция функций

Пусть E – произвольное множество, декартово произведение $E^n = E\times E\times\dots\times E$, объект $f(x_1,\dots,x_n): E^n\to E$ есть функция n переменных (n-местная функция f^n), $P_E=\{f^n: E^n\to E\}$, $n=0,1,\dots$ есть множество всех функций, определенных на множестве E. Нуль-местная функция есть константа из E.

Определение. Пусть F есть некоторое множество функций из P_E. Тогда

1. Всякая функция из F есть суперпозиция над F.
2. Если функция $f(x_1,\dots,x_n)\in F$ и каждое из $A_1,\dots,A_n$ есть либо суперпозиция над F, либо переменная, то $f(A_1,\dots,A_n)$ есть суперпозиция над F.

Определение. Класс M функций из P_E функционально замкнут, если вместе с любыми своими функциями класс M содержит и всякую суперпозицию над ними.

Определение. Замыкание $[M]$ множества M из P_E есть множество всех суперпозиций над M.

Замечание. 1. $M\subseteq[M]$. 2. $[[M]]=[M]$.

3. $M_1\subseteq M_2 \to [M_1]\subseteq[M_2]$.

Пусть $f(x_1,\dots,x_n): E^n\to E$ есть функция n переменных. Примем, что $D(f)$ есть область определения функции f; $R(f)$ или $f(E^n)$ есть область значений функции f; $f(E^n)$ – область тех значений функции f. Если $C^n\subseteq E^n$, то $f(C^n)$ есть область тех значений функции f, когда аргументы функции f пробегают множество C. Запись $f|C$ означает ограничение функции f на множество C. При этом f есть расширение функции $f|C$.

17.8. Операции Мальцева над функциями

Пусть $f(x_1,\dots,x_n)$, $g(x_1,\dots,x_m)$ – функции из P_E. Введем следующие операции (Мальцева).

$\zeta f(x_1,\dots,x_n)=f_\zeta(x_1,\dots,x_n)=f(x_2,\dots,x_n,x_1)$ – циклическая перестановка аргумента;

$\tau f(x_1,\dots,x_n)=f_\tau(x_1,\dots,x_n)=f(x_2,x_1,x_3,\dots,x_n)$ – транспозиция аргументов x_1 и x_2;

$\Delta f(x_1,\dots,x_n)=f_\Delta(x_1,\dots,x_n)=f(x_1,x_1,x_2,\dots,x_{n-1})$ – отождествление переменных;

$\nabla f(x_1,\dots,x_n)=f_\nabla(x_1,\dots,x_n)=f(x_2,\dots,x_{n+1})$ – введение фиктивной переменной;

$f(x_1,\dots,x_n)*g(x_1,\dots,x_m)=f(g(x_1,\dots,x_m),x_{m+1},\dots,x_{m+n-1})$ есть композиция функций f и g.

Утверждение (Мальцев, 1966). 1. Всякая суперозиция функ-

ций над $F \subseteq P_E$ может быть получена из функций системы F с помощью операций $\zeta,\tau,\Delta,\nabla,*$.

2. Функция, полученная из функций системы F с помощью операций $\zeta,\tau,\Delta,\nabla,*$, может быть получена как суперпозиция над функциями из F.

Определение. Алгебра Поста $(\Phi_E,\{\zeta,\tau,\Delta,\nabla,*\})$ есть множество функций $\Phi_E \subseteq P_E$, замкнутое относительно операций $\zeta,\tau,\Delta,\nabla,*$ (то есть замкнутое относительно суперпозиции).

Замечание. Алгебра Поста функций k-значной логики P_k есть алгебра функций, определенных на множестве $E_k=\{0,1,\ldots,k-1\}$.

17.9. Алгебра отношений (реляционная алгебра)

Пусть E – произвольное множество и пусть R_E есть класс всех отношений, определенных на E.

Замечание. Класс всех n-арных отношений, определенных на множестве E, есть множество $\mathcal{P}(E^n)$ всех подмножеств множества E^n. Класс $\mathcal{P}(E^n)$ относительно операций $\cup,\cap,\neg$ образует булеву алгебру (то есть удовлетворяет булевым соотношениям).

17.9.1. Операции Мальцева над отношениями

Пусть $\rho(x_1,\ldots,x_n)$, $\sigma(x_1,\ldots,x_m)$ – отношения, определенные на множестве E. Введем следующие операции (Мальцева).

$\zeta\rho(x_1,\ldots,x_n) = \rho_\zeta(x_1,\ldots,x_n) = \rho(x_2,\ldots,x_n,x_1)$ – циклическая перестановка аргумента;

$\tau\rho(x_1,\ldots,x_n) = \rho_\tau(x_1,\ldots,x_n) = \rho(x_2,x_1,x_3,\ldots,x_n)$ – транспозиция аргументов x_1 и x_2;

$\Delta\rho(x_1,\ldots,x_n) = \rho_\Delta(x_1,\ldots,x_n) = \rho(x_1,x_1,x_2,\ldots,x_{n-1})$ – отождествление переменных;

$\nabla\rho(x_1,\ldots,x_n) = \rho_\nabla(x_1,\ldots,x_n) = \rho(x_2,\ldots,x_{n+1})$ – введение фиктивной переменной;

$\rho(x_1,\ldots,x_n)*\sigma(x_1,\ldots,x_m) = \rho_*(x_1,\ldots,x_{n+m-2}) =$
$\{(a_1,\ldots,a_{n+m-2})\in E^{n+m-2}:\ \exists a\in E\ (a_1,\ldots,a_{n-1},a)\in\rho\ \&$
$(a,a_n,a_{n+1},\ldots,a_{n+m-2})\in\sigma\}$ – свертка отношений ρ и σ.

Замечание. 1. С помощью операций ζ и τ в отношении может быть получена произвольная перестановка переменных.

2. С помощью операций ζ,τ,Δ отождествление переменных может быть осуществлено на любых аргументных местах отношения.

3. С помощью операций ζ,τ,∇ фиктивные переменные могут быть введены на любых аргументных местах отношения.

4. С помощью операций $\zeta,\tau,*$ свертка может осуществляться по любой переменной в обоих отношениях.

5. Кроме перечисленных в теории и практике программирования вводятся и другие операции над отношениями.

Пример. 1. Перестановка переменных. $A=\{0,1,2\}$.

$$\rho(x_1,x_2,x_3,x_4)=\begin{matrix}x_1&x_2&x_3&x_4\end{matrix}\begin{bmatrix}0&1&0&2\\1&0&1&1\\1&0&0&0\end{bmatrix},\quad \rho_1(x_1,x_2,x_3,x_4)=\begin{matrix}x_3&x_4&x_1&x_2\end{matrix}\begin{bmatrix}0&2&0&1\\1&1&1&0\\0&0&1&0\end{bmatrix}.$$

2. Отождествление переменных x_2 и x_4. $A=\{0,1,2\}$.

$$\rho(x_1,\ldots,x_5)=\begin{matrix}x_1&x_2&x_3&x_4&x_5\end{matrix}\begin{bmatrix}0&1&0&1&2\\0&1&1&2&2\\2&2&1&2&0\\1&2&0&1&1\\2&2&1&2&2\end{bmatrix}\rightarrow\begin{matrix}x_1&x_2&x_3&x_4&x_5\end{matrix}\begin{bmatrix}0&1&0&1&2\\2&2&1&2&0\\2&2&1&2&2\end{bmatrix},$$

$$\rho_1(x_1,x_2,x_3,x_5)=\begin{matrix}x_1&x_2&x_3&x_5\end{matrix}\begin{bmatrix}0&1&0&2\\2&2&1&0\\2&2&1&2\end{bmatrix}.$$

3. Введение фиктивной переменной x_0. $A=\{0,1\}$.

$$\rho(x_1,x_2,x_3)=\begin{matrix}x_1&x_2&x_3\end{matrix}\begin{bmatrix}1&0&0\\0&1&1\end{bmatrix}.\quad \nabla\rho=\sigma(x_0,x_1,x_2,x_3)=\begin{bmatrix}0&1&0&0\\1&1&0&0\\0&0&1&1\\1&0&1&1\end{bmatrix}.$$

$\sigma(x_0,x_1,x_2,x_3)=\rho(x_1,x_2,x_3)$,
$\sigma(0,1,0,0)=\rho(1,0,0)$, $\sigma(1,1,0,0)=\rho(1,0,0)$.

4. Свертка. $\rho=\begin{bmatrix}0&2&0\\1&2&1\\0&3&2\\1&1&3\end{bmatrix}$, $\sigma=\begin{bmatrix}0&1\\0&2\\1&3\\2&0\end{bmatrix}$, $\rho*\sigma=\begin{bmatrix}0&2&1\\0&2&2\\1&2&3\\1&1&0\end{bmatrix}$.

Определение. Проекция отношения $\rho\in E^n$ на оси $1\leqslant j_1<\ldots<j_k\leqslant n$ есть k-арное отношение σ, получаемое сохранением в ρ координат $j_1,\ldots,j_k$ и удалением всех остальных.

Пример.

$$\rho=\begin{matrix}x_1&x_2&x_3&x_4&x_5&x_6\end{matrix}\begin{bmatrix}2&0&1&1&0&3\\0&1&2&4&0&1\\0&1&2&4&0&2\end{bmatrix},\quad Pr_{2,4,5}(\rho)=\begin{matrix}x_2&x_4&x_5\end{matrix}\begin{bmatrix}0&1&0\\1&4&0\\1&4&0\end{bmatrix}=\begin{matrix}x_2&x_4&x_5\end{matrix}\begin{bmatrix}0&1&0\\1&4&0\end{bmatrix}.$$

Замечание. Введенные операции могут рассматриваться на

отношениях, определенных на разнородных множествах: $\rho \subseteq A_1 \times \ldots \times A_n$.

17.10. Алгебра отношений k-значной логики

Пусть $E_k = \{0,1,\ldots,k-1\}$.

Определение. Коалгебра Поста $(MO,\{\zeta,\tau,\Delta,\nabla,*\})$ (алгебра отношений R_k, определенных на E_k) есть множество отношений MO, замкнутое относительно операций $\zeta,\tau,\Delta,\nabla,*$.

Определение. Функция $f(x_1,\ldots,x_n)$ из P_k сохраняет h-арное отношение $\rho \subseteq E_k^h$, если $\forall n$ наборов

$$\begin{array}{l}(a_{11}, \ldots, a_{1h}) \in \rho,\\ (a_{21}, \ldots, a_{2h}) \in \rho,\\ \ldots \qquad\qquad \text{набор } (f(\mathbf{a}^1),\ldots,f(\mathbf{a}^h)) \in \rho.\\ (a_{n1}, \ldots, a_{nh}) \in \rho\end{array}$$

$\mathbf{a}^1 \qquad \mathbf{a}^h$

Вместо "f сохраняет ρ" говорят также "f есть полиформизм для ρ" и "ρ есть инвариант для f".

Определение. Пусть $MF \subseteq P_k$ есть некоторое множество функций (k-значной логики) из P_k и $MO \subseteq R_k$ есть некоторое множество отношений из R_k, определенных на E_k.

Множество полиформизмов

$Pol(MO) = \{f \in P_k : \forall \rho \in MO \ f$ есть полиформизм для $\rho\}$.

Множество инвариантов

$Inv(MF) = \{\rho \in R_k : \forall f \in MF \ \rho$ есть инвариант для $f\}$.

Утверждение. 1. Если $(MF,\{\zeta,\tau,\Delta,\nabla,*\})$ есть алгебра функций k-значной логики (алгебра Поста) из P_k, то $(Inv(MF), \{\zeta,\tau,\Delta,\nabla,*\})$ есть алгебра отношений (коалгебра Поста) из R_k.

2. Если $(MO,\{\zeta,\tau,\Delta,\nabla,*\})$ есть алгебра отношений (коалгебра Поста) из R_k, то $(Pol(MO),\{\zeta,\tau,\Delta,\nabla,*\})$ есть алгебра функций k-значной логики (алгебра Поста) из P_k.

Замечание. Между алгебрами функций из P_k и алгебрами отношений из R_k можно установить взаимно однозначное соответствие.

Часть 4. АЛГОРИТМЫ НА ГРАФАХ

18. СПОСОБЫ ЗАДАНИЯ ГРАФОВ

18.1. Графы, мультиграфы, псевдографы

Приведем основные понятия теории графов.

Определение. *(p,q)-граф* есть система объектов $G = (V,E)$, где $V = \{v_1,\ldots,v_p\}$ есть множество *вершин* (узлов); $E = \{e_1 = (v_{i_1},v_{j_1}),\ldots,e_q=(v_{i_q},v_{j_q})\}$ есть множество (неориентированных) *ребер*; $i_k \neq j_k$, $k = 1,2,..,q$.

Пример. Для графа $G = (V,E)$, приведенного на рис.18.1, $V = \{v_1,v_2,v_3,v_4\}$; $E = \{e_1,e_2,e_3,e_4,e_5\}$, где $e_1 = (v_1,v_2)$, $e_2 = (v_1, v_3)$, $e_3 = (v_2,v_3)$, $e_4 = (v_3,v_4)$, $e_5 = (v_2,v_4)$.

Две вершины графа G называются *смежными* (соседними), если они соединены в G ребром. Если граф G имеет ребро $e = (u,v)$, то вершина u и ребро e (равно как и вершина v и ребро e) *инцидентны* (принадлежат друг другу). *Изолированная* вершина не инцидентна ни одному ребру. *Висячая* (концевая) вершина (или лист) инцидентна ровно одному ребру. Два ребра, инцидентные одной вершине, называются *смежными*.

Граф $G_1 = (V_1,E_1)$ *равен* графу $G_2 = (V_2,E_2)$, если $V_1 = V_2$, $E_1 = E_2$. Графы G_1 и G_2 *изоморфны*, если существует взаимно однозначное соответствие $\varphi: V_1 \to V_2$, сохраняющее отношение смежности (соседства) вершин, т.е. если $(u,v) \in E_1 \longleftrightarrow (\varphi(u), \varphi(v)) \in E_2$.

Граф $G_1 = (V_1,E_1)$ есть *подграф* графа $G = (V,E)$, если $V_1 \subseteq V$ и $E_1 \subseteq E$. Если при этом $V_1 = V$, то подграф G_1 называется *остовным подграфом* графа G. Подграф G_1 графа G называется

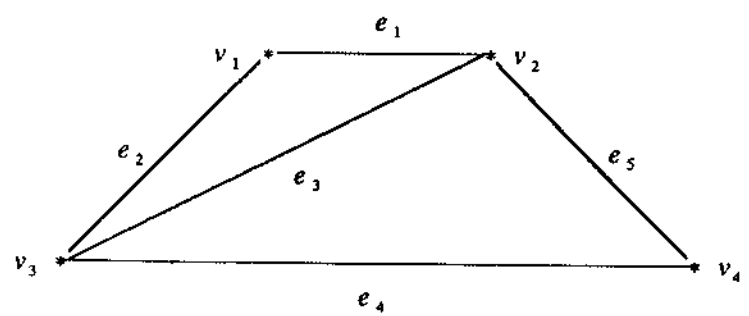


Рис.18.1

порожденным множеством вершин V_1 (натянут на множество вершин V_1), если $V_1 \subseteq V$ & $\forall u,v \in V_1$ $(e = (u,v) \in E \rightarrow e \in E_1)$. Далее, если G_1 есть подграф графа G, то G есть *надграф* графа G_1.

На рис.18.2*а* приведен граф G; на рис.18.2*б* – остовный подграф графа G; на рис.18.2*в* – подграф графа G, натянутый на множество вершин v_1,v_2,v_3,v_4,v_5 графа G.

В *полном* графе K_p все его p вершин смежны. *Петля* есть ребро $e = (u,u)$. *Кратные* (параллельные) ребра соединяют одну и ту же пару вершин.

На рис.18.3 приведены полные графы K_3,K_4,K_5.

Мультиграф (рис.18.4*а*) допускает кратные ребра и не допускает петель. *Псевдограф* (рис.18.4*б*) допускает и кратные ребра, и петли.

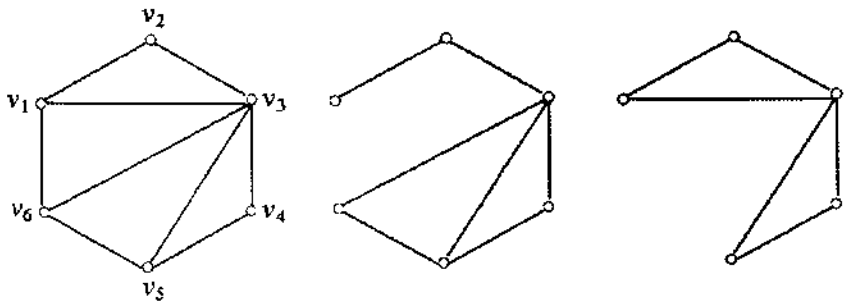


Рис.18.2

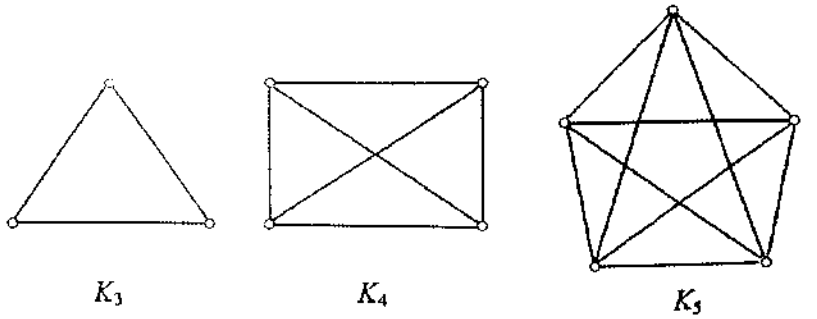


Рис.18.3

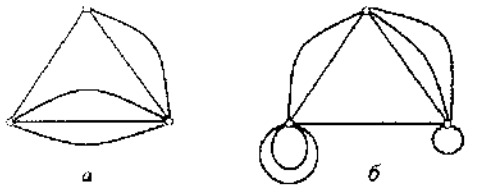


Рис.18.4

Замечание. Граф не допускает ни кратных ребер, ни петель.

Определение. *Ориентированный* граф (орграф) есть граф $G = (V,E)$, в котором множество E дуг есть множество упорядоченных (ориентированных) пар вершин (u,v) (рис.18.5).

Можно говорить об ориентированных мультиграфах и псевдографах.

18.2. Задание графов

Граф $G = (V,E)$ можно задать списком его вершин и ребер. Можно задать и геометрически, нарисовав его на плоскости (или любой другой поверхности) и отождествив его вершины с точками на плоскости, а ребра – с отрезками, соединяющими смежные вершины.

Матрица смежности (соседства) вершин (p,q)-графа $G = (V, E)$ с p вершинами есть квадратная симметричная $p \times p$-матрица

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ & & \dots & \\ a_{p1} & a_{p2} & \dots & a_{pp} \end{bmatrix}, \text{ где } a_{ij} = \begin{cases} 1, & \text{если вершины } v_i,v_j \\ & \text{соседние;} \\ 0 & \text{в противном случае.} \end{cases}$$

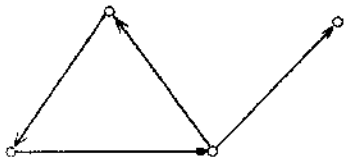


Рис.18.5

Всякому графу соответствует его **бинарная симметричная матрица смежности**. Всякой **бинарной симметричной квадратной матрице** с нулевой диагональю соответствует **некоторый граф**.

Матрица инциденций (p,q)-графа G с p **вершинами и** q ребрами есть $p \times q$-матрица

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ & & \dots & \\ b_{p1} & b_{p2} & \dots & b_{pq} \end{bmatrix}, \text{ где } b_{ij} = \begin{cases} 1, & \text{если вершина } v_i \in e_j; \\ 0 & \text{в противном случае.} \end{cases}$$

Для всякого графа можно построить соответствующую ему бинарную матрицу инциденций. Всякой **бинарной матрице** с двумя единицами в каждом столбце соответствует **некоторый граф**.

На рис.18.6 приведен граф вместе с его матрицами A и B смежности и инциденций соответственно.

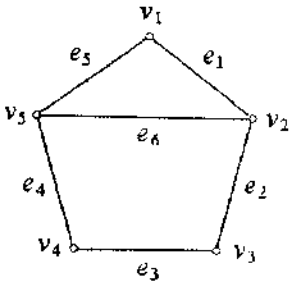
18.3. Операции над графами

Удаление вершины v из графа G **приводит к подграфу** $G-v$ графа G без вершины v и **инцидентных с ней ребер**.

Удаление ребра e из графа $G=(V,E)$ **при сохранении** его вершин приводит к остовному подграфу $G-e = (V,E-\{e\})$ **графа** G.

Добавление ребра $e = (u,v)$ к графу $G = (V,E)$, **содержащему** вершины u,v, приводит к графу $G+e = (V,E \cup \{e\})$.

На рис.18.7 приведены графы G, $G-v$, $G-e$, $G+e'$, где $e' = (v',v'')$.



$$A = \begin{bmatrix} 0&1&0&0&1 \\ 1&0&1&0&1 \\ 0&1&0&1&0 \\ 0&0&1&0&1 \\ 1&1&0&1&0 \end{bmatrix} \begin{matrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{matrix} \qquad B = \begin{matrix} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{matrix} \begin{bmatrix} 1&0&0&0&1&0 \\ 1&1&0&0&0&1 \\ 0&1&1&0&0&0 \\ 0&0&1&1&0&0 \\ 0&0&0&1&1&1 \end{bmatrix} \begin{matrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{matrix}$$

Рис.18.6

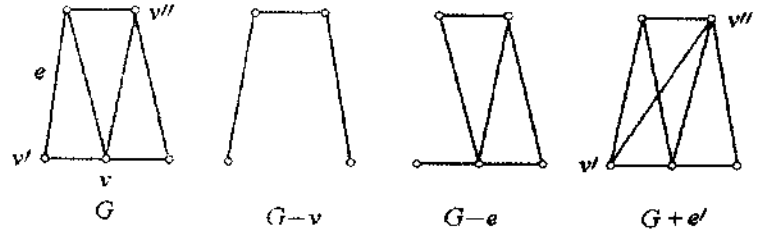


Рис.18.7

18.4. Маршруты, цепи, циклы, связность

Маршрут в графе $G = (V,E)$ есть чередующаяся последовательность вершин и ребер $v_0e_1v_1e_2v_2e_3v_3e_4\dots v_{n-1}e_nv_n$ графа G, для которой каждое ребро инцидентно двум соседним вершинам. При этом v_0 есть начало маршрута, v_n – конец маршрута, n – длина маршрута.

Обозначим через $[u,v]$ маршрут между вершинами u и v. Иногда маршрут отмечается только вершинами, иногда только ребрами. Маршрут нулевой длины состоит из единственной вершины.

Цепь в графе G есть маршрут, в котором все ребра попарно различны (нет повторов ребер, возможны повторы вершин). *Простая цепь* в графе G есть цепь без повторов вершин (а следовательно, и без повторов ребер).

Цикл в графе G есть замкнутая цепь (в которой начало и конец одинаковы). *Простой цикл* не имеет повторов вершин (кроме начальной и конечной), а следовательно, и повторов ребер.

Цепь, простая цепь, цикл, простой цикл в графе G есть некоторые подграфы в графе G. Для графа, изображеного на рис.18.8:

$v_2e_2v_3e_6v_1e_5v_4e_3v_3e_6v_1$, маршрут;
$v_2e_2v_3e_3v_4e_5v_1e_6v_3$, цепь;
$v_2e_2v_3e_6v_1$, простая цепь;
$v_2e_2v_3e_6v_1e_5v_4e_4v_5e_8v_1e_1v_2$, цикл;
$v_2e_2v_3e_6v_1e_1v_2$, простой цикл.

В орграфе маршрут, цепь, простая цепь, цикл, простой цикл становятся ориентированными. Иногда их называют ормаршрут, путь, простой путь, контур, простой контур соответственно.

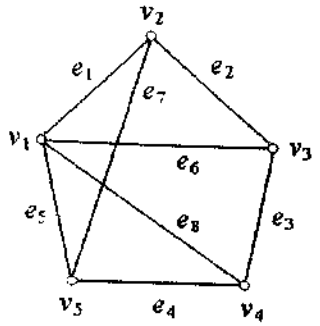


Рис.18.8

По всякому маршруту (циклу) можно построить цепь (простой цикл) с теми же концами, достаточно удалить из них внутренние циклы.

Граф (орграф) G *связен*, если любая пара его вершин u,v соединима цепью (путем) от u к v. В противном случае граф G не связен. *Компонента связности* графа G есть наибольший по числу ребер связный подграф графа G. Граф, состоящий из одних вершин, называется *вполне несвязным*.

Расстояние между двумя вершинами в графе есть длина кратчайшей цепи (*геодезической*) между этими вершинами.

Граф $G=(V,E)$ *нагружен*, если задана функция $w: E \to \mathbb{N}$, сопоставляющая каждому его ребру натуральное число (вес).

18.4.1. Помечивающий алгоритм (Дейкстры) поиска кратчайшего (с наименьшим весом) пути между двумя вершинами s и t в связном нагруженном ориентированном графе

18.4.1.1. Вычисление наименьшего веса пути от s до t

Шаг 1. Присвоим вершине s постоянную (подчеркнутую) пометку $\underline{0}$. Вершину s объявляем активной и помечаем знаком плюс. Всем остальным вершинам присвоим временные (неподчеркнутые) пометки ∞. Переход к шагу 2.

Шаг 2. Если пометка вершины t постоянна (подчеркнута), то алгоритм заканчивает работу. Пометка вершины t равна весу кратчайшего пути от s к t. Постоянные пометки других вершин равны весам кратчайших путей от s до этих вершин. Если

пометка вершины t временная, то переход к шагу 3.

Шаг 3. Изменим временные пометки вершин v, соседних (по дугам) с активной, следующим образом. Присваиваем вершине v временную пометку, равную сумме пометки активной вершины и веса дуги, идущей в вершину v из активной вершины, если эта сумма меньше, чем существующая временная пометка вершины v. В противном случае оставим у вершины v прежнюю пометку. Переход к шагу 4.

Шаг 4. Среди всех вершин с временными пометками найдем вершину с наименьшей пометкой. Если таких вершин несколько, то возьмем любую из них, объявим ее постоянной, а эту вершину – новой активной вершиной, которую помечаем знаком плюс. Прежняя активная вершина свой плюс теряет. Переход к шагу 2.

18.4.1.2. Построение наименьшего пути от s до t

Кратчайший путь от s к t соответствует (в обратном порядке) начинающейся в t и заканчивающейся в s любой последовательности вершин, в которой каждая предыдущая вершина смежна (по дуге) с последующей, причем разность между пометками соседних вершин последовательности равна весу ребра, соединяющему эти вершины.

Пример. Найти кратчайший путь между вершинами s и t в нагруженном связном ориентированном графе $G = (V,E)$, где
$V = \{v_1,v_2,v_3,v_4,v_5,v_6\}$, $s=v_1$, $t=v_6$,
$E = \{\{v_1,v_2,2\},\{v_1,v_3,5\},\{v_2,v_4,3\},(v_3,v_2,1),\{v_3,v_4,1\},$
$\{v_3,v_5,1\},\{v_4,v_6,5\},(v_5,v_4,1),\{v_5,v_6,2\}\}$ (рис.18.9).

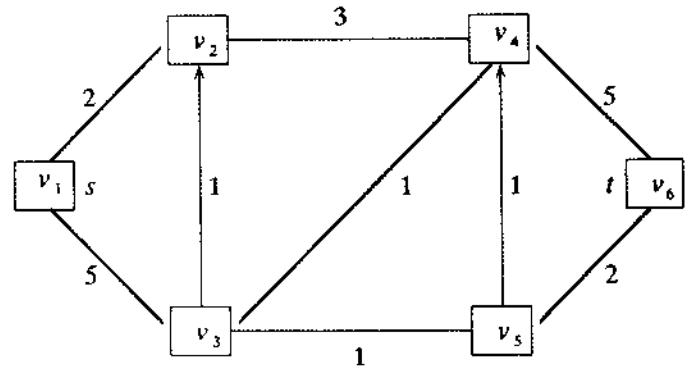


Рис.18.9

Решение. Постоянные пометки подчеркиваем. Активную вершину помечаем знаком плюс.

Вычисление наименьшего веса пути от s до t

Шаг 1. Присваиваем вершине $s=v_1$ постоянную пометку $\underline{0}$. Остальные вершины получают временные пометки ∞. Вершину $s=v_1$ объявляем активной и помечаем знаком плюс (рис.18.10). Переход к шагу 2.

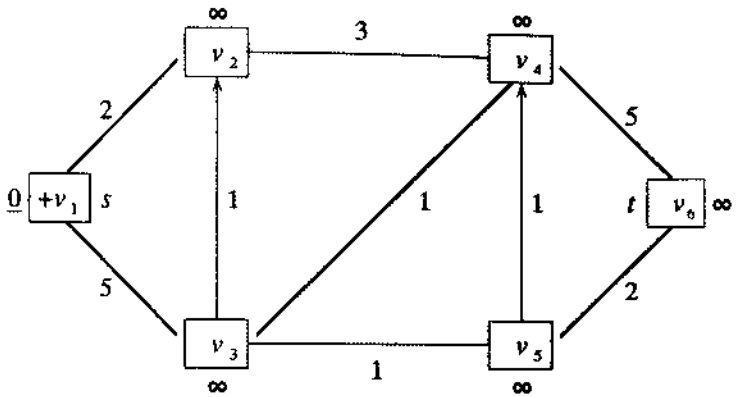


Рис.18.10

Шаг 2. Вершина t постоянной пометки не имеет. Переход к шагу 3.

Шаг 3. Среди всех вершин с временными пометками соседние с активной вершиной $s=v_1$ с пометкой 0 вершины v_2, v_3 имеют временные пометки ∞. Для v_2: $0+2=2<\infty$. Для v_3: $0+5=5<\infty$. Присваиваем для v_2 и v_3 новые временные пометки 2 и 5 соответственно (рис.18.11). Переход к шагу 4.

Шаг 4. Из всех временных пометок пометка 2 для v_2 наименьшая. Объявляем пометку 2 для v_2 постоянной, вершину v_2 объявляем активной и помечаем знаком плюс. Вершина v_1 свой плюс теряет (рис.18.12). Переход к шагу 2.

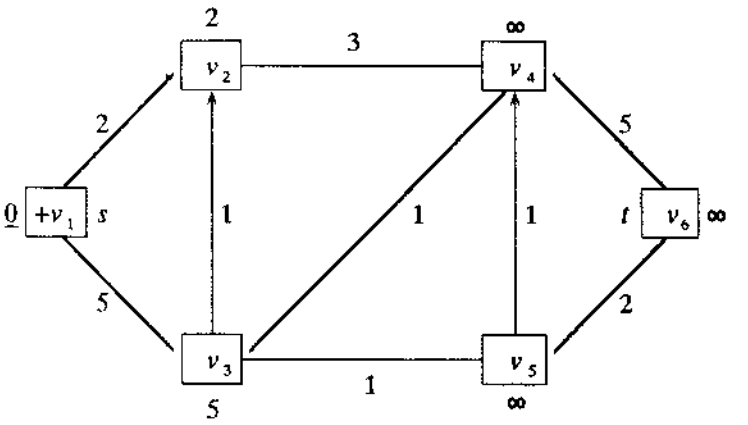


Рис.18.11

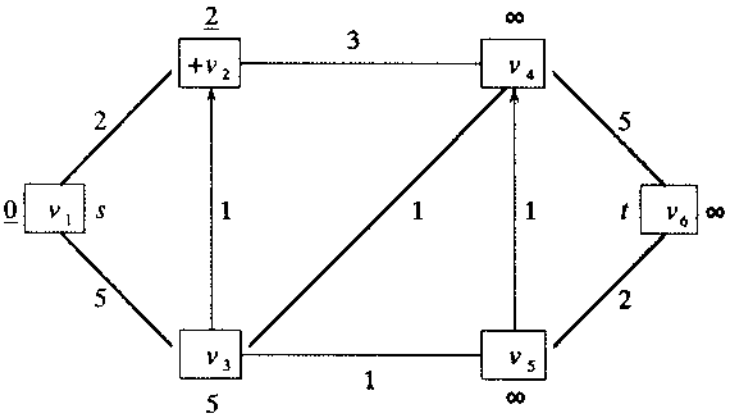


Рис.18.12

Шаг 2. Вершина t постоянной пометки не имеет. Переход к шагу 3.

Шаг 3. Среди всех вершин с временными пометками соседняя с активной вершиной v_2 с пометкой 2 вершина v_4 имеет временную пометку ∞. Для v_4: $2+3=5<\infty$. Присваиваем для v_4 новую временную пометку 5 (рис.18.13). Переход к шагу 4.

Шаг 4. Наименьшие временные пометки 5 у вершин v_3, v_4 одинаковы. Любую из них, например, 5 у v_4, объявляем постоянной, вершину v_4 объявляем активной и помечаем знаком плюс. Вершина v_2 свой плюс теряет (рис.18.14). Переход к шагу 2.

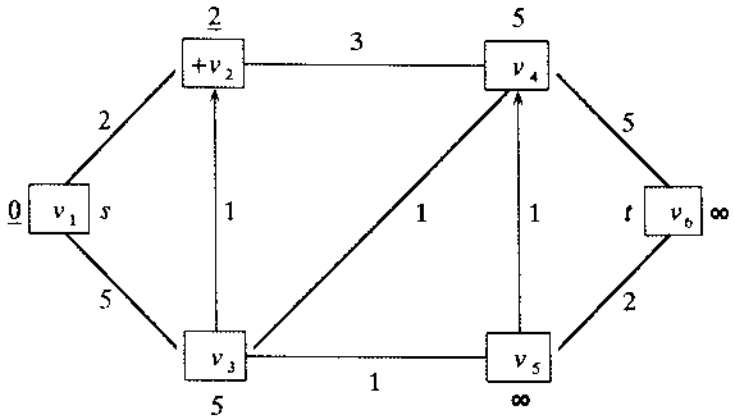


Рис.18.13

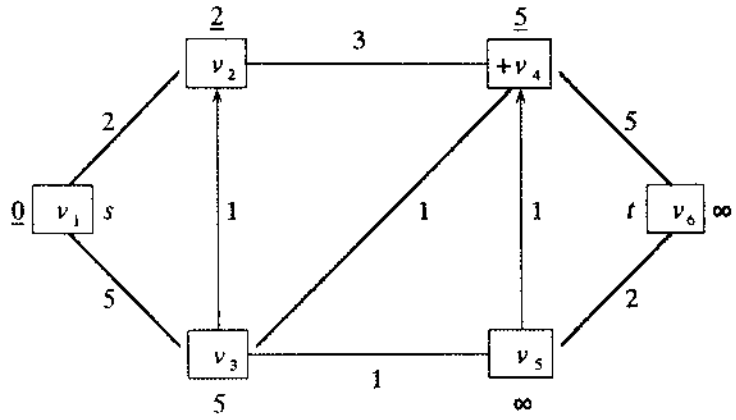


Рис.18.14

Шаг 2. Вершина t постоянной пометки не имеет. Переход к шагу 3.

Шаг 3. Среди всех вершин с временными пометками соседние с активной вершиной v_4 с пометкой 5 вершины v_3, v_6 имеют временные пометки 5 и ∞ соответственно. Для v_3: 5+1=6≥5. Оставляем для v_3 старую пометку 5. Для v_6: 5+5=10<∞. Присваиваем для v_6 новую временную пометку 10 (рис.18.15). Переход к шагу 4.

Шаг 4. Из всех временных (не подчеркнутых) пометок пометка 5 для v_3 наименьшая. Объявляем пометку 5 для v_3 постоянной, вершину v_3 объявлем активной и помечаем знаком плюс.

Вершина v_4 свой плюс теряет (рис.18.16). Переход к шагу 2.

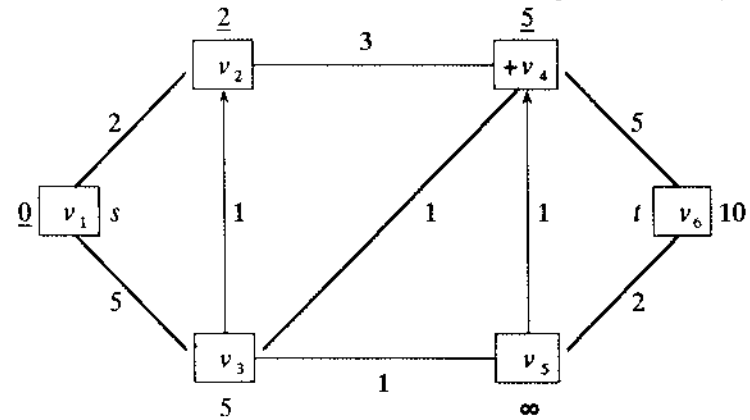


Рис.18.15

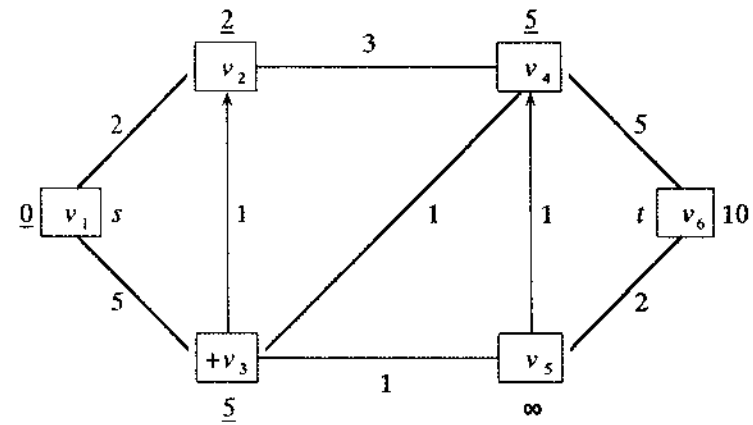


Рис.18.16

Шаг 2. Вершина t постоянной пометки не имеет. Переход к шагу 3.

Шаг 3. Среди всех вершин с временными пометками соседняя с активной вершиной v_3 вершина v_5 имеет временную пометку ∞. Для v_5: 5+1=6<∞. Присваиваем для v_5 новую временную пометку 6 (рис.18.17). Переход к шагу 4.

Шаг 4. Из всех временных пометок пометка 6 для v_5 наименьшая. Объявляем пометку 6 для v_5 постоянной, вершину v_5 объявляем активной и помечаем знаком плюс. Вершина v_3 свой плюс теряет (рис.11.18). Переход к шагу 2.

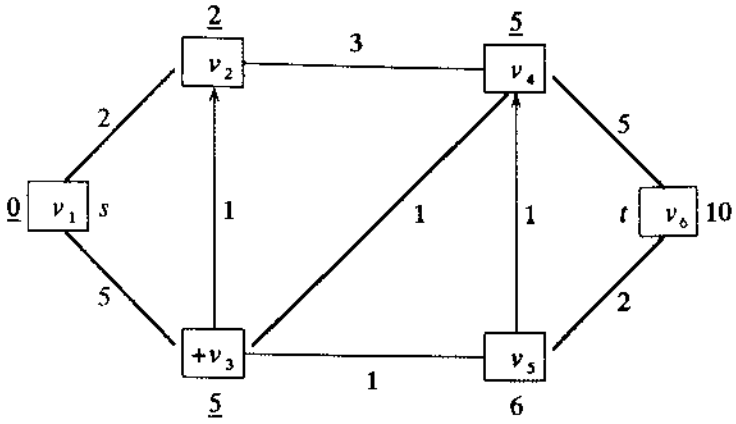


Рис.18.17

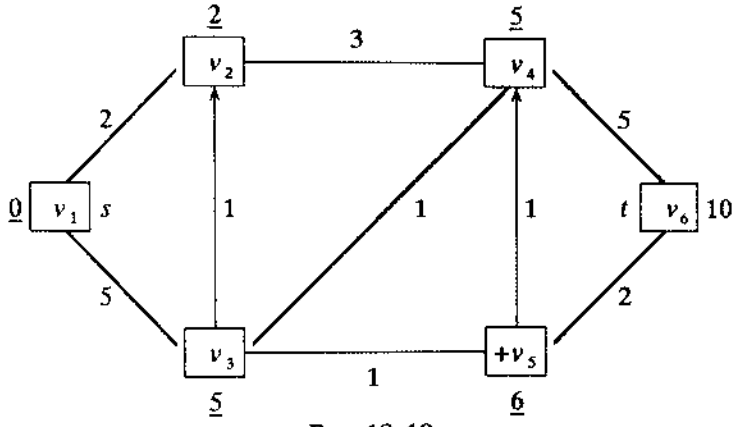


Рис.18.18

Шаг 2. Вершина t постоянной пометки не имеет. Переход к шагу 3.

Шаг 3. Среди всех вершин с временными пометками соседняя с активной вершиной v_5 вершина v_6 имеет временную пометку 10. Для v_6: 6+2=8<10. Присваиваем для v_6 новую временную пометку 8 (рис.18.19). Переход к шагу 4.

Шаг 4. Из всех временных пометок пометка 8 для v_6 наименьшая. Объявляем пометку 8 для v_6 постоянной, вершину v_6 объявляем активной и помечаем знаком плюс. Вершина v_5 свой плюс теряет (рис.18.20). Переход к шагу 2.

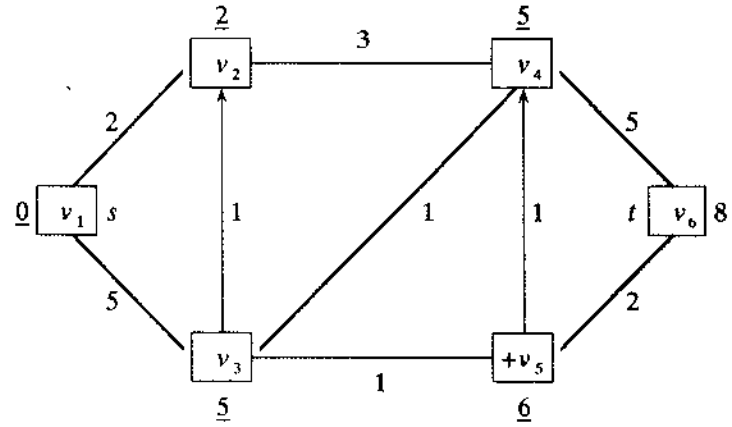


Рис.18.19

Шаг 2. Пометка 8 вершины $t=v_6$ постоянна. Алгоритм заканчивает работу. Пометка 8 вершины t равна весу кратчайшего пути от s до t.

Построение кратчайшего пути от s до t

Пусть $f^{-1}(v)$ есть множество всех вершин v', смежных с v; $d(v)$ есть пометка вершины v; $c(v_i,v_j)$ есть вес ребра (v_i,v_j). $f^{-1}(t) = \{v_4,v_5\}$,
$d(t)-d(v_4)=8-5=3\neq 5=c(v_4,t)$, $d(t)-d(v_5)=8-6=2=c(v_5,t)=2$.
v_5,t есть подпоследовательность кратчайшего пути.

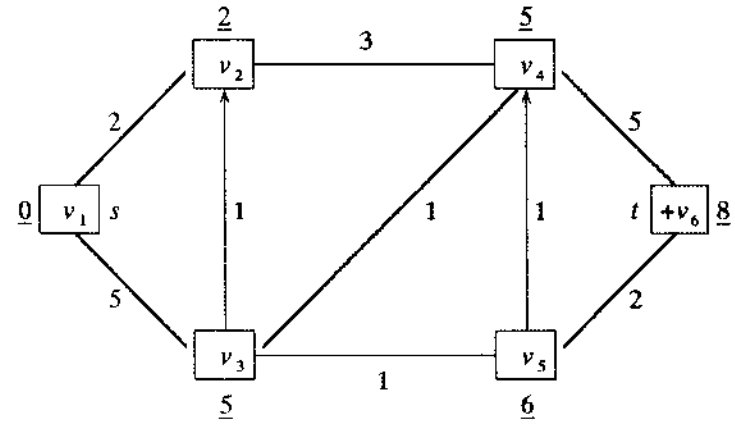


Рис.18.20

$f^{-1}(v_5) = \{v_3\}$,.
$d(v_5)-d(v_3)=6-5=1=c(v_3,v_5)=1$.
v_3,v_5,t есть подпоследовательность кратчайшего пути.

$f^{-1}(v_3) = \{v_1,v_4\}$,
$d(v_3)-d(v_1)=5-0=5=c(v_1,v_3)=5$;. $d(v_3)-d(v_4)=5-5=0\neq 1=c(v_3,v_4)$.
s,v_3,v_5,t есть кратчайший путь от s до t.

Ответ. Путь $s=v_1 \to v_3 \to v_5 \to v_6=t$ от s до t кратчайший. Его (наименьший) вес есть 8.

19. ОБХОДЫ ГРАФОВ

19.1. Эйлеровы графы

Определение. *Степень* вершины графа есть число инцидентных (т.е. принадлежащих) ей ребер. Граф G *четен*, если каждая его вершина имеет четную степень.

Определение. Цикл (контур) в графе (орграфе) G называется *эйлеровым*, если он проходит без повторов ребер (дуг) через каждое ребро (дугу) графа G. Граф (орграф) G называется *эйлеровым*, если он имеет эйлеров цикл (контур).

Эйлеров граф связен, ибо эйлеров цикл связывает все вершины графа. Эйлеров цикл ориентирует ребра графа в направлении обхода.

Теорема. Связный граф G является эйлеровым тогда и только тогда, когда граф G четен.

Доказательство. Пусть G есть связный граф и C – его эйлеров цикл. Каждое прохождение вершины при движении по циклу C привносит двойку в степень этой вершины. Так как каждое ребро графа G появляется в цикле C только один раз, то любая вершина в G имеет четную степень, т.е. граф G четен.

Пусть связный граф G четен. Тогда G имеет простой цикл C. Удалим из G ребра цикла C (сохранив в G их вершины). Снова получим четный граф, так как либо от вершины в G удаляют два ребра цикла (если эта вершина принадлежит циклу C), либо число ребер в ней не меняется (если вершина не принадлежит C). Из получившегося четного графа, каждая компонента связности которого четна, можно снова выделить простой цикл, и так далее, пока не придем к графу, в котором нет ребер. Таким образом граф G допускает представление в виде объединения простых циклов, попарно непересекающихся по ребрам: $G = C_1 \cup C_2$

$\cup \ldots \cup C_k$. Пусть C' есть один из таких простых циклов. Если G состоит только из этого простого цикла C', то G есть эйлеров граф. Если G имеет другие простые циклы, то в силу связности графа G существует простой цикл C'', имеющий общую вершину v с циклом C' (при отсутствии у C' и C'' общих ребер). Тогда $C' \cup C''$ есть цикл с началом в вершине v, составленный сначала из ребер цикла C', а затем из ребер цикла C''. Подобным образом к циклу $C' \cup C''$ последовательно присоединяем другие простые циклы. В результате получим цикл, проходящий через все ребра графа G. Построенный цикл является эйлеровым. Следовательно, граф G тоже эйлеров. Теорема доказана.

Замечание. Теорема и ее доказательство без изменений переносятся на мультиграфы и псевдографы. При этом петли относим к одному из циклов, проходя их все при попадании в вершину, на которую эти петли навешены. Так как петля прибавляет к степени вершины двойку, то любое количество петель на четность вершины не влияет. Аналогичная теорема справедлива и для орграфов.

Теорема. Связный орграф G эйлеров тогда и только тогда, когда для каждой вершины v в G число входящих в v дуг (полустепень захода) равно числу исходящих из v дуг (полустепень исхода).

Пример. Построим эйлеров цикл в четном графе
$G = (\{1,2,3,4,5,6\},\{(1,2),(1,6),(2,3),(2,5),(2,6),(3,4),(3,5),(3,6),(4,5),(5,6)\})$.

Выбираем цикл в G. $C_1 = 2352$; $G_1 = G-C_1 = \{(1,2),(1,6),(2,6),(3,4),(3,6),(4,5),(5,6)\}$.

Выбираем цикл в G_1. $C_2 = 63456$; $G_2 = G_1-C_2 = \{(1,2),(1,6),(2,6)\}$.

Выбираем цикл в G_2. $C_3 = 1261$; $G_3 = G_2-C_3 = \varnothing$.

Из циклов C_1,C_2,C_3 компонуем эйлеров цикл. Выбираем два цикла C_1=2352, C_2=34563 с общей вершиной 3 и вставляем C_2 в C_1 на место вершины 3; получаем цикл C_4=23456352. Циклы C_4, C_3 объединяем по общей вершине 6; получаем C_5=23456126352. Цикл C_5 является эйлеровым циклом.

Эйлеров цикл в четном графе можно построить, начав его любым ребром, а затем последовательно надстраивая его вправо смежными ребрами, одновременно удаляя выбранные ребра из графа и следя за тем, чтобы при очередном удалении ребра из графа он не распался на несвязные компоненты, или не очутились в изолированной вершине, не исчерпав при этом всех ре-

бер графа.

Пример. Построим эйлеров цикл в графе $G=(V,E)$ с множеством вершин $V=\{1,2,3,4,5,6,7,8\}$ и со следующими ребрами:

$e_1=(1,2), e_2=(2,8), e_3=(8,6), e_4=(6,4), e_5=(4,2), e_6=(2,3),$

$e_7=(3,4), e_8=(4,5), e_9=(5,6), e_{10}=(6,7), e_{11}=(7,8), e_{12}=(8,1).$

Мы перечислили ребра в порядке их удаления из графа. Построенная последовательность ребер $e_1,e_2,e_3,\ldots,e_{12}$ составляет эйлеров цикл. Заметим, что после удаления ребра e_4 нельзя убрать ребро e_8, ибо полученный тогда граф распадется на две несвязные компоненты. После удаления ребра e_2 нельзя удалять ребро e_{12}, ибо тогда мы попадем в изолированную вершину 1, не исчерпав всех ребер графа.

Все эйлеровы циклы графа можно построить, начав от любой вершины графа и перебирая дерево всех путей в графе без повторов ребер на каждом пути. Всякая циклическая ветвь в дереве путей, содержащая все ребра графа, составит эйлеров цикл. Этот алгоритм прост, но не эффективен, как всякий переборный алгоритм.

19.2. Полные циклы и последовательности де Брейна

Пусть E_2^n есть множество всех наборов $(a_1,a_2,\ldots,a_n)$ из 0 и 1 длины n.

Определение. Набор из 0 и 1 длины 2^n есть *полный цикл* (де Брейна), если множество его последовательно прочитанных кусков длины n исчерпывает все множество наборов длины n из 0 и 1.

Пример. $n = 2$ (рис 19.1).

Замечание. Цикл де Брейна (как и всякий цикл) располагается по кругу: конец цикла совпадает с его началом.

Теорема. Для всякого n существует полный цикл.

Доказательство. Построим ориентированный псевдограф G следующим образом. Вершины в G пометим наборами из 0 и 1 длины $n-1$. Тогда G имеет 2^{n-1} вершин. Каждый набор $(a_1,a_2,\ldots,a_n)$ из E_2^n длины n определяет дугу, соединяющую вершины $(a_1,a_2,\ldots,a_{n-1})$ и $(a_2,a_3,\ldots,a_n)$. Граф G связен, ибо любая пара вершин $(a_1,\ldots,a_{n-1})$ и $(b_1,\ldots,b_{n-1})$ в G соединима следующим образом:

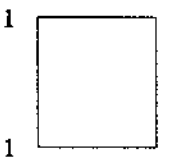
$$(a_1,\ldots,a_{n-1}) \to (a_2,\ldots,a_{n-1},b_1) \to (a_3,\ldots,a_{n-1},b_1,b_2) \to$$

$$\ldots\to (a_{n-1},b_1,\ldots,b_{n-2}) \to (b_1,b_2,\ldots,b_{n-1}).$$

В каждую вершину $(a_2,\ldots,a_n)$ входят две дуги (рис.19.2), помеченные наборами $(0,a_2,\ldots,a_n)$; $(1,a_2,\ldots,a_n)$, и выходят две дуги, помеченные наборами $(a_2,\ldots,a_n,0)$; $(a_2,\ldots,a_n,1)$. Следовательно, степень каждой вершины в G четна; полустепени захода и исхода в каждой вершине G равны, и потому орграф G эйлеров; следовательно, он имеет эйлеров контур, в котором каждая вершина проходится дважды (ибо ее степень равна 4). Тогда первые компоненты наборов, которыми помечены вершины в этом обходе, и дадут полный цикл де Брейна.

Замечание. Теорема справедлива и для наборов $(a_1,a_2,\ldots,a_n)$ из E_k^n длины n с элементами из $E_k = \{0,1,\ldots,k-1\}$.

Пример. Построим цикл де Брейна для n=4. Вершины орграфа есть все наборы из 0 и 1 длины 3:



Полный цикл 0011 дает (по кругу) все наборы длины 2: (0,0); (0,1); (1,1); (1,0).

Рис.19.1

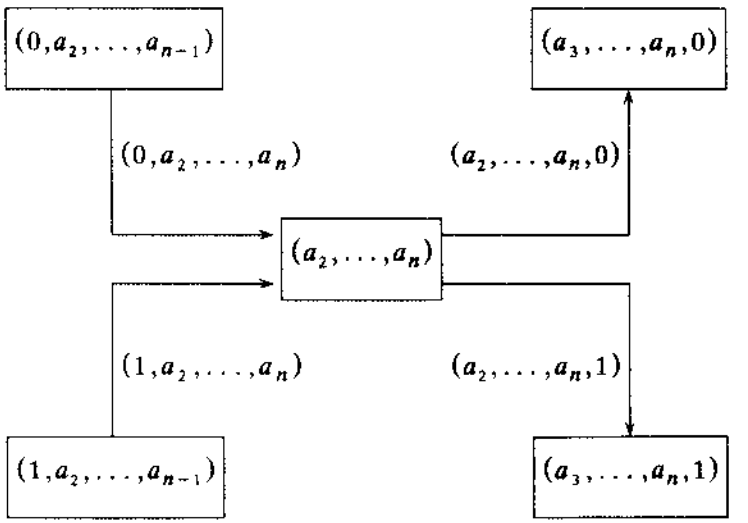


Рис.19.2

000,001,010,011,100,101,110,110;

их можно рассматривать как двоичную запись чисел 0,1,2,3,4, 5,6,7 соответственно. Ребра графа

e_0=(000,000), e_1=(000,001), e_2=(001,010), e_3=(001,011),
e_4=(010,100), e_5=(010,101), e_6=(011,110), e_7=(011,111),
e_8=(100,000), e_9=(100,001), e_{10}=(101,010), e_{11}=(101,011),
e_{12}=(110,100), e_{13}=(110,101), e_{14}=(111,110), e_{15}=(111,111).

Эйлеров цикл

$0e_0 0e_1 1e_2 2e_5 5e_{10} 2e_4 4e_9 1e_3 3e_7 7e_{15} 7e_{14} 6e_{13} 5e_{11} 3e_6 6e_{12} 4e_8 0$;

0 0 0 0 1 0 1 0 0 1 1 1 1 0 1 1 0
0 0 0 1 0 1 0 0 1 1 1 1 0 1 1 0 0
0 0 1 0 1 0 0 1 1 1 1 0 1 1 0 0 0

Составляем цикл де Брейна, взяв от каждого кода вершины первый элемент (старший разряд). Ниже приведены все наборы длины 4, порождаемые циклом де Брейна при его обходе по кругу.

```
        0 0 0 0 1 0 1 0 0 1 1 1 1 0 1 1 0
 0      0 0 0 0
 1        0 0 0 1
 2          0 0 1 0
 5            0 1 0 1
10              1 0 1 0
 4                0 1 0 0
 9                  1 0 0 1
 3                    0 0 1 1
 7                      0 1 1 1
15                        1 1 1 1
14                          1 1 1 0
13                            1 1 0 1
11                              1 0 1 1
 6                                0 1 1 0
12                                  1 1 0 0
 8                                    1 0 0 0
```

19.3. Гамильтоновы графы

Цикл C в графе G называется *гамильтоновым циклом*, если C проходит без повторов вершин через все вершины графа G. Граф G называется *гамильтоновым графом*, если G имеет гамильтонов цикл.

Гамильтонов граф допускает обход всех своих вершин, при этом каждая вершина проходится единожды. Определения гамильтоновых и эйлеровых графов весьма схожи, но методы их исследования существенно различны. Пока не найдены простые критерии гамильтоновости графа, отличные от переборного критерия; доказаны лишь некоторые достаточные условия гамильтоновости, два из которых приведем здесь без доказательства.

Теорема (Поша). Если связный граф G:

1) имеет $p \geqslant 3$ вершин;

2) для любого n из того, что $1 \leqslant n < (p-1)/2$ следует, что число вершин со степенями, не превосходящими n, меньше n;

3) из нечетности p следует, что число вершин степени $(p-1)/2$ меньше $(p-1)/2$, то G есть гамильтонов граф.

Ограничивая условия теоремы Поша, получаем следующие более простые, но менее сильные достаточные условия гамильтоновости графа.

Следствие 1 (Оре). Если число вершин графа G $p \geqslant 3$ и сумма степеней вершин $deg(u) + deg(v) \geqslant p$ для всякой пары несмежных вершин u и v в G, то граф G гамильтонов.

Следствие 2 (Дирак). Если число вершин графа G $p \geqslant 3$ и $deg(v) \geqslant p/2$ для всякой вершины v графа G, то граф G гамильтонов.

На рис.19.3 приведены примеры эйлеровых и гамильтоновых графов.

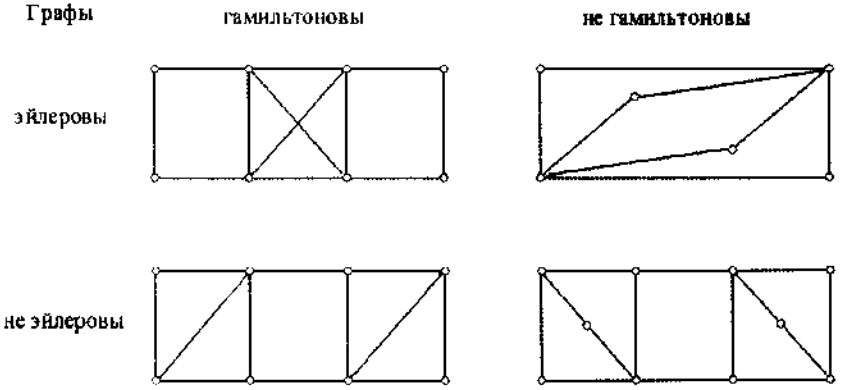


Рис.19.3

19.4. Коды Грея

Определение. k–арный код Грея порядка n есть последовательность всех k^n наборов длины n из 0 и 1, в которых каждый следующий набор отличается от предыдущего только в одном разряде.

k–арный код Грея порядка n строится индукцией по порядку n следующим образом.

Базис. n=1. Последовательность наборов $0,1,\ldots,k-1$ длины 1 есть k–арный код Грея порядка n=1.

Предположение индукции. Предположим, что k–арный код Грея $C_0,C_1,\ldots,C_{k^{n-1}}$ уже построен.

Шаг индукции. Последовательность наборов

$$
\begin{array}{l}
C_0 0, C_1 0, \ldots, C_{k^{n-2}} 0, C_{k^{n-1}} 0,\\
C_{k^{n-1}} 1, C_{k^{n-2}} 1, \ldots, C_1 1, C_0 1,\\
C_0 2, C_1 2, \ldots, C_{k^{n-2}} 2, C_{k^{n-1}} 2,\\
C_{k^{n-1}} 3, C_{k^{n-2}} 3, \ldots, C_1 3, C_0 3,\\
\qquad \ldots\\
\begin{cases}
C_0(k-1), C_1(k-1), \ldots, C_{k^{n-2}}(k-1), C_{k^{n-1}}(k-1), & \text{если } n \text{ не четно},\\
C_{k^{n-1}}(k-1), C_{k^{n-2}}(k-1), \ldots, C_1(k-1), C_0(k-1), & \text{если } n \text{ четно}.
\end{cases}
\end{array}
$$

Пример. 1. Коды Грея для k=4, n=1,2,3.

n=1. 0,1,2,3.

n=2. 00,10,20,30, 31,21,11,01, 03,13,23,33.

n=3. 000,100,200,300,310,210,110,010,030,130,230,330,
331,231,131,031,011,111,211,311,301,201,101,001,
002,102,202,302,312,212,112,012,032,132,232,332.

2. Бинарные коды Грея для k=2, n=1,2,3,4.

n=1. 0,1.

n=2. 00,10, 11,01.

n=3. 000,100,110,010, 011,111,101,001.

n=4. 0000,1000,1100,0100,0110,1110,1010,0010,
0011,1011,1111,0111,0101,1101,1001,0001.

Замечание. Пусть в графе $G=(V,E)$ вершины V есть наборы из E_2^n, а ребро связывает две вершины, если они отличаются только в одном разряде, то код Грея есть гамильтонов цикл для G.

20. ДЕРЕВЬЯ

20.1. Деревья и лес

Определение. *Дерево* есть связный граф, не имеющий циклов. Совокупность деревьев есть *лес*.

Определение. Ребро в графе G называется *циклическим*, если оно принадлежит хотя бы одному циклу графа G, и *ациклическим* в противном случае. Граф G *ацикличен*, если каждое его ребро ациклично.

Замечание. При удалении из связного графа циклического ребра граф остается связным; при удалении ациклического ребра – несвязным.

Каждая компонента связности леса есть дерево.

Пусть G есть связный граф. Тогда следующие утверждения эквивалентны.

1. Граф G есть дерево.
2. Граф G не имеет циклов.
3. Граф G не имеет циклических ребер.
4. Все ребра в графе G ацикличны.
5. Граф G ацикличен.

20.2. Характеристические свойства деревьев

Теорема. Для всякого (p,q)-графа G следующие утверждения эквивалентны.

1. Граф G есть дерево.
2. Любая пара вершин в G соединима единственной простой цепью.
3. Граф G связен и $q - p + 1 = 0$.
4. Граф G ацикличен и $q - p + 1 = 0$.
5. а) граф G ацикличен;
   б) если любую пару несмежных вершин в G соединить ребром e, то получившийся граф $G+e$ будет иметь в точности один простой цикл.

Доказательство. 1 → 2. Пусть граф G есть дерево. Так как граф G связен, то всякие две вершины u,v в G соединимы простой цепью Z. Эта цепь единственна, ибо при наличии другой цепи Z' между u и v (рис.20.1) получим, что дерево G имеет цикл, чего нет.

2 → 3. Так как любая пара вершин в G соединима (единственной простой) цепью, то граф G связен. Соотношение $q - p + 1 = 0$ докажем индукцией по числу p вершин в G.

Базис. $p = 1$. Так как граф G состоит из единственной вершины, то $q - p + 1 = 0 - 1 + 1 = 0$.

Предположение индукции. Пусть соотношение $q-p+1=0$ справедливо для всех графов с числом вершин меньше p.

Шаг индукции. Покажем, что равенство $q-p+1=0$ справедливо для всех графов с числом вершин p. Пусть связный граф G имеет p вершин. Удалим из G ребро $e = (u,v)$, сохранив в G его концы. Получившийся граф $G' = G-e$ не связен (рис.20.2), ибо если G' связен, то между u и v в G', а потому и в G, существует простая цепь. Другая простая цепь в G между u и v есть ребро e. Противоречие с единственностью простой цепи в G между u и v. Итак, граф G' не связен и имеет в точности две компоненты связности с числом вершин в каждой компоненте p_1, p_2 и числом ребер q_1, q_2; при этом $p = p_1+p_2$, $q = q_1+q_2+1$. По предположению индукции $q_i-p_i+1 = 0$, $i=1,2$. Тогда для графа G сумма $q-p+1 = (q_1+q_2+1) - (p_1+p_2)+1 = (q_1-p_1+1) + (q_2-p_2+1) = 0+0 = 0$.

Шаг индукции установлен. Равенство $q-p+1 = 0$ доказано.

Итак, граф G связен и $q-p+1 = 0$.

3 → 4. Покажем, что граф G ацикличен. Допустим противное: граф G цикличен. Тогда G имеет простой цикл C длины $n \leqslant q$ (длина цикла есть число его ребер). Цикл C имеет n вершин и n ребер. Пусть $U = \{u_1,u_2,\ldots,u_{p-n}\}$ (рис.20.3) есть оставшиеся вершины в G (т.е. вершины вне цикла C). Пусть v есть вершина цикла C. Так как граф G связен, то вершина v цикла C соединима с любой вершиной u_i из U. Пусть $[u_1,v],[u_2,v],\ldots,[u_{p-n},v]$ есть геодезические (т.е. цепи наименьшей длины). Ребра из $E' = \{e_1,e_2,\ldots,e_{p-n}\}$, лежащие на этих геодезических и инцидентные вершинам $u_1,u_2,\ldots,u_{p-n}$ соответственно, попарно различны (как инцидентные различным вершинам).

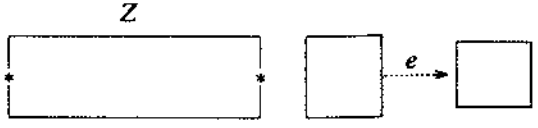


Рис.20.1

Рис.20.2

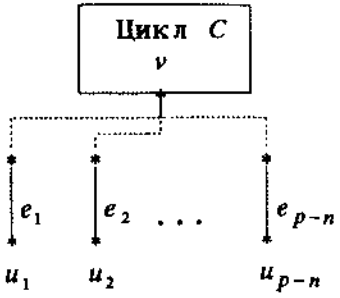


Рис.20.3

Пусть граф $G' = C \cup E'$. Так как $G' \subseteq G$, то число ребер в G' $n + (p - n) \leqslant q$; поэтому $p \leqslant q$, откуда $q - p \geqslant 0$ и потому $q - p + 1 \geqslant 1$. Противоречие с $q - p + 1 = 0$. Следовательно, граф G ацикличен.

Итак, граф G ацикличен и $q - p + 1 = 0$.

4 → 5. Пусть граф G ацикличен и $q-p+1=0$. Покажем сначала, что граф G есть дерево. Так как граф G ацикличен, то каждая его компонента связности есть дерево. Пусть G имеет k компонент связности (k деревьев T_i), причем дерево T_i имеет p_i вершин и q_i ребер. Тогда граф G имеет $p = \sum_{i=1}^{k} p_i$ вершин и $q = \sum_{i=1}^{k} q_i$ ребер. Так как для каждого дерева T_i по 1 → 3 $q_i-p_i+1 = 0$, то для графа G $\sum_{i=1}^{k}(q_i-p_i+1) = \sum_{i=1}^{k} q_i - \sum_{i=1}^{k} p_i + \sum_{i=1}^{k} 1 = q-p+k = 0$, что вместе с $q-p+1 = 0$ дает $k = 1$, т.е. граф G имеет только одну компоненту связности, а потому граф G есть дерево.

Соединим в дереве G любую пару несмежных вершин u,v ребром $e=(u,v)$. Ребро e вместе с единственной (по 1 → 2) простой цепью $[u,v]$ в G дает единственный простой цикл в $G+e$.

Итак: а) граф G ацикличен; б) если любую пару несмежных вершин в G соединить ребром e, то получим граф $G+e$, имеющий в точности один простой цикл. Условие 5 показано.

5 → 1. Пусть выполнены условия 5. Граф G связен, ибо если G не связен, то добавляя к G ребро e между двумя компонентами связности, получим ациклический граф (ибо ребро e ациклично), не имеющий циклов в противоречие с условием 5. Следовательно, граф G связен и ацикличен, т.е. G есть дерево.

Условие 1 установлено. Теорема доказана.

Следствие 1. Связный (p,q)-граф G есть дерево тогда и только тогда, когда $q - p + 1 = 0$.

Доказательство следует из того, что по доказанной теореме условия 1 и 3 эквивалентны.

Следствие 2. Пусть (p,q)-граф G связен. Тогда G имеет единственный простой цикл тогда и только тогда, когда $q - p$

$+ 1 = 1$.

Доказательство. Пусть связный (p,q)-граф G имеет единственный простой цикл C. Удалим из G его циклическое ребро e (лежащее в C). Граф $G' = G-e$ связен и не имеет простых циклов, ибо если G' имеет простой цикл C', то G имеет два простых цикла C и C', различающихся ребром e: цикл C ребро e имеет, а цикл C' нет. Следовательно, граф G' есть дерево с тем же числом вершин p. Для дерева G' $(q-1)-p+1 = 0$, откуда $q-p+1 = 1$.

Пусть теперь для связного (p,q)-графа G имеет место соотношение $q-p+1 = 1$. Тогда G не есть дерево (ибо для дерева $q-p+1 = 0$). Следовательно, граф G имеет циклическое ребро e, ибо без него граф G был бы деревом. Удалим из G это циклическое ребро e, входящее в некоторый простой цикл C. Полученный связный граф $G' = G-e$ имеет p вершин и $q-1$ ребер. Тогда $(q-1)-p+1 = (q-p+1)-1 = 0$ (ибо по условию $q-p+1 = 1$); следовательно G' есть дерево. Возвращение в дерево G' удаленного ребра e приводит к связному графу G, который в силу эквивалентности условий 1 и 5 теоремы имеет единственный простой цикл.

Замечание. Выбрав в дереве T произвольную вершину v (корень дерева) и достраивая от нее граф T вниз, получим изображение для T, в котором вершины группируются по ярусам (рис.20.4).

20.3. Каркасы и хорды в связном графе

Определение. *Каркас* (стягивающее дерево, остов) в связном графе G есть наименьшее по числу ребер дерево в G, сохраняющее связность между всеми вершинами G.

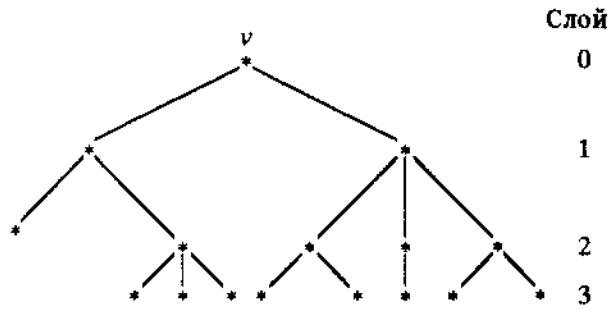


Рис.20.4

Каркас графа G можно получить, последовательно удаляя (стирая) из G его циклические ребра. Каркас графа G содержит все ациклические ребра в G, ибо их нельзя удалить из G, не нарушив связности получившегося графа. Каркас в графе G находится неоднозначно.

Каркас графа G можно получить также, последовательно надстраивая ребрами из G произвольно взятое в G ребро до дерева, являющегося каркасом. Если граф G является нагруженным (каждому ребру графа G приписано некоторое неотрицательное число – вес ребра, его стоимость), то имеет смысл разыскивать наименьший каркас (каркас с наименьшей суммой весов ребер). Такой наименьший каркас можно получить, последовательно надстраивая ребрами из G произвольно взятое в G ребро с наименьшим весом до дерева, являющегося каркасом. При этом надстройку каждый раз следует выполнять ребром с наименьшим возможным весом, избегая появления циклов.

Пример. Построим наименьший каркас для нагруженного графа

$G=(V,E)=(\{a,b,c,d,e,f,g\},\{(a,b,1),(a,g,2),(b,c,7),$
$(b,d,6),(b,f,8),(b,g,3),(c,d,2),(c,g,9),(d,e,1),(d,f,9),$
$(d,g,1),\ (e,f,4),(e,g,5),(f,g,9)\})$.

Исходим из ребра (a,b,1). Последовательное его расширение ребрами с наименьшим весом (избегаем при этом появления циклов) приводит нас к каркасу

$$\{(a,b,1),(a,g,2),(d,g,1),(d,e,1),(c,d,2),(e,f,4)\}$$

с наименьшим весом 11.

Определение. Если $K = (V,E_0)$ есть каркас графа $G = (V,E)$, то *хорда* есть произвольное ребро из $E - E_0$.

Множество хорд есть семейство удаленных в процессе построения каркаса K циклических ребер. Каждая хорда (последовательно удаленное циклическое ребро) соответствует некоторому простому циклу в G.

Пример. На рис.20.5 приведен граф G и его каркас, полученный последовательным удалением из G циклических ребер. Тогда $M = \{e_1,e_2,\ldots,e_8\}-\{e_1,e_2,e_5,e_7,e_8\} = \{e_3,e_4,e_6\}$ есть множество хорд графа G.

Утверждение. Число хорд в связном (p,q)-графе $G = (V,E)$ равно $q - p + 1$.

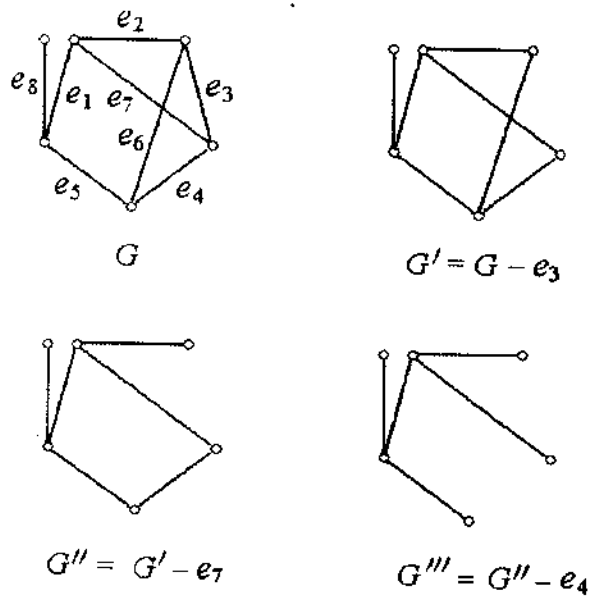


Рис.20.5

Доказательство. Пусть (p,q')-граф $K = (V,E_0)$ есть каркас (связного) (p,q)-графа $G = (V,E)$. Так как каркас K есть дерево, то $q'-p+1 = 0$. Поэтому число хорд в графе G равно $\nu = |E| - |E_0| = q - q' = q - q' + (q' - p + 1) = q - p + 1$, где $|E|$ означает мощность множества E.

Замечание. Каркас K имеет число ребер

$$q' = q - \nu = q - (q - p + 1) = p - 1.$$

Чтобы сделать две вершины в связном графе G несвязными, достаточно удалить из G все множество хорд и одно ребро, (например, принадлежащее одной из двух этих вершин) на пути между ними.

Множество хорд графа G, дополненных до простых циклов в G, составляют фундаментальную систему циклов в G. Фундаментальную систему циклов можно построить также, последовательно добавляя каждую хорду к каркасу графа G и разыскивая тот единственный цикл, который через эту хорду проходит.

21. ЦИКЛЫ В ГРАФАХ

21.1. Линейное пространство двоичных наборов

Напомним: *поле* $(F,\{+,\cdot\})$ есть множество F элементов произвольной природы с определенными на нем двумя операциями: сложением $a + b$: $F \times F \to F$ и умножением $a \cdot b$: $F \times F \to F$, удовлетворяющими следующим аксиомам: $\forall a,b,c \in F$

1. $(a + b) + c = a + (b + c)$.
2. $a + b = b + a$.
3. $\exists 0 \in F\ (a + 0 = a)$.
4. $\forall a \in F\ \exists (-a) \in F\ a + (-a) = 0$.

} F есть абелева группа по сложению

5. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
6. $a \cdot b = b \cdot a$.
7. $\exists e \in F\ \forall a \in F\ a \cdot e = a$.
8. $\forall a \in F - \{0\}\ \exists a^{-1} \in F\ a \cdot a^{-1} = e$.

} F без нуля есть абелева группа по умножению

9. $a \cdot (b + c) = a \cdot b + a \cdot c$.

Иногда поле $(F,\{+,\cdot\})$ обозначается через F.

Пример. $E_2 = \{0,1\}$. Сложение и умножение (по модулю два) зададим согласно следующим таблицам.

x	y	$x+y$
0	0	0
0	1	1
1	0	1
1	1	0

x	y	$x \cdot y$
0	0	0
0	1	0
1	0	0
1	1	1

Элемент -1, равный $0-1$, равен такому c, что $0 = 1 + c$, откуда $c = 1$; поэтому $-1 = 1$ (по модулю 2). Объект $F = (E_2, \{+,\cdot\})$ есть поле.

Напомним: *линейное пространство* L над полем F есть множество L с двумя операциями: сложением $\mathbf{x} + \mathbf{y}$: $L \times L \to L$ и умножением на константу $\mathbf{x} \cdot a$: $L \times F \to L$ (при этом $a \cdot \mathbf{x} = \mathbf{x} \cdot a$), удолетворяющими следующим аксиомам: $\forall \mathbf{x},\mathbf{y},\mathbf{z} \in L$, $\forall a,b \in F$

выполняются следующие равенства.

1) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.
2) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
3) $\exists\ \mathbf{0} \in L\ (\mathbf{x} + \mathbf{0} = \mathbf{x})$.
4) $\forall \mathbf{x} \in L\ \exists (-\mathbf{x}) \in L\ (\mathbf{x} + (-\mathbf{x}) = \mathbf{0})$.
5) $a \cdot (\mathbf{x} + \mathbf{y}) = a \cdot \mathbf{x} + a \cdot \mathbf{y}$.
6) $(a + b) \cdot \mathbf{x} = a \cdot \mathbf{x} + b \cdot \mathbf{x}$.
7) $a \cdot (b \cdot \mathbf{x}) = (a \cdot b) \cdot \mathbf{x}$.
8) $e \cdot \mathbf{x} = \mathbf{x}$, где e есть единица поля F.

Пример. Пусть E_2^n есть множество всех векторов $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ длины n из 0 и 1. Введем линейное пространство L_2^n векторов E_2^n над двухэлементным полем $F = (E_2, \{+, \cdot\})$ со сложением и умножением по модулю 2, определив (поразрядное) сложение и умножение векторов на константу из $\{0,1\}$, и положив $\forall \mathbf{a}, \mathbf{b} \in E_2^n$, $\forall c \in E_2$

$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$;
$c \cdot \mathbf{a} = (c \cdot a_1,\ c \cdot a_2,\ \ldots,\ c \cdot a_n)$.

Например, в E_2^4 для $\mathbf{a} = (1,0,0,1)$, $\mathbf{b} = (0,1,1,1)$ получим $\mathbf{a} + \mathbf{b} = (1,1,1,0)$; $0 \cdot \mathbf{a} = (0,0,0,0)$; $1 \cdot \mathbf{a} = \mathbf{a}$.

21.2. Линейное пространство подграфов данного графа

Пусть $G = (V, E)$ есть (p,q)-граф с множеством $V = \{v_1, \ldots, v_p\}$ из p вершин и с множеством $E = \{e_1, \ldots, e_q\}$ из q ребер. Пусть $H \subseteq G$ есть остовный подграф графа G. Подграфу H поставим в соответствие его характеристический вектор

$\mathbf{a}_H = (a_1, \ldots, a_q)$, где

$$a_i = \begin{cases} 1, & \text{если ребро } e_i \in H, \\ 0, & \text{если ребро } e_i \notin H, \end{cases} \quad i = 1,2,\ldots,q.$$

Между множеством L_G^q всех остовных подграфов графа G и множеством всех наборов из E_2^q можно установить взаимно однозначное соответствие, при котором вектору $\mathbf{a} = (a_1, \ldots, a_q)$ соответствует подграф $H_{\mathbf{a}} = (V, E_{\mathbf{a}})$, в котором ребро $e_i \in E_{\mathbf{a}} \longleftrightarrow a_i = 1$.

На множестве всех (остовных) подграфов графа G определим сложение и умножение на константу из $E_2 = \{0,1\}$ следующим образом. Пусть H, H_1, H_2 есть подграфы графа G и $\mathbf{a}_H, \mathbf{a}_{H_1}, \mathbf{a}_{H_2}$

есть им соответствующие характеристические векторы. Пусть $\lambda \in E_2$. Тогда

подграф $H_1 + H_2$ соответствует вектору $\mathbf{a}_{H_1} + \mathbf{a}_{H_2}$;

подграф $\lambda \cdot H$ соответствует вектору $\lambda \cdot \mathbf{a}_H$.

Подграф $0 \cdot H$ соответствует нуль-вектору $0 \cdot \mathbf{a}_H$ и потому содержит лишь вершины из G и ни одного ребра. Подграф $1 \cdot H$ совпадает с H, ибо его характеристический вектор $1 \cdot \mathbf{a}_H = \mathbf{a}_H$.

Линейное пространство наборов E_2^q изоморфно системе подграфов L_G^q и потому система L_G^q есть тоже линейное пространство. Размерность обоих линейных пространств равна q (число ребер в G).

На рис.21.1 приведены граф G, его подграфы H_1, H_2, сумма $H_1 + H_2$, их характеристические векторы $\mathbf{a}_{H_1}, \mathbf{a}_{H_2}, \mathbf{a}_{H_1} + \mathbf{a}_{H_2}$.

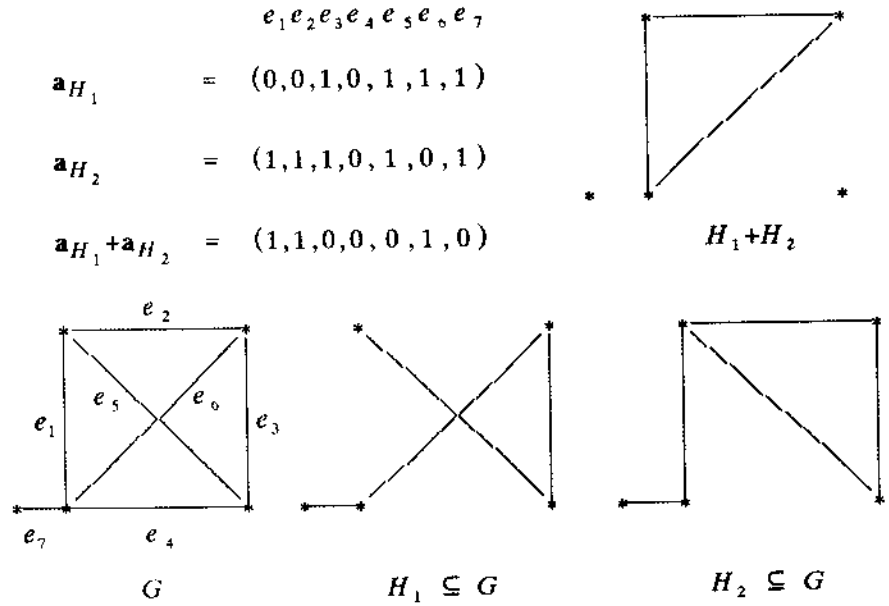


Рис.21.1

21.3. Подпространство четных подграфов

Пусть G есть (p,q)-граф и L_G^q есть линейное пространство всех (остовных) подграфов графа G.

Теорема. Множество $L_{чет}$ всех четных подграфов графа G образует подпространство линейного пространства L_G^q всех подграфов графа G.

Доказательство. Покажем, что множество $L_{чет}$ замкнуто относительно линейных операций сложения и умножения на константу.

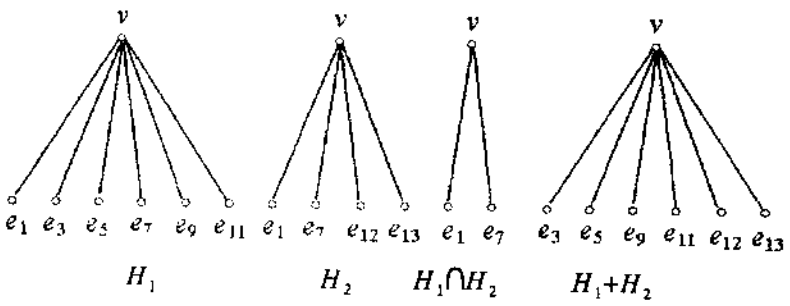
Пусть H_1, H_2 есть четные подграфы графа G, v – произвольная вершина в G, $s(v)$, $s_1(v)$, $s_2(v)$, $s_{12}(v)$ – степени вершины v в графах H_1+H_2, H_1, H_2, $H_1 \cap H_2$ соответственно (рис.21.2). Пусть $\mathbf{a}$, $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^{12}$ есть характеристические векторы этих подграфов. Для ситуации, изображенной на рис. 21.2, имеем следующее.

Степени: $s_1(v) = 6$, $s_2(v) = 4$, $s_{12}(v) = 2$.

Степень вершины v в графе H_1+H_2 складывается из числа ребер в вершине v графа H_1 плюс число ребер в вершине v графа H_2 минус удвоенное число ребер в вершине v графа $H_1 \cap H_2$, т.е. $s(v) = s_1(v) + s_2(v) - 2 \cdot s_{12}(v)$.

Получилось, что степень любой вершины v в графе H_1+H_2 есть число четное. Поэтому сумма H_1+H_2 есть четный граф.

Если H есть четный подграф графа G, то $0 \cdot H$ и $1 \cdot H$ есть четные подграфы графа G.



$$e_1\ e_2\ e_3\ e_4\ e_5\ e_6\ e_7\ e_8\ e_9\ e_{10}e_{11}e_{12}e_{13}\ \ldots$$

$$\mathbf{a}^1 = (1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ \ldots),$$

$$\mathbf{a}^2 = (1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ \ldots),$$

$$\mathbf{a}=\mathbf{a}^1+\mathbf{a}^2 = (0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 1,\ 1,\ \ldots).$$

Рис.21.2

Итак, множество $L_{чет}$ замкнуто относительно линейных операций (сложения и умножения на константу).

Все восемь аксиом линейного пространства выполняются для $L_{чет}$, ибо они выполняются для соответствующих векторов из E_2^q. Поэтому множество $L_{чет}$ есть подпространство линейного пространства L_G^q. Теорема доказана.

Замечание. Всякий (простой) цикл в графе G есть четный подграф графа G.

Определение. Матрица множества (простых) циклов графа G есть матрица, строки которой есть характеристические векторы (простых) циклов этого множества.

Утверждение. Пусть $C_1,\ldots,C_k$ есть циклы в графе G, попарно непересекающиеся по ребрам. Тогда справедливо следующее.

1. Объединение $C_1 \cup \ldots \cup C_k$ циклов совпадает с их суммой $C_1 + \ldots + C_k$.

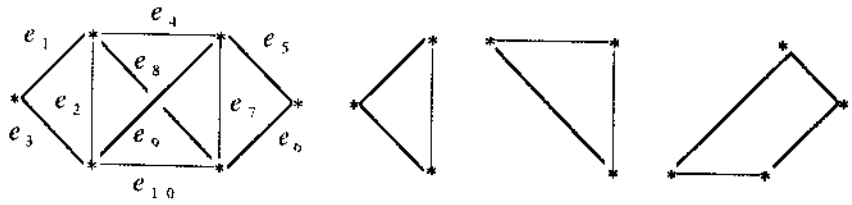
2. Циклы $C_1,\ldots,C_k$ линейно независимы.

Доказательство. На рис.21.3 приведены граф G и три его цикла C_1, C_2, C_3, попарно непересекающиеся по ребрам, и их характеристические векторы $\mathbf{a}_{C_1}, \mathbf{a}_{C_2}, \mathbf{a}_{C_3}$ соответственно. Видим, что каждый столбец их матрицы имеет не более одной единицы. Графы $C_1 \cup C_2$ и $C_1 + C_2$ совпадают.

1. Так как циклы $C_1,\ldots,C_k$ не имеют попарно общих ребер, то матрица их характеристических векторов $\mathbf{a}_{C_1},\ldots,\mathbf{a}_{C_k}$ в каждом столбце имеет не более одной единицы. Поэтому при сложении строк матрицы (т.е. при поразрядном сложении векторов по модулю два) ни одна единица (т.е. ни одно вхождение ребра) не пропадает. Тогда сложение характеристических векторов (сложение циклов) совпадает с объединением соответствующих циклов по общим вершинам.

2. Каждая строка в матрице характеристических векторов линейно независима от остальных строк, ибо если строка $\mathbf{a}_{C_j}$ в разряде j имеет 1, то в разряде j остальные строки матрицы имеют 0, и единица в разряде j вектора $\mathbf{a}_{C_j}$ не может быть получена никакой линейной комбинацией нулей в разряде j остальных векторов. Поэтому все строки в матрице, а, следовательно, и все циклы $C_1,\ldots,C_k$ линейно независимы.

Лемма. Множество всех простых циклов графа G составляет порождающую систему подпространства $L_{чет}$ всех четных подграфов графа G.



Граф G Цикл C_1 Цикл C_2 Цикл C_3

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}
$\mathbf{a}_{C_1}$ =	1	1	1	0	0	0	0	0	0	0
$\mathbf{a}_{C_2}$ =	0	0	0	1	0	0	1	1	0	0
$\mathbf{a}_{C_3}$ =	0	0	0	0	1	1	0	0	1	1
$\mathbf{a}_{C_1}+\mathbf{a}_{C_2}$ =	1	1	1	1	0	0	1	1	0	0

Рис.21.3

Доказательство. Пусть H есть произвольный четный подграф графа G. По теореме Эйлера граф H представим как объединение некоторых простых циклов $C_1,\ldots,C_k$ из H с попарно непересекающимися множествами ребер. По доказанному выше утверждению $H = C_1+\ldots+C_k$. Всякий цикл в H есть цикл в G. Поэтому граф H есть линейная комбинация множества простых циклов в G, причем в этой линейной комбинации коэффициенты 1 имеют простые циклы $C_1,\ldots,C_k$; остальные простые циклы из G имеют коэффициенты нуль. Лемма доказана.

Теорема. Подпространство $L_{чет}$ четных подграфов линейного просранства L_G^q всех (остовных) подграфов связного (p,q)-графа G имеет размерность $dim(L_{чет}) = q - p + 1$.

Доказательство. Пусть $G = (V,E)$ есть связный (p,q)-граф с p вершинами и q ребрами. Пусть подграф-дерево $T = (V,E_T)$ есть некоторый каркас графа G. Число ν хорд в G равно $q - p + 1$. Каждый простой цикл четен. Каждая хорда $e = (u,v)$ вместе с единственной простой цепью $[u,v]$ в дереве T образует простой цикл. Векторы таких циклов для разных хорд образуют линейно независимую систему Σ, ибо каждый из циклов содержит ребро (свою хорду), не принадлежащее ни одному из остальных циклов системы Σ. Мощность $|\Sigma| = \nu = q - p + 1$.

С другой стороны, каждый четный подграф H графа G линейно выражается через циклы системы Σ. В самом деле, пусть $e_1,\ldots,e_r$ есть все хорды в графе H. Прибавим к четному подграфу H графа G те циклы $C_1,\ldots,C_r$ из Σ, которые содержат хорды $e_1,\ldots,e_r$ из H. Тогда суммарный четный граф $H' = H + C_1 + \ldots + C_r$ не содержит ни одной хорды, т.е. H' есть подграф каркаса T, а всякое дерево имеет висячие вершины с нечетной степенью 1. Но граф H' четен как сумма четных графов. Поэтому граф H' пуст (не имеет ребер), т.е. в линейном пространстве L_G^q элемент H' есть нуль. Тогда $H' = H + C_1 + \ldots + C_r = 0$, откуда $H = -C_1-\ldots-C_r$, поэтому $H = C_1 + \ldots + C_r$ (ибо $-1 = 1$ по модулю два). Итак, всякий четный подграф H графа G линейно выражается через циклы системы Σ.

Получили, что Σ есть линейно независимая система циклов в G, причем любой четный подграф графа G линейно выражается через циклы системы Σ. Следовательно, Σ есть базис пространства $L_{чет}$ четных подграфов графа G. Так как мощность $|\Sigma| = q - p + 1$, то $dim(L_{чет}) = q - p + 1$.

Определение. Фундаментальная система циклов есть базис линейного пространства всех четных подграфов данного графа.

Замечание. Пусть несвязный (p,q)-граф G имеет k компонент связности $G_1,\ldots,G_k$, причем каждое G_i есть связный (p_i,q_i)-граф, $i = 1,\ldots,k$. Тогда $p = \sum_{i=1}^{k} p_i$, $q = \sum_{i=1}^{k} q_i$. Базис пространства всех четных подграфов графа G получается объединением базисов связных компонент в G. Поэтому

$$dim(L_{чет\ в\ G}) =$$

$$dim(\sum_{i=1}^{k}(L_{чет\ в\ G_i})) = \sum_{i=1}^{k}(q_i - p_i + 1) = \sum_{i=1}^{k} q_i - \sum_{i=1}^{k} p_i + \sum_{i=1}^{k} 1 =$$

$q - p + k$; т.е. $dim(L_{чет\ в\ G}) = q - p + k$.

21.4. Циклический ранг графа

Определение. *Циклический ранг* $CR(G)$ (*цикломатическое число*) графа G есть размерность подпространства четных подграфов линейного пространства всех подграфов графа G.

Справедливы следующие утверждения.

1. Для связного (p,q)-графа G:

а) $CR(G) = dim(L_{чет}) = q - p + 1$;

б) $CR(G)$ равно рангу матрицы простых циклов в G;

в) $CR(G)$ равно числу хорд в G;

г) $CR(G)$ равно максимальному числу линейно независимых простых циклов в G.

2. $CR(G) = dim(L_{чет}) \geqslant 0$.

3. Для связного (p,q)-графа G следующие утверждения эквивалентны:

а) граф G есть дерево;

б) $q - p + 1 = 0$;

в) $dim(L_{чет}) = 0$;

г) $CR(G) = 0$;

д) граф G не имеет циклов;

е) число хорд в G равно 0.

4. Для связного (p,q)-графа G следующие утверждения эквивалентны:

а) граф G имеет единственный цикл;

б) $q - p + 1 = 1$;

в) $dim(L_{чет}) = 1$;

г) $CR(G) = 1$;

д) ранг матрицы простых циклов в G равен 1;

е) число хорд в G равно 1.

5. Если (p,q)-граф G имеет k компонент связности G_i, $i = 1,2,\ldots,k$, причем каждое G_i есть (p_i,q_i)-граф, то $CR(G) =$

$$dim(L_{чет\ в\ G}) = \sum_{i=1}^{k} dim(L_{чет\ в\ G_i}) = \sum_{i=1}^{k} (q_i - p_i + 1) = q - p + k.$$

6. Фундаментальную систему циклов (базис подпространства $L_{чет}$ линейного пространства четных подграфов графа G) можно построить, взяв с каждой хордой графа G один из проходящих через эту хорду простых циклов. Фундаментальную систему циклов можно построить также, последовательно выделяя в G цикл, удаляя затем из G произвольное ребро (хорду) этого цикла, снова выделяя в получившемся графе цикл, и так далее, пока выделение циклов в последовательно получающихся графах возможно. Система получившихся циклов составит фундаментальную систему циклов графа G. Оставшийся после последовательного удаления из G хорд граф образует каркас графа G.

Пример 1. В ненагруженном графе G с помощью алгоритма удаления циклических ребер найти фундаментальную систему циклов, соответствующее множество хорд, каркас, все фундаментальные сечения (разрезы).

$G = (V,E)=(\{1,2,3,4,5,6\},\{(1,2),(1,4),(1,5),(1,6),(2,3),(2,5),(3,4),(3,6),(4,5),(5,6)\})$.

Решение.

Граф	Цикл	Удаляемое ребро
G	C_1=12341	e_1=(2,3)
$G_1=G-e_1$	C_2=1451	e_2=(1,5)
$G_2=G_1-e_2$	C_3=34563	e_3=(5,6)
$G_3=G_2-e_3$	C_4=14361	e_4=(3,6)
$G_4=G_3-e_4$	C_5=12541	e_5=(2,5)

Граф $G_5=G_4-e_5$ циклов не имеет. Множество $\{C_1,C_2,C_3,C_4,C_5\}$ составляет фундаментальную систему циклов графа G. Множество $\{e_1,e_2,e_3,e_4,e_5\}$ содержит все хорды графа G. Граф $G_5 = \{(1,2),(1,4),(1,6),(3,4),(4,5)\}$ есть каркас графа G. Всякий фундаметальный разрез составят хорды плюс одно произвольное ребро каркаса. Все фундаментальные разрезы:

$H \cup \{(1,2)\}$, $H \cup \{(1,4)\}$, $H \cup \{(1,6\}$, $H \cup \{(3,4)\}$, $H \cup \{(4,5)\}$.

Пример 2. В ненагруженном графе G с помощью алгоритма надстраивания ребер найти каркас, соответствующее множество хорд, фундаментальную систему циклов, все фундаментальные сечения (разрезы).

$G=(V,E)=(\{a,b,c,d,e,f,g\},\{(a,b),(a,g),(b,c),(b,d),(b,f),(b,g),(c,d),(c,g),(d,e),(d,f),(d,g),(e,f),(e,g),(f,g)\})$.

Решение. Исходим из ребра (a,b). Последовательное его расширение ребрами (избегаем при этом появления циклов) приводит нас к каркасу (стягивающему дереву)

$T = \{(a,b),(a,g),(d,g),(d,e),(c,d),(e,f)\}$.

Множество хорд

$H=E-T = \{(b,c),(b,d),(b,f),(b,g),(c,g),(d,f),(e,g),(f,g)\}$

Последовательно возвращаем в каркас по одной хорде и получаем фундаментальную систему из восьми циклов:

C_1=abcdga, C_2=abdga, C_3=abfedga, C_4=abga, C_5=cdgc, C_6=defd, C_7=degd, C_8=defgd.

Всякий фундаметальный разрез составят хорды плюс одно произвольное ребро каркаса. Все фундаментальные разрезы:

$H \cup \{(a,b)\}$, $H \cup \{(a,g)\}$, $H \cup \{(d,g)\}$, $H \cup \{(d,e)\}$, $H \cup \{(c,d)\}$, $H \cup \{(e,f)\}$.

Пример 3. В нагруженном графе G найти кратчайший каркас и

соответствующие множество хорд, фундаментальную систему циклов, все фундаментальные сечения (разрезы).

$G=(V,E)=(\{a,b,c,d,e,f,g\},\{(a,b,1),(a,g,2),(b,c,7),$
$(b,d,6),(b,f,8),(b,g,3),(c,d,2),(c,g,9),(d,e,1),(d,f,9),$
$(d,g,1),(e,f,4),(e,g,5),(f,g,9)\})$.

Решение. Исходим из ребра (a,b,1). Последовательное его расширение ребрами с наименьшим весом (избегаем при этом появления циклов) приводит нас к каркасу (стягивающему дереву)

$$T = \{(a,b,1),(a,g,2),(d,g,1),(d,e,1),(c,d,2),(e,f,4)\}$$

с наименьшим весом 11.

$H=E-T = \{(b,c,7),(b,d,6),(b,f,8),(b,g,3),(c,g,9),$
$(d,f,9),(e,g,5),(f,g,9)\}$ есть множество хорд.

Последовательно возвращаем в каркас по одной хорде и получаем фундаментальную систему из восьми циклов:

C_1=abcdga, C_2=abdga, C_3=abfedga, C_4=abga, C_5=cdgc,
C_6=defd, C_7=degd, C_8=defgd.

Всякий фундаметальный разрез составят хорды плюс одно произвольное ребро каркаса. Все фундаментальные разрезы:

$H\cup\{(a,b,1)\}$, $H\cup\{(a,g,2)\}$, $H\cup\{(d,g,1)\}$, $H\cup\{(d,e,1)\}$,
$H\cup\{(c,d,2)\}$, $H\cup\{(e,f,4)\}$.

21.5. Матричная теорема о деревьях

Теорема. (Кирхгоф). Пусть граф $G = (V,E)$ имеет множество вершин $V=\{v_1,\ldots,v_p\}$ и ребер E. Пусть

A есть матрица смежности (соседства вершин) графа G,

M есть матрица, полученная из матрицы $-A$ заменой элемента i главной диагонали на степень вершины v_i, то есть на число ребер, принадлежащих вершине v_i.

Стягиваюшее дерево графа G есть наименьшее по числу ребер подграф-дерево графа G, соединяющее все вершины в G.

Все алгебраические дополнения матрицы M равны между собой и их общее значение равно числу стягивающих деревьев (каркасов) графа G.

Пример. Для графа G с ниже заданной матрицей смежности (соседства вершин) вычисления дают следующее.

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{array}\!\!\begin{array}{c} \begin{array}{cccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}, \quad M = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{array}\!\!\begin{array}{c} \begin{array}{cccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \end{array} \\ \begin{bmatrix} 4 & -1 & 0 & -1 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 3 & -1 & 0 & -1 \\ -1 & 0 & -1 & 3 & -1 & 0 \\ -1 & -1 & 0 & -1 & 4 & -1 \\ -1 & 0 & -1 & 0 & -1 & 3 \end{bmatrix} \end{array},$$

$$A_{42} = (-1)^{4+2}\begin{vmatrix} 4 & 0 & -1 & -1 & -1 \\ -1 & -1 & 0 & -1 & 0 \\ 0 & 3 & -1 & 0 & -1 \\ -1 & 0 & -1 & 4 & -1 \\ -1 & -1 & 0 & -1 & 3 \end{vmatrix} = 135.$$

Все другие алгебраические дополнения матрицы M равны 135.

22. ДВУДОЛЬНЫЕ ГРАФЫ И ПАРОСОЧЕТАНИЯ

22.1. Двудольные графы

Определение. Граф $G = (V,E)$ называется *двудольным* (графом), если множество V его вершин допускает разбиение на два непересекающихся подмножества V_1 и V_2 (две доли), причем каждое ребро графа соединяет вершины из разных долей.

Обозначим через $G = (V_1,V_2,E)$ двудольный граф G с долями V_1 и V_2. Будем считать, что $|V_1| \leqslant |V_2|$.

Определение. Двудольный граф $G = (V_1,V_2,E)$ есть *полный* двудольный граф, если каждая вершина из V_1 соединена ребром с каждой вершиной из V_2.

Обозначим через $K_{n,m}$ полный двудольный граф $G = (V_1,V_2, E)$, у которого $n = |V_1|$, $m = |V_2|$. *Звезда* есть полный двудольный граф $K_{1,p}$. Несколько примеров двудольных графов приведено на рис.22.1.

Теорема. Граф G двудолен тогда и только тогда, когда все простые циклы в G имеют четную длину (четное число ребер).

Доказательство. Пусть граф $G = (V,E) = (V_1,V_2,E)$ двудолен и пусть $C = v_1e_1v_2e_2v_3e_3v_4\ldots v_{n-1}e_{n-1}v_n$ (причем $v_1 = v_n$) есть простой цикл в G длины $n-1$. Тогда $v_1,v_3,v_5,v_7,\ldots \in V_1$, а $v_2,v_4,v_6,\ldots \in V_2$, т.е. вершины с нечетными индексами лежат в V_1, а вершины с четными индексами лежат в V_2 (рис. 22.2, 22.3). Так как $v_n = v_1 \in V_1$, то число n нечетно. Поэтому длина $n-1$ цикла C есть число четное.

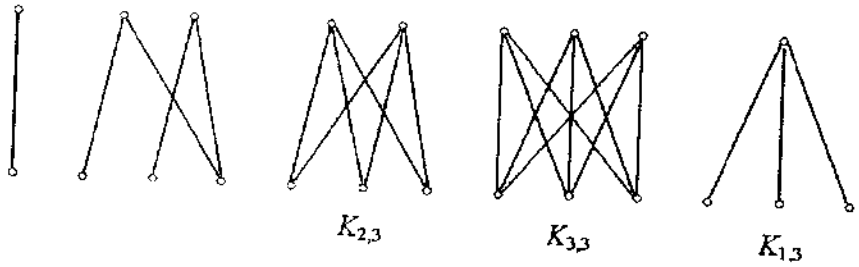


Рис.22.1

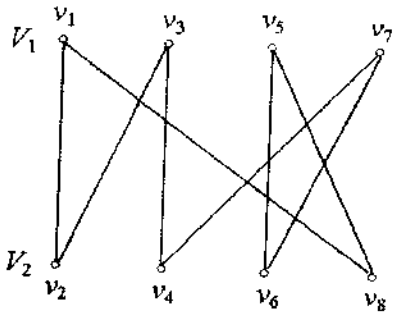


Рис.22.2

Рис.22.3

Пусть теперь $G = (V,E)$ есть граф, все простые циклы которого имеют четную длину. Можно считать, что граф G связен, ибо каждую компоненту связности графа G можно рассматривать по отдельности.

Пусть $v_1 \in V$ есть произвольная вершина в связном графе G, и пусть множество вершин $V_1 = \{v \in V$: расстояние $d(v,v_1)$ между v и v_1 четно$\}$. Тогда $V_2 = V - V_1$ есть множество вершин, находящихся от v_1 на нечетном расстоянии. Покажем, что никакие две вершины в V_2 не соединимы в G ребром. Допустим противное: существует пара вершин $u,v \in V_2$, соединимых в G ребром $e = (u,v)$. Пусть $[v_1,u]$ и $[v_1,v]$ на рис.22.3 есть две геодезические (нечетной длины). Тогда $[v_1,u] \cup \{e\} \cup [v,v_1]$ есть простой цикл в G нечетной длины. Противоречие с условием. Следовательно, любые две вершины в V_2 не смежны. Аналогично можно показать, что любые две вершины в V_1 не смежны. Следовательно, ребра в графе G соединяют вершины из разных классов, т.е. граф G двудолен.

22.2. Паросочетания

Определение. *Паросочетание* есть двудольный граф, в котором никакие два ребра не являются смежными (рис.22.4). Паросочетание P для графа $G = (V,E)$ есть любое множество попарно несмежных ребер в G. P есть *наибольшее паросочетание* для G, если число ребер в P наибольшее среди всех паросочетаний для G. P есть *максимальное* (тупиковое) *паросочетание* для G, если к P нельзя добавить ни одного ребра из G, не нарушив паросочетаемости. P есть *совершенное паросочетание* для G, если P имеет $|V|$ вершин.

Замечание. Наибольшее паросочетание максимально. Совершенное паросочетание является и наибольшим, и максимальным (рис.22.5).

Простая цепь C ненулевой длины в G, ребра которой попеременно лежат и не лежат в P, называется чередующейся цепью (относительно паросочетания P). Эта цепь C называется P-увеличителем, если первое и последнее ребро цепи C лежат вне P. С помощью P-увеличителя C паросочетание P можно переделать в другое паросочетание P' для G с числом ребер в P' на единицу больше, чем в P. Для этого достаточно все ребра в C, лежащие вне P, добавить к P, а ребра в C, лежащие в P, удалить из P.

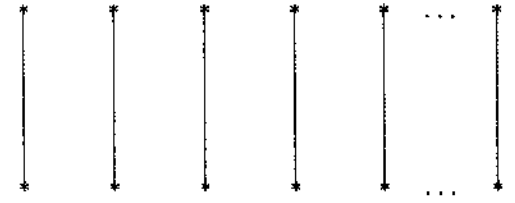
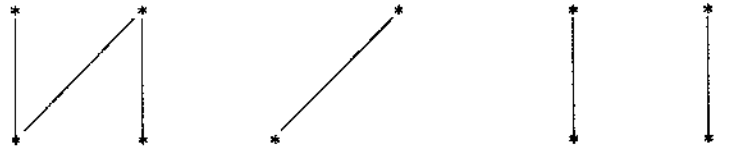


Рис.22.4



Граф G | Максимальное паросочетание в G | Совершенное паросочетание в G

Рис.22.5

Для получившегося паросочетания P' можно снова искать увеличитель, и так далее, последовательно расширяя получающиеся паросочетания, пока это возможно.

Теорема. Если паросочетание P для графа G не является наибольшим, то граф G имеет P-увеличитель.

Доказательство. Если паросочетание P (заданное множеством P своих ребер) для G не является наибольшим, то G имеет другое паросочетание P' с большим числом ребер, чем в P. В подграфе графа G, порожденном множеством ребер $(P-P') \cup (P'-P)$, степень любой вершины не больше 2, следовательно, каждая компонента связности этого подграфа есть простая цепь или простой цикл с чередованием ребер из P и из P'. Среди этих компонент обязательно найдется простая цепь нечетной длины, начинающаяся и оканчивающаяся ребрами из P', ибо в противном случае было бы $|P'| \leqslant |P|$ вопреки предположению; эта цепь и является искомым P-увеличителем в G.

Следствие. Если граф G не имеет P-увеличителя, то паросочетание P для графа G является наибольшим.

Приведем без доказательства следующее утверждение.

Теорема. Граф $G = (V,E)$ имеет совершенное паросочетание тогда и только тогда, когда $\forall\, V' \subseteq V$ $(n(G-V') \leqslant |V'|)$, где $n(G-V')$ есть число тех компонент связности подграфа $G - V'$ графа G, которые имеют нечетное число вершин.

Определение. Паросочетание P для двудольного графа $G = (V_1,V_2,E)$ есть любое множество попарно несмежных ребер в G. P есть *наибольшее паросочетание* для G, если число ребер в P наибольшее среди всех паросочетаний для G. P есть *максимальное* (тупиковое) *паросочетание* для G, если к P нельзя добавить ни одного ребра из G, не нарушив свойства паросочетаемости. P есть *совершенное паросочетание* для G, если P имеет $|V_1|$ ребер.

22.2.1. Алгоритм построения совершенного паросочетания для двудольного графа

Пусть $G = (U,V,E)$ есть двудольный граф. Выберем исходное паросочетание P_1, например, одно ребро графа G. Допустим, что паросочетание $P_i = (U_i,V_i,E_i)$ для графа G построено.

Построим паросочетание P_{i+1} для G следующим образом.

1. Выбираем u из U не из P_i. Если такой вершины u нет, то P_i есть совершенное паросочетание. Строим в G чередующуюся цепь $\mu_i = [u_1,v_1,u_2,v_2,\ldots,u_p,v_p]$ с $u_1 = u$, в которой всякое

ребро (u_i,v_i) не принадлежит E_i, а всякое ребро (v_i,u_{i+1}) принадлежит E_i. Если такой цепи нет, то совершенного паросочетания граф G не имеет, а паросочетание P_i является для G максимальным (тупиковым). Цепь μ_i есть P_i-увеличитель.

2. Удаляем из P_i все ребра (v_i,u_{i+1}) и добавляем все ребра (u_i,v_i) цепи μ_i. Получившееся паросочетание P_{i+1} на одно ребро длиннее паросочетания P_i. Переходим к п. 1.

Пример. Построим совершенное паросочетание для двудольного графа

$$G = (U,V,E) = (\{x_1,x_2,x_3,x_4,x_5,x_6\},\{y_1,y_2,y_3,y_4,y_5,y_6,y_7\},$$
$$\{(x_1,y_1),(x_1,y_2),(x_1,y_5),(x_2,y_1),(x_2,y_3),(x_2,y_5),(x_3,y_1),$$
$$(x_3,y_6),(x_4,y_3),(x_4,y_4),(x_4,y_6),(x_4,y_7),(x_5,y_5),(x_5,y_7),$$
$$(x_6,y_4),(x_6,y_6),(x_6,y_7)\}).$$

Шаг 1. Выбираем исходное паросочетание $P_1 = \{(x_1,y_1)\}$.

P_1-увеличитель (чередующаяся цепь)

$$\mu_1 = [x_2,y_1,x_1,y_5].$$
$$0\ \ 1\ \ 0$$
$$1\ \ 0\ \ 1$$

Единственная единица в первой строке из нулей и единиц означает, что соответствующее этой единице ребро (y_1,x_1) лежит в P_1. Убираем это ребро из P_1, а вместо него добавляем два ребра $(x_2,y_1),(x_1,y_5)$, соответствующие двум единицам второй строки из нулей и единиц. В результате получим следующее паросочетание P_2, число ребер в котором на одно больше чем в P_1.

Шаг 2. $P_2 = \{(x_1,y_5),(x_2,y_1)\}$.

$$\mu_2 = [x_3,y_1,x_2,y_3].$$
$$0\ \ 1\ \ 0$$
$$1\ \ 0\ \ 1$$

Удаляем из P_2 ребро (x_2,y_1) и добавляем вместо него ребра (x_3,y_1), (x_2,y_3).

Шаг 3. $P_3 = \{(x_1,y_5),(x_2,y_3),(x_3,y_1)\}$.

$$\mu_3 = [x_4,y_4].$$

Добавляем в P_3 ребро (x_4,y_4).

Шаг 4. $P_4 = \{(x_1,y_5),(x_2,y_3),(x_3,y_1),(x_4,y_4)\}$.

$$\mu_4 = [x_5,y_5,x_1,y_1,x_3,y_6].$$
$$0\ \ 1\ \ 0\ \ 1\ \ 0$$
$$1\ \ 0\ \ 1\ \ 0\ \ 1$$

Удаляем из P_4 ребра (x_1,y_5), (x_3,y_1) и добавляем вместо них ребра (x_5,y_5), (x_1,y_1), (x_3,y_6).

Шаг 5. $P_5 = \{(x_1,y_1),(x_2,y_3),(x_3,y_6),(x_4,y_4),(x_5,y_5\}$.

$$\mu_5 = [x_6,y_6,x_3,y_1,x_1,y_5,x_5,y_7].$$
$$0\ 1\ 0\ 1\ 0\ 1\ 0$$
$$1\ 0\ 1\ 0\ 1\ 0\ 1$$

Удаляем из P_5 ребра (x_3,y_6), (x_1,y_1), (x_5,y_5) и добавляем вместо них ребра (x_6,y_6), (x_3,y_1), (x_1,y_5), (x_5,y_7).

Шаг 6. $P_6 = \{(x_1,y_5),(x_2,y_3),(x_3,y_1),(x_4,y_4),(x_5,y_7),(x_6,y_6)\}$. P_6 есть искомое совершенное паросочетание для исходного графа.

22.3. Системы различных представителей

Определение. *Система различных представителей* (СРП) для семейства конечных множеств $S = \{A_1,A_2,\ldots,A_i,\ldots,A_m\}$ есть система попарно различных элементов $\{a_1,a_2,\ldots,a_i,\ldots,a_m\}$, для которой $a_i \in A_i$, $i = 1,2,\ldots,m$.

Примеры.

1. A_1={1,4}; A_2 = {1,2,5}; A_3 = {5,6}; СРП = {1,2,5}.
2. A_1 = {1,3}; A_2 = {1,3}; A_3 = {3,4}; A_4 = {1,4}. СРП нет.

По семейству множеств $S = \{A_1,\ldots,A_m\}$ построим двудольный граф $G = (V_1,V_2,E)$, положив (рис.22.6)

$$V_1 = S = \{A_1,A_2,\ldots,A_m\};\ V_2 = \bigcup_{i=1}^{m} A_i = \{a_1,a_2,\ldots,a_r\};$$

$$\text{ребро } e = (A_i,a_j) \in E \longleftrightarrow a_j \in A_i.$$

Семейство S имеет СРП тогда и только тогда, когда граф G имеет совершенное паросочетание.

Для двудольного графа $G = (V_1,V_2,E)$ можно построить семейство множеств $S(u) = \{v \in V_2 : \text{ребро } e = (u,v) \in E\}$, $u \in V_1$. Тогда семейство множеств S имеет СРП тогда и только тогда, когда двудольный граф G имеет совершенное паросочетание. Так что вопрос о наличии совершенного паросочетания у двудольного графа G эквивалентен вопросу о наличии СРП для семейства множеств S, порожденным графом G описанным выше способом.

Теорема (Радо). Пусть $A_1,\ldots,A_m$ есть семейство S конечных множеств. Семейство S имеет СРП тогда и только тогда, когда

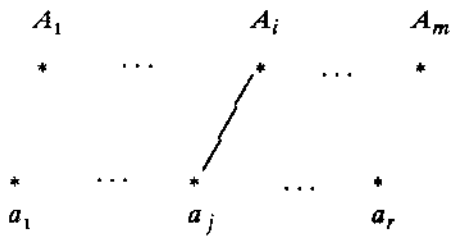


Рис.22.6

$\forall k = 1,2,\ldots,m$ объединение любых k множеств этого семейства имеет по крайней мере k элементов.

Коротко: семейство S имеет СРП тогда и только тогда, когда $\forall I \subseteq \{1,2,\ldots,m\}$ $(|I| \leqslant |\bigcup_{i\in I} A_i|)$.

Доказательство. Пусть $a \in A_m$, $b \in A_{m-1}$,

$$S' = \{A_1',\ldots,A_{m-1}',A_m'\} = \{A_1,\ldots,A_{m-1},A_m-\{a\}\};$$
$$S'' = \{A_1'',\ldots,A_{m-1}'',A_m''\} = \{A_1,\ldots,A_{m-1}-\{b\},A_m-\{a\}\};$$

Лемма. Если

1) $\forall\ I \subseteq \{1,2,\ldots,m\}$ $(|\bigcup_{i\in I} A_i| \geqslant |I|)$,

2) $|A_m| \geqslant 2$,

то $\exists a \in A_m\ \forall I \subseteq \{1,2,\ldots,m\}$ $(|\bigcup_{i\in I} A_i'| \geqslant |I|)$.

Доказательство. Допустим противное:

$$\forall a \in A_m\ \exists I \subseteq \{1,2,\ldots,m\}\ (|\bigcup_{i\in I} A_i'| < |I|). \qquad (22.1)$$

Пусть $A_m = \{a_1,\ldots,a_r\}$; $M = \{1,2,\ldots,m\}$. Из условия (22.1) имеем

$$\text{для } a = a_1 \in A_m\ \exists I = J_1' \subseteq M\ (|\bigcup_{i\in J_1'} A_i'| < |J_1'|);$$

$$\text{для } a = a_2 \in A_m\ \exists I = J_2' \subseteq M\ (|\bigcup_{i\in J_2'} A_i'| < |J_2'|); \qquad (22.2)$$

$$\ldots$$

$$\text{для } a = a_r \in A_m\ \exists I = J_r' \subseteq M\ (|\bigcup_{i\in J_r'} A_i'| < |J_r'|)$$

Возможны следующие случаи:

а) $\exists l \in \{1,\ldots,r\}$ $(m \notin J_l')$, т.е. $J_l' \subseteq \{1,2,\ldots,m-1\}$. Тогда

по формуле (12.2) будем иметь

$$\text{для } a=a_i\in A_m\ \exists I=J_i'\subseteq\{1,2,\ldots,m-1\}\ \left(\left|\bigcup_{i\in J_i'} A_i'\right| < |J_i'|\right). \quad (22.3)$$

Из (22.1) для $I=J_i' \subseteq \{1,2,\ldots,m-1\}$ имеем $\left|\bigcup_{i\in J_i'} A_i\right| = \left|\bigcup_{i\in J_i'} A_i'\right| \geqslant |J_i'|$. Противоречие с (22.3);

б) $\forall I \subseteq \{1,2,\ldots,r\}$ $(m \in J_i')$. Пусть $J_i = J_i' - \{m\}$. Тогда $J_i' = J_i \cup \{m\}$, $J_i \subseteq \{1,2,\ldots,m-1\}$, $|J_i'| = |J_i|+1$. Так как $|A_m| \geqslant 2$, то A_m имеет два различных элемента a,b. Из формулы (22.1) имеем:

$$\text{для } a \in A_m\ \exists I = J' \subseteq \{1,2\ldots,m\}\ \left(\left|\bigcup_{i\in J'} A_i'\right| < |J'|\right).$$

При $J = J'-\{m\}$; $J \subseteq \{1,2,\ldots,m-1\}$; $|J'| = |J|+1$; для $i\in J$ $A_i' = A_i$; $\left|\bigcup_{i\in J'} A_i'\right| = \left|\left(\bigcup_{i\in J} A_i'\right) \cup (A_m-\{a\})\right| = \left|\left(\bigcup_{i\in J} A_i\right) \cup (A_m - \{a\})\right|$. Поэтому

$$\text{для } a\in A_m\ \exists J\subseteq\{1,2\ldots,m-1\}\ \left(\left|\left(\bigcup_{i\in J} A_i\right)\cup(A_m-\{a\}\right|\leqslant|J|\right). \quad (22.4)$$

Аналогично

$$\text{для } b\in A_m\ \exists K\subseteq\{1,2\ldots,m-1\}\ \left(\left|\left(\bigcup_{i\in K} A_i\right)\cup(A_m-\{b\}\right|\leqslant|K|\right). \quad (22.5)$$

Сложим (22.4) и (22.5), в результате получим: $|J| + |K| \geqslant$

$$\left|\left(\bigcup_{i\in J} A_i\right) \cup (A_m-\{a\})\right| \cup \left|\left(\bigcup_{i\in K} A_i\right) \cup (A_m-\{b\}\right| \geqslant$$

$$\left|\left(\bigcup_{i\in J} A_i\right) \cup A_m\right| \cup \left|\bigcup_{i\in K} A_i\right| \underset{\text{по (12.2)}}{\geqslant} |J|+1+|K|, \text{ откуда}$$

$|J|+|K| \geqslant |J|+|K|+1$. Противоречие. Поэтому наше предположение неверно и заключение леммы справедливо.

Продолжим доказательство теоремы.

Необходимость очевидна, ибо если семейство $A_1,A_2,\ldots,A_m$ имеет СРП $\{a_1,a_2,\ldots,a_m\}$, то

$$\forall I \subseteq \{1,2,\ldots,m\}\ \left|\bigcup_{i\in I} A_i\right| \geqslant \left|\bigcup_{i\in I}\{a_i\}\right| = |I|.$$

Пусть теперь для конечной системы S множеств $A_1,A_2,\ldots,A_m$ выполняется условие $\forall I \subseteq \{1,2,\ldots,m\}$ $\left(\left|\bigcup_{i\in I} A_i\right| \geqslant |I|\right)$. Покажем, что система S имеет СРП. Индукция по наибольшему из чисел $|A_i|$.

Базис. $\max\limits_{i=1,\ldots,m} |A_i| = 1$. Тогда $|A_1| = \ldots = |A_m| = 1$. Одноэлементные множества $A_1,\ldots,A_m$ тривиально имеют СРП.

Предположение индукции. Пусть утверждение достаточности справедливо для всех семейств S множеств $A_1,\ldots,A_m$, для которых $\max\limits_{i=1,\ldots,m} |A_i| < q$.

Шаг индукции. Покажем, что утверждение достаточности справедливо для всех семейств S с $\max\limits_{i=1,\ldots,m} |A_i| = q \geqslant 2$. Пусть d есть число всех множеств семейства S, которые имеют мощность q. Пусть для простоты $d = 2$ и $|A_m| = |A_{m-1}| = q$. Применяя тогда два раза лемму, получаем

$$\exists a \in A_m \quad \forall I \subseteq \{1,2,\ldots,m\}\ \left(\left|\bigcup_{i\in I} A_i'\right| \geqslant |I|\right),$$

$$\exists b \in A_{m-1}\ \forall I \subseteq \{1,2,\ldots,m\}\ \left(\left|\bigcup_{i\in I} A_i''\right| \geqslant |I|\right),$$

где $S' = \{A_1',\ldots,A_{m-1}',A_m'\} = \{A_1,\ldots,A_{m-1},A_m-\{a\}\}$,

$S'' = \{A_1'',\ldots,A_{m-1}'',A_m''\} = \{A_1,\ldots,A_{m-2},A_{m-1}-\{b\},\ A_m-\{a\}\}$.

Так как в системе $S'' = \{A_1'',\ldots,A_m''\}$ $\max\limits_{i=1,\ldots,m} |A_i''| = q-1 < q$ и $\forall I \subseteq \{1,2,\ldots,m\}$ $\left(\left|\bigcup_{i\in I} A_i''\right| \geqslant |I|\right)$, то по предположению индукции семейство $S'' = \{A_1'',\ldots,A_{m-2}'',A_{m-1}'',A_m''\} = \{A_1,\ldots,A_{m-2},A_{m-1}-\{b\},\ A_m-\{a\}\}$ имеет СРП $\{a_1,\ldots,a_{m-2},a_{m-1},a_m\}$, которая есть СРП и для семейства S. Теорема доказана.

Пусть $G = (V_1,V_2,E)$ есть двудольный граф; вершина $u \in V_1$; множество вершин $A \subseteq V_1$. Положим множество $S(u) = \{v\in V_2 : u$ смежна с $v\}$; $S(A) = \bigcup\limits_{u\in A} S(u)$.

Теорема (Холла). Пусть $G = (V_1, V_2, E)$ есть двудольный граф. Тогда граф G имеет совершенное паросочетание тогда и только тогда, когда $\forall\ A \subseteq V_1\ (|A| \leqslant |S(A)|)$.

Доказательство. Граф G имеет совершенное паросочетание тогда и только тогда, когда семейство множеств $S(u) = \{v \in V_2$: u смежна с $v\}$, $u \in V_1$, имеет СРП. Последнее по теореме Радо справедливо тогда и только тогда, когда

$$\forall A \subseteq V_1\ (|A| \leqslant |\bigcup_{u \in A} S(u)|), \text{ т.е. когда}$$

$$\forall A \subseteq V_1\ (|A| \leqslant |S(A)|).$$

Пример. Для указанных множеств найти систему различных представителей. A_1={1,2,5}, A_2={1,3,5}, A_3={1,6}, A_4={3,4,6, 7}, A_5={5,7}, A_6={4,6,7}.

Решение. Пусть множества вершин

$$U = \{A_1, A_2, A_3, A_4, A_5, A_6\},\ V = \bigcup_{i=1}^{6} A_i = \{1,2,3,4,5,6,7\},$$

множество E ребер таково, что $(A_i, j) \in E \longleftrightarrow j \in A_i$. Тогда

$E = \{(A_1,1),(A_1,2),(A_1,5),(A_2,1),(A_2,3),(A_2,5),$
$(A_3,1),(A_3,6),(A_4,3),(A_4,4),(A_4,6),(A_4,7),(A_5,5),$
$(A_5,7),(A_6,4),(A_6,6),(A_6,7)\}$).

Двудольный граф $G = (U,V,E)$ есть двудольный граф из параграфа 22.2.1. Его совершенное паросочентание

$P = \{(A_1,5),(A_2,3),(A_3,1),(A_4,4),(A_5,7),(A_6,6)\}$.

Система различных представителей:

$5 \in A_1\{1,2,5\}$, $3 \in A_2$={1,3,5}, $1 \in A_3$={1,6},
$4 \in A_4$={3,4,6,7}, $7 \in A_5$={5,7}, $6 \in A_6$={4,6,7}.

23. ПЛАНАРНЫЕ ГРАФЫ

23.1. Плоские графы

Определение. Граф G изоморфно укладывается на плоскость, если его можно изобразить на плоскости так, чтобы никакие его ребра не пересекались (кроме соединений ребер в вершинах). Граф G, уложенный на плоскости, называется *плоским изображением* графа G. Граф G *плоский*, если он изображен на плоскости без пересечений ребер. Граф G *планарен* (рис.23.1),

если G изоморфно укладывается на плоскость.

Всякий плоский граф планарен. Всякий подграф планарного (плоского) графа планарен (плоский). Плоский граф иногда называют плоской картой.

Определение. *Грань* плоского графа есть часть плоскости, ограниченная простым циклом графа. Конечная грань называется внутренней. Бесконечная грань называется внешней. Простой цикл, ограничивающий грань, называется *границей грани*.

Дерево имеет единственную (бесконечную) внешнюю грань.

На рис.23.2 изображен плоский граф с внутренними гранями 1,2,3 и внешней гранью 4.

Аналогичные определения можно ввести для мультиграфов и псевдографов. Возможны укладки графов без пересечений ребер и на другие поверхности.

23.2. Формула Эйлера

Теорема. В любом связном плоском графе числа p,q,r его вершин, ребер, граней соответственно связаны равенством (Эйлера) $p - q + r = 2$.

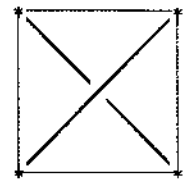
Доказательство. (Индукция по числу q ребер в графе).

Базис. $q = 0$. Так как граф состоит из единственной вершины, то $p - q + r = 1 - 0 + 1 = 2$.

Предположение индукции. Допустим, что формула Эйлера справедлива для всех связных плоских графов с числом ребер меньше q.

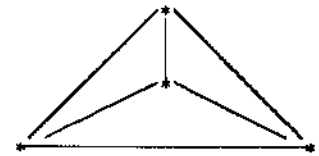
Шаг индукции. Покажем, что формула Эйлера справедлива для всех связных плоских графов с числом ребер q. Пусть плоский граф G имеет p вершин, q ребер и r граней. Возможны следующие случаи.

1. Граф G имеет простой цикл. Удалим из G циклическое ребро e. Граф $G' = G - e$ связен, имеет p вершин, $q-1$ ребер и



Планарный граф G

Рис.23.1



Плоское изображение G

Рис.23.2

$r-1$ граней. По предположению индукции для графа G' формула Эйлера справедлива; тогда $p - (q - 1) + (r - 1) = 2$; отсюда $p - q + r = 2$.

2. Связный граф G простых циклов не имеет. Тогда граф G есть дерево с единственной внешней гранью (рис.23.3). Все ребра в G ациклични. Удалим из G любое ребро e. Получим несвязный граф $G' = G - e$ с двумя компонентами связности G_1 и G_2 и числами p_1,q_1,r_1 и p_2,q_2,r_2 вершин, ребер и граней соответственно (рис.23.3). Так как $q_1,q_2 < q$, то по предположению индукции для графов G_1,G_2 формула Эйлера справедлива:

$$\left.\begin{array}{l} p_1 - q_1 + r_1 = 2 \\ p_2 - q_2 + r_2 = 2 \end{array}\right\} \rightarrow \left.\begin{array}{l} p_1 - q_1 + 1 = 2 \\ p_2 - q_2 + 1 = 2 \end{array}\right\} \rightarrow \text{(сложим)}$$

$$(p_1 + p_2) - (q_1 + q_2) + 2 = 4 \rightarrow p - (q_1 + q_2 + 1) + 1 = 2 \rightarrow$$

$$p - q + 1 = 2 \rightarrow p - q + r = 2.$$

Шаг индукции установлен. Теорема доказана.

Следствие. Если G есть связный плоский (p,q)-граф; каждая внутренняя грань в G есть n-цикл (простой цикл длины n), то $q \leqslant (n(p - 2))/(n-2)$.

Доказательство. Так как каждая внутренняя грань в G имеет n ребер, внешняя грань в G имеет не менее n ребер и любое ребро в G принадлежит двум граням, то $n \cdot r \leqslant 2q$, ибо каждое ребро слева входит дважды (рис.23.4). Тогда $r \leqslant (2 \cdot q)/n$. По формуле Эйлера $p - q + (2 \cdot q)/n \geqslant 2$, ибо $(2 \cdot q)/n \geqslant r$. Тогда $q \leqslant (n(p-2))/(n-2)$.

Утверждение. Граф K_5 не планарен.

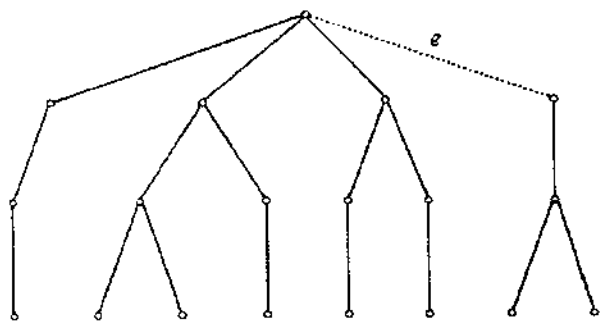
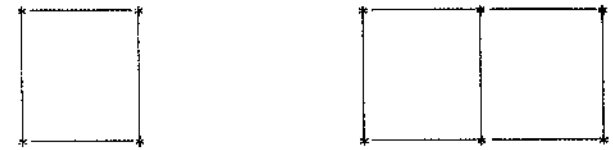


Рис.23.3



$nr=2q$ $nr\leqslant 2q$

Рис.23.4

Доказательство. Допустим противное: граф K_5 планарен и G есть его плоское изображение. Так как графы G и K_5 изоморфны, то каждая грань в G есть 3-цикл. Подставляя $n = 3$, $p = 5$, $q = 10$ в формулу следствия, получим $10 \leqslant (3 \cdot (5-2))/(3-2) = 9$. Противоречие. Поэтому граф K_5 планарным не является.

Утверждение. Граф $K_{3,3}$ не планарен.

Доказательство. Допустим противное: граф $K_{3,3}$ планарен и имеет плоское изображение G. Так как $K_{3,3}$ и G изоморфны, то каждая грань в G есть 4-цикл. Подставляя в формулу следствия значения $n = 4$, $p = 6$, $q = 9$, получим $9 \leqslant (4 \cdot (6-2))/(4-2) = 8$. Противоречие. Следовательно, граф $K_{3,3}$ не планарен.

Определение. Граф G есть *максимальный* планарный (плоский) граф, если граф G планарный (плоский), но при добавлении к G любого ребра он перестает быть планарным (плоским).

Замечание. Любая грань максимального плоского (p,q)-графа G есть 3-цикл (треугольник). Подставляя в формулу следствия $n = 3$, получим $q \leqslant (3(p-2))/(3-2) = 3p - 6$, т.е. для максимального плоского (p,q)-графа G число ребер $q \leqslant 3p - 6$. Тогда для немаксимального плоского (p,q)-графа G (который получается удалением некоторых ребер из некоторого максимального графа) число ребер $q \leqslant 3p-6$ тем более. Так как планарный граф изоморфен некоторому плоскому графу, то для планарного (p,q)-графа G число ребер $q \leqslant 3p-6$ тоже.

Утверждение. Каждый планарный (p,q)-граф имеет вершину степени $s \leqslant 5$.

Доказательство. Допустим противное: все вершины планарного (p,q)-графа G имеют степень $s \geqslant 6$. Так как любая вершина в G имеет степень $s \geqslant 6$ и так как каждое ребро соединяет две вершины, то в G число ребер $q \geqslant (6p)/2 = 3p$ и потому $q = 3p + k$ для некоторого $k \geqslant 0$. Подставляя это значение q в неравенство $q \leqslant 3p - 6$, получим $3p+k \leqslant 3p-6$, откуда $0 \leqslant k \leqslant -6$. Противоречие. Следовательно, планарный (p,q)-граф G имеет вершину степени $s \leqslant 5$.

23.3. Критерий планарности Понтрягина–Куратовского

Вершину степени 2 в графе назовем проходной. Операция добавления проходной вершины в ребро $e = (u,v)$ в графе $G = (V,E)$ состоит в удалении ребра e из E, добавлении к V новой вершины w и в добавлении к графу $G-e$ двух новых ребер (u,w) и (w,v).

Операция удаления проходной вершины v в графе $G = (V,E)$ состоит в удалении вершины v из V и замене двух ей инцидентных (принадлежащих) в E ребер $e' = (u,v)$, $e'' = (v,w)$ на одно ребро $e = (u,w)$.

Операция стягивания ребра $e = (u,v)$ в графе $G = (V,E)$ состоит в удалении ребра e из E и в объединении (склеивании) инцидентных ребру e вершин.

Два графа *гомеоморфны*, если один из них может быть получен из другого применением конечного числа раз операций добавления и исключения проходных вершин и стягивания ребра. Если графы гомеоморфны, то они планарны или не планарны одновременно.

Теорема (Понтрягина–Куратовского). Граф G планарен тогда и только тогда, когда G не содержит подграфов, гомеоморфных графам K_5 или $K_{3,3}$.

Например, граф Петерсена (рис.23.5) не планарен.

Аналогично стоят вопросы укладки графов и на другие поверхности, например, на сферу, тор и так далее.

23.3.1. Алгоритм построения плоского изображения графа

Изложим алгоритм построения плоского изображения графа. Пусть $G=(V,E)$ есть исходный граф, плоское изображение которого нам требуется построить (если оно имеется). Будем предполагать, что граф G связен, не имеет висячих вершин и точек сочленения, т.е. вершин, удаление которых из G вместе с принадлежащими им ребрами приводит к несвязному графу.

Пусть $G' = (V',E')$ есть некоторый плоский подграф графа G. Остаток графа G относительно G' есть граф $R = (V'',E'') = (V-V',E'')$ – подграф графа G, порожденный подмножеством вершин $V-V'$, т.е. R состоит из всех тех ребер графа G, концы которых не лежат в V' (т.е. лежат вне V').

Кусок P графа G относительно его подграфа G' есть один из следующих объектов:

1) компонента связности остатка R относительно G', дополненная теми ребрами графа G, которые соединяют вершины этой компоненты и вершины V' графа G';

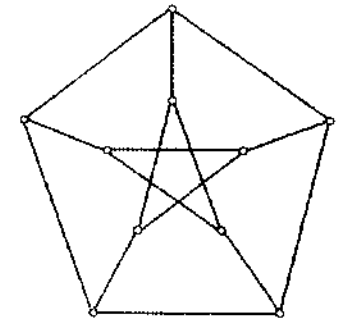


Рис.23.5.

2) одно ребро из $E-E'$ с концами, лежащими в V'.

Контактные точки куска P есть вершины, общие для P и G'. Грань F в G' совместима с куском P, если все контактные точки куска P принадлежат грани F.

Пусть G_1 есть некоторый простой цикл графа G. Поместим на плоскости его плоское изображение.

Допустим что плоский граф G_i уже построен. Плоский граф G_{i+1} получим следующим образом.

1. Построим остаток R_i графа G относительно G_i.

2. Построим все куски графа G относительно G_i. Если ни одного такого куска построить не удается, то G_i есть плоское изображение графа G.

3. Для каждого куска выписать все грани, которые с ним совместимы. При этом возможны три случая:

а) существует кусок, не совместимый ни с одной гранью плоского графа G_i; тогда граф G на плоскость не укладывается;

б) существует кусок, совместимый с единственной гранью графа G_i; тогда выбираем этот кусок;

в) каждый из кусков совместим по крайней мере с двумя гранями графа G_i; тогда выбираем любой из таких кусков.

4. В выбранном куске P находим такую цепь μ, один или оба конца которой (и только они) принадлежат G_i. Построим граф G_{i+1}, дополнив граф G_i ребрами цепи μ, проведя μ внутри любой из совместимых с куском P граней. Плоский граф G_{i+1} построен. Переходим к пункту 1.

В случае неоднозначности проведения цепи μ будем проводить ее во внутренней грани.

Пример. Построим плоское изображение графа

$G = (\{1,2,3,4,5,6\},\{(1,2),(2,6),(5,6),(1,3),(1,4),(1,5),(2,4),(3,5),(3,6),(4,6)\})$.

Шаг 1. Выбираем в G плоский цикл $G_1 = [1,2,6,5,1]$.

Граф G_1 определяет две грани:

$F_{10} = [1,2,6,5,1]$, внешняя;

$F_{11} = [1,2,6,5,1]$, внутренняя.

Остаток R_1 графа G относительно G_1 распадается в две компоненты связности: $R_{11} = (\{3\},\varnothing)$ и $R_{12} = (\{4\},\varnothing)$.

Строим куски графа G относительно G_1 и их контактные точки. $P_{11} = (\{1,3,5,6\},\{(1,3),(3,5),(3,6)\})$; $\{1,5,6\}$;

$P_{12} = (\{1,2,4,6\},\{(1,4),(2,4),(4,6)\})$; $\{1,2,6\}$.

Кусок P_{11} совместим с гранями F_{10}, F_{11}.

Кусок P_{12} совместим с гранями F_{10}, F_{11}.

Цепь $\mu_1 = [1,4,2]$ в куске P_{12} помещаем в грани F_{11} графа G_1.

Шаг 2. Плоский граф $G_2 = (\{1,2,4,5,6\},\{(1,5),(5,6),(2,6),(1,2),(2,4),(1,4)\})$.

Граф G_2 определяет грани:

$F_{20} = [1,2,6,5,1]$; $F_{21} = [1,4,2,6,5,1]$; $F_{22} = [1,4,2,1]$.

Остаток R_2 для G относительно G_2 принимает вид: $R_2=(\{3\},\varnothing)$.

Строим куски графа G относительно G_2 и их контактные точки.

$P_{21} = (\{1,3,5,6\},\{(1,3),(3,5),(3,6)\})$; $\{1,5,6\}$;

$P_{22} = (\{4,6\},\{(4,6)\})$; $\{4,6\}$.

Кусок P_{21} совместим с гранями F_{20}, F_{21}.

Кусок P_{22} совместим с гранью F_{21}.

Цепь $\mu_2=[4,6]$ в куске P_{22} помещаем в грани F_{21} графа G_2.

Шаг 3. Плоский граф $G_3 = (\{1,2,6,5,4\},\{(1,5),(5,6),(2,6),(1,2),(2,4),(1,4),(4,6)\})$.

Граф G_3 определяет грани:

$F_{30} = [1,2,6,5,1]$; $F_{31} = [1,4,6,5,1]$;

$F_{32} = [1,4,2,1]$; $F_{33} = [4,6,2,4]$.

Остаток R_3 для G относительно G_3 принимает вид: $R_3=(\{3\},\varnothing)$.

Строим куски графа G относительно G_3 и их контактные точки.

$P_{31} = (\{1,3,5,6\},\{(1,3),(3,5),(3,6)\})$; $\{1,5,6\}$;

Кусок P_{31} совместим с гранями F_{30}, F_{31}.

Цепь $\mu_3=[1,3,5]$ в куске P_{31} помещаем в грани F_{31} графа G_3.

Шаг 4. Плоский граф $G_4 = (\{1,2,6,5,4,3\},\{(1,5),(5,6),$

$(2,6),(1,2),(2,4),(1,4),(4,6),(3,5),(1,3)\})$.

Граф G_4 определяет грани:

$F_{40} = [1,2,6,5,1]$; $F_{41} = [1,3,5,6,4,1]$; $F_{42} = [1,4,2,1]$;

$F_{43} = [4,6,2,4]$; $F_{44} = [1,3,5,1]$.

Остаток R_4 для G относительно G_4 принимает вид: $R_4 = \varnothing$.

Строим куски графа G относительно G_4 и их контактные точки.

$P_{41} = (\{3,6\},\{(3,6)\})$; $\{3,6\}$;

Кусок P_{41} совместим с гранью F_{41}.

Цепь $\mu_4=[3,6]$ в куске P_{41} помещаем в грани F_{41} графа G_4.

Шаг 5. Плоский граф $G_5 = (\{1,2,6,5,4,3\},\{(1,5),(5,6),(2,6),(1,2),(2,4),(1,4),(4,6),(3,5),(1,3),(3,6)\})$.

Ни одного куска относительно графа G_5 построить не удается. Следовательно, граф G_5 есть плоская укладка графа G. Последовательные графы G_1,G_2,G_3,G_4,G_5 приведены на рис.23.6.

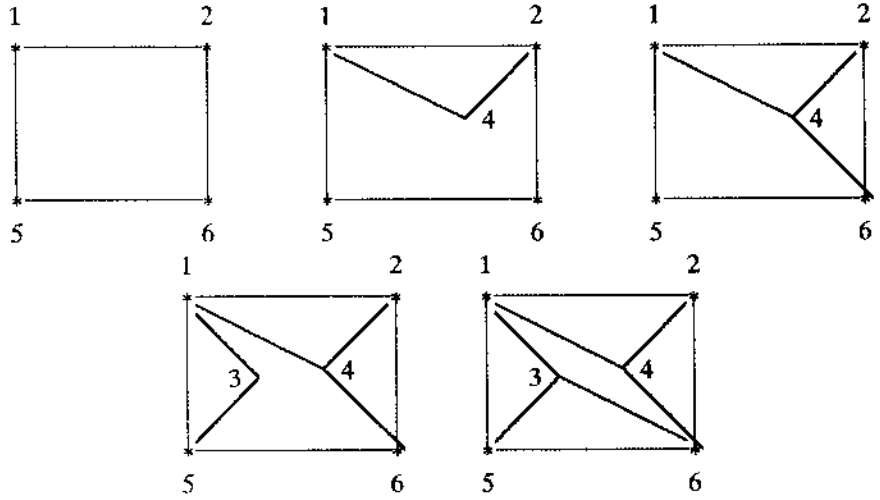


Рис.23.6

24. РАСКРАСКА ГРАФОВ

24.1. Хроматическое число и хроматический класс

Определение. Вершины (ребра) графа G *правильно раскрашены*, если каждой вершине (ребру) графа G сопоставлен некоторый цвет, причем любым двум смежным вершинам (ребрам) сопоставлены разные цвета.

Замечание. Всякий подграф правильно раскрашенного графа

правильно раскрашен.

Определение. Граф G k-раскрашиваем, если его можно правильно раскрасить не более, чем в k цветов.

Определение. Хроматическое число графа G есть наименьшее число $\chi(G)$ красок, с помощью которых можно правильно раскрасить вершины графа G. Хроматический класс графа G есть наименьшее число $\chi'(G)$ красок, с помощью которых можно правильно раскрасить ребра графа G.

24.2. Раскраска вершин

Определение. Граф G с хроматическим числом $\chi(G) = 2$ называется *бихроматическим*.

Теорема. Граф G двудолен тогда и только тогда, когда G есть бихроматический граф.

Доказательство. Пусть $G = (V_1,V_2,E)$ есть двудольный граф. Вершины из V_1 окрашиваем в один цвет, а вершины из V_2 – в другой. Полученная раскраска правильная, ибо соседние вершины, одна из V_1, а другая из V_2 окрашены в разный цвет.

Пусть теперь G есть бихроматический граф. Тогда множество его вершин можно разбить на два класса:

V_1, вершины из G, окрашенные в один цвет;

V_2, вершины из G, окрашенные в другой цвет.

Ребра из G могут соединять вершины только из разных классов. Следовательно, $G = (V_1,V_2,E)$ есть двудольный граф.

Замечание. Следующие утверждения эквивалентны.

1. Граф G является бихроматическим.
2. Граф G двудолен.
3. Все простые циклы в G имеют четную длину.

Дерево есть бихроматический граф, ибо вершины четных ярусов графа можно окрасить в один цвет, а вершины нечетных ярусов в другой.

Утверждение. Пусть K_p есть полный граф с p вершинами. Тогда хроматическое число $\chi(K_p) = p$.

Доказательство. Индукция по p.

Базис. $p = 3$. $\chi(K_3) = 3$.

Предположение индукции. Предположим, что $\chi(K_{p-1}) = p-1$.

Шаг индукции. Покажем, что $\chi(K_p) = p$. В самом деле, пусть v есть некоторая вершина в K_p. Удалим вершину v из K_p вместе с инцидентными ей ребрами. Тогда граф $K_p-v = K_{p-1}$ $p-1$-раскрашиваем по предположению индукции $p-1$ красками $1,2,\ldots,$ $p-1$. Вершина v в K_p должна быть окрашена в новый цвет p. Поэтому $\chi(K_p) = p$.

Следствие. Существуют графы со сколь угодно большим хроматическим числом.

24.3. Верхняя и нижняя оценки хроматического числа. Внутренне и внешне устойчивые множества вершин графа

Теорема (верхняя оценка). Если граф G имеет максимальную степень вершин, равную s, то хроматическое число $\chi(G) \leqslant s+1$.

Доказательство. Индукция по числу p вершин.

Базис. Если число вершин в графе G не превосходит $s+1$, то утверждение теоремы тривиально, ибо $s+1$ вершин можно правильно раскрасить в $s+1$ цветов, по одному цвету на каждую вершину.

Предположение индукции. Допустим, что теорема верна для графов с числом вершин $k < p$, причем $p \geqslant s+1$.

Шаг индукции. Покажем, что если граф G имеет p вершин, то хроматическое число $\chi(G) \leqslant s+1$. Пусть v есть произвольная вершина в G. Удалим из G вершину v вместе с инцидентными ей ребрами. Получившийся граф $G' = G-v$ имеет число вершин $k < p$ и потому по предположению индукции граф G' $s+1$-раскрашиваем. В графе G вершина v имеет не более s соседних вершин, окрашенных не более, чем в s цветов. Вершине v припишем один из оставшихся цветов. Тогда граф G правильно раскрашиваем с числом цветов $r \leqslant s+1$, т.е. граф G имеет хроматическое число $\chi(G) \leqslant s+1$.

24.3.1. Внутренне устойчивые множества вершин графа

Определение. Подмножество S вершин графа $G = (V,E)$ *внутренне устойчиво*, если никакие две вершины из S не смежны в G. *Число внутренней устойчивости* графа G

$$\alpha(G) = \max \{|S| : S \subseteq V \text{ и } S \text{ внутренне устойчиво в } G\}.$$

Внутренне устойчивое множество вершин S называется (*максимально*) *тупиковым*, если всякое строгое надмножество множества S внутренне устойчивым уже не является. При этом S называется наибольшим, если среди всех внутренне устойчивых множеств вершин в G оно имеет наибольшую мощность.

Пусть S есть внутренне устойчивое множество вершин графа $G = (V,E)$ и ребро $e=(u,v) \in E$. С каждой вершиной $v \in V$ свяжем логическую переменную x_v и пусть x_v означает, что $v \notin S$.

Построим логическую формулу

$$\&_{(u,v)\in E}\ (x_u \vee x_v). \qquad (24.1)$$

Взяв по одной переменной в каждой скобке формулы (24.1), получим некоторую конъюнкцию $x_{u_1}x_{u_2}\ldots x_{u_q}$, для которой $V - \{u_1,u_2,\ldots,u_q\}$ является внутренне устойчивым множеством вершин. Формула (24.1) есть условие внутренней устойчивости вершин в G. Пусть $\vee_i K_i$ есть минимальная ДНФ D для формулы (24.1). Каждому дизъюнктивному слагаемому $K_i = x_{w_1}x_{w_2}\ldots x_{w_m}$ в D соответствует максимальное внутренне устойчивое множество $S_i = V - \{w_1,w_2,\ldots,w_m\}$. Множества S_i исчерпывают все максимальные внутренне устойчивые множества вершин графа G. Из них можно выбрать все наибольшие.

24.3.2. Алгоритм вычисления всех наибольших внутренне устойчивых множеств вершин графа $G = (V,E)$

1. Построить формулу

$$F = \&_{(u,v)\in E}\ (x_u \vee x_v),$$

условие внутренней устойчивости графа G.

2. Построить минимальную ДНФ D формулы F.

3. Для каждого дизъюнктивного слагаемого $K = x_u x_v \ldots x_w$ в D получить соответствующее ему максимальное внутренне устойчивое множество вершин $S = V - \{u,v,\ldots,w\}$.

4. Из полученных максимальных внутренне устойчивых множеств вершин выбрать все наибольшие.

Пример. Построим все наибольшие внутренне устойчивые множества вершин для графа

$G = (V,E) = (\{1,2,3,4,5,6,7\},\{(1,2),(1,3),(1,5),$
$(1,6),(2,3),\ (3,4),(3,6),(4,5),(4,7),(5,6),(6,7)\})$.

Условие внутренней устойчивости графа G

$$F = \&_{(u,v)\in E}\ (u \vee v) = (1\vee2)(1\vee3)(1\vee5)(1\vee6)(2\vee3)(3\vee4)\&$$

$$(3\vee6)(4\vee5)(4\vee7)(5\vee6)(6\vee7) = 1357\vee23456\vee23567\vee1246\vee1346.$$

Максимальными внутренне устойчивыми множествами вершин будут множества:

$V - \{1,3,5,7\} = \{2,4,6\}$; $V - \{2,3,4,5,6\} = \{1,7\}$;
$V - \{2,3,5,6,7\} = \{1,4\}$; $V - \{1,2,4,6\} = \{3,5,7\}$;

$V - \{1,3,4,6\} = \{2,5,7\}$.

Выбираем из них наибольшие: $\{2,4,6\}$; $\{3,5,7\}$; $\{2,5,7\}$.

Теорема (нижняя оценка). Пусть $G = (V,E)$ есть связный (p, q)-граф. Пусть $\alpha(G)$ есть число внутренней устойчивости графа G. Тогда хроматическое число $\chi(G) \geqslant p/\alpha(G)$.

Доказательство. Граф G $\chi(G)$-раскрашиваем. Пусть

$V_1,V_2,\ldots,V_{\chi(G)}$ есть множества вершин, окрашенных в цвета $1, 2,\ldots, \chi(G)$ соответственно.

Вершины из V_i внутренне устойчивы, ибо они окрашены в один и тот же цвет и потому не смежны. Поэтому $|V_i| \leqslant \alpha(G)$, $i = 1,2,\ldots,\chi(G)$. Множества $V_1,V_2,\ldots,V_{\chi(G)}$ попарно не пересекаются и в сумме дают все множество V. Тогда

$$p = \sum_{i=1}^{\chi(G)} |V_i| \leqslant \sum_{i=1}^{\chi(G)} \alpha(G) = \alpha(G)\cdot\chi(G);$$

отсюда имеем, что хроматическое число $\chi(G) \geqslant p/\alpha(G)$.

Замечание. Из теорем о верхней и нижней оценках для хроматического числа $\chi(G)$ графа G имеем

$$p/\alpha(G) \leqslant \chi(G) \leqslant s + 1.$$

24.3.3. Внешне устойчивые множества вершин графа

Определение. Множество T вершин графа $G = (V,E)$ называется *внешне устойчивым* (в G), если $\forall v \notin T\ \exists u \in T\ (e=(u,v) \in E)$. *Число внешней устойчивости* графа G

$$\beta(G) =$$
$$\min\{|T| : T \subseteq V \text{ и } T \text{ есть внешне устойчивое множество в } G\}.$$

Внешне устойчивое множество вершин T называется (*минимально*) *тупиковым*, если T не содержит в себе строго ни одного подмножества, являющегося внешне устойчивым. Внешне устойчивое множество вершин называется *наименьшим*, если среди всех внешне устойчивых множеств вершин в G оно имеет наименьшую мощность.

С понятием внешне устойчивого множества можно связать следующую практическую задачу. Пусть имеем карту городов с дорогами между ними. Следует построить наименьшее число складов с не более чем одним складом в каждом городе так, чтобы от каждого города вела прямая дорога к одному из скла-

дов. Задача решается отысканием наименьшего внешне устойчивого множества вершин графа городов с дорогами между ними.

Пусть T есть внешне устойчивое множество вершин графа $G = (V,E)$. С каждой вершиной $u \in V$ свяжем логическую переменную x_u и пусть x_u означает, что $u \in T$. Пусть в множестве $\{u,v,\ldots,w\}$ u есть вершина из V, а $v,\ldots,w$ есть все те вершины, которые смежны с u. Построим логическую формулу

$$\&_{u \in V} (x_u \vee x_v \vee \ldots x_w). \qquad (24.2)$$

Взяв по одной переменной в каждой скобке формулы (24.2), получим некоторую конъюнкцию $x_{u_1}x_{u_2}\ldots x_{u_p}$, для которой $\{u_1, u_2,\ldots,u_p\}$ есть внешне устойчивое множество вершин. Формула (24.2) является условием внешней устойчивости вершин в G. Пусть $\vee_i K_i$ есть минимальная ДНФ D для формулы (24.2). Каждому дизъюнктивному слагаемому $K_i = x_{w_1}x_{w_2}\ldots x_{w_m}$ в D соответствует минимальное внешне устойчивое множество $T_i = \{w_1, w_2,\ldots,w_m\}$. Множества T_i исчерпывают все минимальные внешне устойчивые множества вершин графа G. Из них можно выбрать все наименьшие.

24.3.4. Алгоритм вычисления всех наименьших внешне устойчивых множеств вершин графа $G = (V,E)$

1. Построить формулу

$$F = \&_{u \in V} (x_u \vee (\vee_{(u,v) \in E} x_v)),$$

условие внешней устойчивости графа G.

2. Построить минимальную ДНФ D формулы F.

3. Для каждого дизъюнктивного слагаемого $K = x_u x_v \ldots x_w$ в D получить соответствующее ему минимальное внешне устойчивое множество вершин $S = \{u,v,\ldots,w\}$.

4. Из полученных минимальных внешне устойчивых множеств вершин выбрать все наименьшие.

Пример. Вычислим все наименьшие внешне устойчивые множества вершин графа

$G = (V,E) = (\{1,2,3,4,5,6,7\},\{(1,2),(1,3),(1,5),(1,6),$
$(2,3),(3,4),(3,6),(4,5),(4,7),(5,6),(6,7)\})$.

Условие внешней устойчивости для графа G

$F = \&_{u \in V} (u \vee (\vee_{(u,v) \in E} v)) = (1V2V3V5V6)(2V1V3)(3V1V2V4V6)\&$
$(4V3V5V7)(5V1V4V6)(6V1V3V5V7)(7V4V6) = 156 \vee 17 \vee 246 \vee$
$247 \vee 257 \vee 245 \vee 256 \vee 267 \vee 357 \vee 36 \vee 34 \vee 14.$

Все минимальные внешне устойчивые множества:

$\{1,5,6\}$, $\{1,7\}$, $\{2,4,6\}$, $\{2,4,7\}$, $\{2,5,7\}$, $\{2,4,5\}$,
$\{2,5,6\}$, $\{2,6,7\}$, $\{3,5,7\}$ $\{3,6\}$, $\{3,4\}$, $\{1,4\}$.

Из полученных множеств выбираем наименьшие по мощности. Они и составят все наименьшие внешне устойчивые множества вершин: $\{1,7\}$; $\{3,6\}$; $\{3,4\}$; $\{1,4\}$.

24.4. Оптимальная раскраска вершин графа

Пусть граф $G = (V,E)$ правильно раскрашен. Вершины, окрашенные в один цвет, образуют внутренне устойчивое множество вершин в G. Хроматическое число графа G можно определить как минимальное число внутренне устойчивых множеств, в сумме дающих все множество V. Такое минимальное покрытие можно найти следующим образом. Пусть $S_1,S_2,\ldots,S_r$ есть все максимальные внутренне устойчивые множества вершин в G. С каждым S_i свяжем логическую переменную x_{S_i} и пусть x_{S_i} означает, что вершина $v \in S_i$. Построим логическую формулу – условие оптимальной раскраски вершин графа G

$$\&_{v \in V} (\vee_{v \in S_i, i=1,\ldots,r} x_{S_i}). \qquad (24.3)$$

Взяв по одной переменной в каждой скобке формулы (24.3), получим некоторую конъюнкцию $x_{S_a}x_{S_b}\ldots x_{S_c}$, для которой семейство внутренне устойчивых множеств $\{S_a,S_b,\ldots,S_c\}$ в сумме покрывает все множество V. Пусть $\vee_i K_i$ есть минимальная ДНФ D для формулы (24.3). Пусть дизъюнктивному слагаемому $K_i = x_{S_{j_1}}x_{S_{j_2}}\ldots x_{S_{j_k}}$ в D соответствует наименьшее по длине k семейство $L_i = \{S_{j_1},S_{j_2},\ldots, S_{j_k}\}$. Хроматическое число $\chi(G) = k$. Ему соответствует следующая оптимальная раскраска вершин графа G. В цвета $1,2,\ldots,k$ последовательно окрашиваем семейства вершин

$$S_{j_1},\ S_{j_2}-S_{j_1},\ S_{j_3}-(S_{j_1} \cup S_{j_2}),\ldots,S_{j_k}-(S_{j_1} \cup \ldots \cup S_{j_{k-1}})$$

соответственно.

24.4.1. Алгоритм оптимальной раскраски (p,q)-графа $G = (V,E)$

1. Построить все максимальные внутренне устойчивые множества вершин $S_1, S_2, \ldots, S_r$.

2. Построить логическую формулу F – условие оптимальной раскраски графа G:

$$\&_{v \in V} \left(\vee_{v \in S_i, i=1,\ldots,r} x_{S_i} \right).$$

3. Построить минимальную ДНФ D для F.

4. Каждому дизъюнктивному слагаемому $K_i = x_{S_a} x_{S_b} \ldots x_{S_c}$ в D соответствует минимальное семейство $L_i = \{S_a, S_b, \ldots, S_c\}$ внутренне устойчивых множеств $S_a, S_b, \ldots, S_c$. Из всех L_i выбираем наименьшее по длине k семейство $\{S_{j_1}, S_{j_2}, \ldots, S_{j_k}\}$. Хроматическое число $\chi(G) = k$. Ему соответствует следующая оптимальная раскраска вершин графа G. В цвета $1,2,\ldots,k$ последовательно окрашиваем семейства вершин S_{j_1}, $S_{j_2} - S_{j_1}$, $S_{j_3} - (S_{j_1} \cup S_{j_2})$, ..., $S_{j_k} - (S_{j_1} \cup \ldots \cup S_{j_{k-1}})$ соответственно.

Пример. Построим оптимальные раскраски графа

$G=(V,E)=(\{1,2,3,4,5,6,7,8\},\{(1,2),(1,3),(1,5),(1,6),$
$(2,3),(3,4),(3,6),(4,5),(4,7),(5,6),(6,7),(7,8)\})$.

Составим по условию внутренней устойчивости вершин графа G решеточное выражение

$$F = \&_{(u,v) \in E} (u \vee v) = (1 \vee 2)(1 \vee 3)(1 \vee 5)(1 \vee 6)(2 \vee 3)(3 \vee 4) \&$$
$$(3 \vee 6)(4 \vee 5)(4 \vee 7)(5 \vee 6)(6 \vee 7)(7 \vee 8) =$$

$23567 \vee 12467 \vee 12468 \vee 13467 \vee 13468 \vee 1357 \vee 234568.$

Рассматривая полученные дизъюнктивные слагаемые как множества и дополняя их до множества вершин V, получим, что множество

$S = \{S_1, S_2, S_3, S_4, S_5, S_6, S_7\} = \{\{1,4,8\},\{3,5,8\},\{3,5,7\},$
$\{2,5,8\},\{2,5,7\},\{2,4,6,8\},\{1,7\}\}$

есть список всех максимальных (тупиковых) внутренне устойчивых множеств вершин графа G. Составляем решеточное выражение – условие оптимальной раскраски вершин графа

$$R = \&_{v \in V} \left(\vee_{v \in S_i} i \right) = (1 \vee 7)(4 \vee 5 \vee 6)(2 \vee 3)(1 \vee 6)(2 \vee 3 \vee 4 \vee 5)6 \&$$

$$(3 \vee 5 \vee 7)(1 \vee 2 \vee 4 \vee 6) = 672 \vee 736 \vee 2651 \vee 631.$$

Из полученных дизъюнктивных слагаемых выбираем наименьшие по длине: 672, 736, 631. Построим оптимальные раскраски вершин графа по множествам $\{S_6, S_7, S_2\}$, $\{S_7, S_3, S_6\}$, $\{S_6, S_3, S_1\}$. Хроматическое число $\chi(G)=3$, т.е. для правильной раскраски вершин графа необходимо три краски. Возможны следующие варианты оптимальной раскраски вершин.

1. Вершины $L_1 = S_6 = \{2,4,6,8\}$ покрасим цветом 1;
вершины $L_2 = S_7 - S_6 = \{1,7\}$ – цветом 2;
вершины $L_3 = S_2 - (S_6 \cup S_7) = \{3,5\}$ – цветом 3.

2. $L_1 = S_7 = \{1,7\}$; $L_2 = S_3 - S_7 = \{3,5\}$;
$L_3 = S_6 - (S_7 \cup S_3) = \{2,4,6,8\}$.

3. $L_1 = S_6 = \{2,4,6,8\}$; $L_2 = S_3 - S_6 = \{3,5,7\}$; $L_3 = S_1 - (S_6 \cup S_3) = \{1\}$.

24.5. Раскрашивание планарных графов

Теорема (о пяти красках). Всякий связный планарный граф G раскрашиваем не более чем пятью красками.

Доказательство. Индукция по числу вершин p графа.

Базис. Для графа с числом вершин $k \leqslant 5$ теорема очевидна, ибо всякие пять вершин 5-раскрашиваемы пятью различными красками.

Предположение индукции. Допустим, что всякий связный планарный граф с числом вершин $k<p$ 5-раскрашиваем.

Шаг индукции. Покажем, что всякий связный планарный граф с p вершинами 5-раскрашиваем. Так как граф G связен и планарен, то он имеет вершину v степени $s \leqslant 5$. Удалим из G вершину v вместе с инцидентными ей ребрами. Полученный планарный граф $G' = G - v$ имеет число вершин $k<p$, и потому по предположению индукции все компоненты связности графа G' можно раскрасить не более чем пятью красками. Возможны следующие случаи.

1. Степень вершины v не более 4. Пусть для определенности $deg(v) = 4$. Смежные с v вершины v_1, v_2, v_3, v_4 получат в G' не более четырех красок. Вершину v в графе G окрасим в любую из оставшихся красок.

2. $deg(v)=5$. Если смежные с v вершины v_1, v_2, v_3, v_4, v_5 имеют в совокупности $r \leqslant 4$ красок, то вершину v окращиваем в оставшийся цвет. Пусть теперь вершины v_1, v_2, v_3, v_4, v_5 окрашены в пять цветов c_1, c_2, c_3, c_4, c_5 соответственно. Натянем на

вершины v_1,v_2,v_3,v_4,v_5 из G подграф H, соединив вершины v_1,v_2,v_3,v_4,v_5 в H ребрами так, как они соединены в G. Подграф H планарный, следовательно, H не содержит K_5. Поэтому в H среди вершин v_1,v_2,v_3,v_4,v_5 существуют две несмежные вершины (ибо если все вершины v_1,v_2,v_3,v_4,v_5 смежны, то граф H есть K_5, чего нет). Пусть для определенности вершины v_1,v_2 не смежны (рис.24.1). Склеим v_1,v_2 с вершиной v. Полученный связный планарный граф по предположению индукции 5-раскрашиваем. При этом четыре вершины v,v_3,v_4,v_5 получат $r \leqslant 4$ цвета. Расклеим назад вершины v,v_1,v_2. Вершинам v_1,v_2 оставим их цвет. Тогда вершины v_1,v_2,v_3,v_4,v_5 имеют $r \leqslant 4$ цвета. Вершину v перекрасим в один из оставшихся цветов.

Шаг индукции установлен. Теорема доказана.

Замечание. Широко известна проблема четырех красок: любой плоский граф 4-раскрашиваем. Появилось много ошибочных доказательств этой гипотезы. Последнее сообщение о положительном решении проблемы четырех красок опубликовано в 1977 г.:

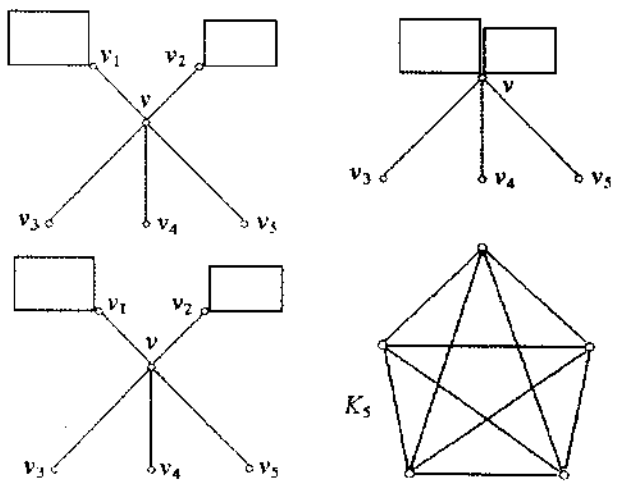


Рис.24.1

25. ПОТОКИ В ТРАНСПОРТНЫХ СЕТЯХ

15.1. Двухполюсные сети

Определение. (Двухполюсная) *сеть* $S = (V,E,s,t)$ есть орграф $G = (V,E)$ с двумя выделенными вершинами (полюсами): s есть входная вершина сети (*исток*); t есть выходная вершина сети (*сток*).

Определение. *Внутренние* вершины сети есть вершины сети, отличные от полюсов. *Полюсная* дуга сети инцидентна одному из полюсов.

Пусть $S = (V,E,s,t)$ есть сеть и $v \in V$. Введем следующие обозначения (рис.25.1):

$D^+(v)$, множество дуг сети, исходящих из v;

$D^-(v)$, множество дуг сети, входящих в v;

$D(v)$, множество дуг сети, инцидентных вершине v.

Очевидно, что $D(v) = D^+(v) \cup D^-(v)$.

Пусть $A \subseteq V$ есть некоторое множество вершин сети S. Положим $D(A) = \bigcup\limits_{v\in A} D(v)$; $D^+(A) = \bigcup\limits_{v\in A} D^+(v)$; $D^-(A) = \bigcup\limits_{v\in A} D^-(v)$.

Определение.

$e = (u,v)$ есть *граничная дуга* для $A \subseteq V \longleftrightarrow u\in A \;\&\; v\notin A \;\vee\; u\notin A \;\&\; v\in A$.

$e = (u,v)$ есть *внутренняя дуга* для $A \subseteq V \longleftrightarrow u\in A \;\&\; v\in A$.

Пусть

ВН($D(A)$) есть множество внутренних дуг из $D(A)$;
ГР($D(A)$) есть множество граничных дуг из $D(A)$;
ВН($D^-(A)$) есть множество внутренних дуг из $D^-(A)$;
ГР($D^-(A)$) есть множество граничных дуг из $D^-(A)$;
ВН($D^+(A)$) есть множество внутренних дуг из $D^+(A)$;
ГР($D^+(A)$) есть множество граничных дуг из $D^+(A)$.

Верны следующие равенства:

$D(A)$ = ВН($D(A)$) $\cup$ ГР($D(A)$);
$D^-(A)$ = ВН($D^-(A)$) $\cup$ ГР($D^-(A)$);
$D^+(A)$ = ВН($D^+(A)$) $\cup$ ГР($D^+(A)$).

Заметим, что ВН($D^-(A)$) = ВН($D^+(A)$) = ВН($D(A)$), ибо для дуги $e = (u,v)$ имеем соотношение

$$e \in \begin{cases} \text{ВН}(D^+(A)) \text{ по вершине } u; \\ \text{ВН}(D^-(A)) \text{ по вершине } v; \\ \text{ВН}(D(A)) \text{ по вершинам } u,v, \end{cases}$$

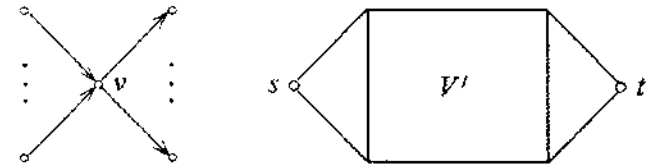


Рис.25.1

где три множества составлены из одних и тех же дуг.

25.2. Дивергенция

Пусть $S = (V,E,s,t)$ есть сеть. Зададим функцию $f: E \to \mathbb{R}_+$, где $\mathbb{R}_+$ есть множество неотрицательных вещественных чисел.

Пусть $V' = V - \{s,t\}$ есть множество внутренних вершин сети S.

Определение. *Дивергенция* функции f в (внутренней) вершине $v \in V'$ есть величина (число) $div_f(v) = \sum\limits_{e \in D^+(v)} f(e) - \sum\limits_{e \in D^-(v)} f(e)$.

Дивергенция функции f на множестве (внутренних) вершин $A \subseteq V'$ есть величина (число) $div_f(A) = \sum\limits_{v \in A} div_f(v)$. В истоке и стоке

$$div_f(s) = \sum_{e \in D^+(s)} f(e); \qquad div_f(t) = \sum_{e \in D^-(t)} f(e).$$

Утверждение (основное свойство дивергенции). Для $A \subseteq V'$

$$div_f(A) = \sum_{e \in D^+(A)} f(e) - \sum_{e \in D^-(A)} f(e).$$

Доказательство. $div_f(A) = \sum\limits_{v \in A} div_f(v) =$

$$\sum_{v \in A} \Big(\sum_{e \in D^+(v)} f(e) - \sum_{e \in D^-(v)} f(e) \Big) = \sum_{v \in A} \sum_{e \in D^+(v)} f(e) - \sum_{v \in A} \sum_{e \in D^-(v)} f(e) =$$

$$\sum_{e \in D^+(A)} f(e) - \sum_{e \in D^-(A)} f(e).$$

25.3. Потоки в сетях

Определение. *Транспортная сеть* $S = (V,E,s,t,c)$ есть сеть (V,E,s,t), в которой $D^-(s) = \varnothing$, $D^+(t) = \varnothing$ и для которой определена функция $c: E \to \mathbb{R}_+$ – *пропускная способность* дуг. Вершина s есть исток сети, вершина t есть сток сети.

Слово транспортная иногда будем опускать.

Определение. *Поток* в сети $S = (V,E,s,t,c)$ есть функция $f: E \to \mathbb{R}_+$, удовлетворяющая следующим условиям:

1) $0 \leqslant f(e) \leqslant c(e)$;

2) для всякой внутренней вершины v сети S $div_f(v) = 0$.

Пусть $S = (V,E,s,t,c,f)$ есть транспортная сеть с пропускной способностью дуг $c: E \to \mathbb{R}_+$ и потоком $f: E \to \mathbb{R}_+$.

Свойства дивергенции потока

Пусть $V' = V - \{s,t\}$ есть множество внутренних вершин сети S.

1. $div_f(V') = \sum\limits_{v \in V'} div_f(v) = 0$. Если $A \subseteq V'$, то $div_f(A) = 0$.

2. $div_f(s) = div_f(t)$. В самом деле, ввиду $D^+(V') = \text{ВН}(D^+(V')) \cup \Gamma P(D^+(V'))$ (рис.25.1) получаем $0 = div_f(V') =$

$$\sum_{e \in D^+(V')} f(e) - \sum_{e \in D^-(V')} f(e) =$$

$$\underbrace{\sum_{e \in \Gamma P(D^+(V'))} f(e)}_{e \in D^-(t)} + \underbrace{\sum_{e \in \text{ВН}(D^+(V'))} f(e) - \sum_{e \in \text{ВН}(D^-(V'))} f(e)}_{\text{сумма} = 0} - \underbrace{\sum_{e \in \Gamma P(D^-(V'))} f(e)}_{e \in D^+(s)}$$

$= div_f(t) - div_f(s)$, откуда $div_f(s) = div_f(t)$.

Определение. *Величина потока* f в сети $S = (V,E,s,t,c,f)$ есть величина (число)

$$M_f = div_f(s) = \sum_{e \in D^+(s)} f(e) \; (= div_f(t) = \sum_{e \in D^-(t)} f(e)).$$

25.4. Сечения в сетях

Определение. *Сечение* (или *разрез*) в сети $S = (V,E,s,t)$ есть множество $W \subseteq E$ дуг, при удалении которых получившаяся сеть $S - W$ становится несвязной (рис.25.2), причем полюсы s и t попадают в разные компоненты связности. Сечение W *простое*, если при возвращении в сеть $S - W$ любой дуги e из W сеть $(S - W) + e$ становится связной (по этой дуге).

Сечение W в сети S можно построить, разделив множество вершин в S на два подмножества A и B, причем $s \in A$, $t \in B$, и взяв в качестве W все дуги с началом в A и с концом в B, а также все дуги с началом в B и с концом в A.

При неориентированной связности далее будем говорить о связности орграфа по ребрам, не обращая внимания на их направление.

Определение. Дуга e простого сечения называется *прямой*, если в цепи $[s,t]$, проходящей через дугу e, она ориентирована от s к t; и *обратной*, если дуга e ориентирована в цепи $[s,t]$ от t к s.

На рис.25.2 в цепи $[s,t]$ дуга e' прямая, а e'' обратная.

Направленность дуги зависит от выбора простого сечения. В фиксированном простом сечении W ориентация дуги не зависит от выбора цепи: дуга либо прямая, либо обратная. Для каждой дуги e из простого сечения W можно указать цепь $[s,t]$, которая проходит через дугу e и не проходит через остальные дуги сечения W.

Определение. Пусть W есть простое сечение в сети S. Пусть
$W_{\text{пр}}$ есть множество прямых дуг в сечении W;
$W_{\text{об}}$ есть множество обратных дуг в сечении W.
Тогда *пропускная способность простого* сечения W есть число

$$c(W) = \sum_{e \in W_{\text{пр}}} c(e).$$

Простое сечение W *минимально*, если W имеет минимальную пропускную способность. *Пропускная способность сети* есть пропускная способность минимального простого сечения этой сети.

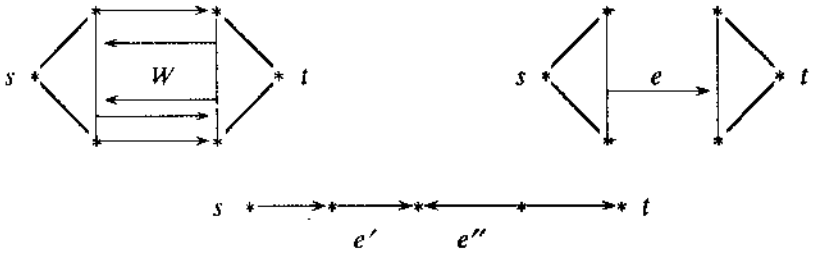


Рис.25.2

25.5. Величина потока и пропускная способность сети

Пусть $S = (V,E,s,t,c,f)$ есть транспортная сеть, где $c: E \to \mathbb{N}$, $f: E \to \mathbb{N}$ есть целочисленные функции пропускных способностей дуг и потока в сети S ($\mathbb{N} = \{0,1,2,\ldots\}$). Пусть W есть простое сечение сети S; $W_{\text{пр}}$, $W_{\text{об}}$ есть множества прямых и обратных дуг в сечении W соответственно.

Утверждение 1. Величина потока $M_f = \sum_{e \in W_{\text{пр}}} f(e) - \sum_{e \in W_{\text{об}}} f(e)$.

Доказательство. Пусть K_s есть компонента связности в сети $S - W$, содержащая исток s. Пусть V_s есть множество вершин компоненты K_s (рис.25.3). Тогда величина потока

$$M_f = div_f(s) =$$

(прибавляем равную нулю дивергенцию по внутренним вершинам сети из $V_s - \{s\}$, ибо $div_f(V_s - \{s\}) = 0$ по основному свойству дивергенции) $div_f(V_s)$ = (по основному свойству дивергенции)

$$\sum_{e \in D^+(V_s)} f(e) - \sum_{e \in D^-(V_s)} f(e) = \text{(так как}$$

$D^+(V_s) = \text{ВН}(D^+(V_s)) \cup \text{ГР}(D^+(V_s)) = \text{ВН}(D^+(V_s)) \cup W_{\text{пр}}$,
$D^-(V_s) = \text{ВН}(D^-(V_s)) \cup \text{ГР}(D^-(V_s)) = \text{ВН}(D^-(V_s)) \cup W_{\text{об}}$,
то продолжая выше написанное равенство, имеем)

$$\sum_{e \in W_{\text{пр}}} f(e) + \underbrace{\sum_{e \in \text{ВН}(D^+(V_s))} f(e) - \sum_{e \in \text{ВН}(D^-(V_s))} f(e)}_{\text{сумма} = 0} - \sum_{e \in W_{\text{об}}} f(e) =$$

$$\sum_{e \in W_{\text{пр}}} f(e) - \sum_{e \in W_{\text{об}}} f(e).$$

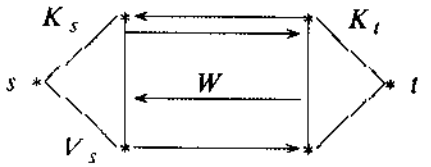


Рис.25.3

Утверждение 2. Если c_{min} есть пропускная способность сети (т.е. пропускная способность $c(W_{min})$ минимального простого сечения W_{min}), то величина потока $M_f \leqslant c_{min}$.

Доказательство. $M_f = \sum\limits_{e\in W_{\text{пр}}} f(e) - \sum\limits_{e\in W_{\text{об}}} f(e) \leqslant \sum\limits_{e\in W_{\text{пр}}} f(e) \leqslant \sum\limits_{e\in W_{\text{пр}}} c(e) = c(W_{min}) = c_{min}$.

Замечание. Если W_{min} есть минимальное сечение в сети S с пропускной способностью $c(W_{min}) = c_{min}$ и если поток f в сети S имеет максимально возможную величину $M_f = c_{min} = c(W_{min}) = \sum\limits_{e\in W_{\text{пр}}} c(e)$, то при этом $M_f = \sum\limits_{e\in W_{\text{пр}}} f(e) - \sum\limits_{e\in W_{\text{об}}} f(e) = \sum\limits_{e\in W_{\text{пр}}} c(e)$ возможно лишь тогда, когда на дугах e из $W_{\text{пр}}$ $f(e) = c(e)$, а на дугах e из $W_{\text{об}}$ поток $f(e) = 0$; так что максимально возможный поток M_f, равный c_{min}, нагружает в минимальном сечении прямые дуги до их пропускных способностей, а обратные дуги не нагружает совсем (нагружает нулями).

25.6. Максимальный поток

Пусть $S = (V,E,s,t,c,f)$ есть транспортная сеть с пропускной способностью $c: E \to \mathbb{N}$ и потоком $f: E \to \mathbb{N}$; $c_{\min}$ есть максимально возможная величина потока M_f.

Определение. Дуга $e\in E$ в сети S *насыщена*, если $f(e)=c(e)$.

Теорема 1. Пусть $\mu = [s,t]$ есть путь от s до t. Если все дуги этого пути не насыщены, то поток f можно увеличить так, что одна из дуг пути μ окажется насыщенной.

Доказательство. Пусть $\delta = \min\limits_{e\in\mu} (c(e)-f(e))$. Увеличивая на δ поток f в каждой дуге e из μ, придем к потоку $f'=f+\delta$. Дуга e, на которой $c(e)-f(e) = \delta$, оказывается насыщенной.

Пример. $c(e_1) = 5$, $c(e_2) = 3$, $c(e_3) = 6$, $f(e_1) = 2$, $f(e_2) = 2$, $f(e_3) = 4$ (рис.25.4). Дуга e_2 оказалась насыщенной.

Определение. Поток f в сети S полный, если всякий путь от s до t имеет насыщенную дугу.

дуга e_2 насыщена

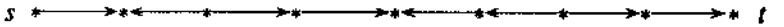
e_1	e_2	e_3	
5	3	6	$c(e)$
2	2	4	$f(e)$
3	1	2	$\delta = \min(c(e)-f(e)) = 1$
3	3	5	$f'(e) = f(e) + \delta$

Рис.25.4

Теорема 2. Пусть $\nu = [s,t]$ есть цепь между s и t в сети S с полным потоком f. Пусть $\vec{e}$ есть прямая дуга в цепи ν, а дуга $\overleftarrow{e}$ обратная. Для дуги $e\in\nu$ положим $\delta(\vec{e}) = c(\vec{e}) - f(\vec{e})$; $\delta = \min(\delta(\vec{e}))$ по всем $\vec{e} \in \nu$; $\eta = \min(f(\overleftarrow{e}))$ по всем $\overleftarrow{e} \in \nu$; $\varepsilon = \min(\delta,\eta)$. Если $\varepsilon > 0$, то увеличивая на ε поток на каждой прямой дуге $\vec{e} \in \nu$ и уменьшая на ε поток на каждой обратной дуге $\overleftarrow{e} \in \nu$, получим новый поток $f' = f + \varepsilon$.

Доказательство следует из формулировки теоремы.

Пример. На рис.25.5 приведена цепь в сети S между s и t, причем цепь имеет и прямые, и обратные дуги. Величины $\delta = 2$, $\eta = 3$, $\varepsilon = \min(\delta,\eta) = \min(2,3) = 2$. Дуги e_1,e_7,e_8 оказались насыщенными.



$\vec{e}_1$	$\overleftarrow{e}_2$	$\vec{e}_3$	$\vec{e}_4$	$\overleftarrow{e}_5$	$\overleftarrow{e}_6$	$\vec{e}_7$	$\vec{e}_8$	
4	5	11	8	9	4	7	6	$c(e)$
2	3	3	4	9	3	5	4	$f(e)$
2	–	8	4	–	–	2	2	$\delta(\vec{e})$
–	3	–	–	9	3	–	–	$f(\overleftarrow{e})$
4	1	5	6	7	1	7	6	$f'(e)$

Рис.25.5

Теорема 3. Если в сети S не существует цепи $\nu = [s,t]$ между s и t с $\varepsilon > 0$, то поток f максимален и его величина $M_f = c_{min}$.

Доказательство. Исключим из сети S все дуги $e \in E$, для которых $f(\vec{e}) = c(\vec{e})$ или $f(\overleftarrow{e}) = 0$. В результате получим сеть S', в которой полюсы s и t не связны, ибо в случае их связности в S' существовала бы цепь ν между s и t, в которой $f(\vec{e}) < c(\vec{e})\ \forall \vec{e} \in \nu$ и $f(\overleftarrow{e}) > 0\ \forall \overleftarrow{e} \in \nu$. Но тогда в S для цепи ν число $\varepsilon > 0$. Противоречие с условием.

Итак, полюсы s и t в S' не связны. Пусть K_s и K_t есть компоненты связности в S', содержащие полюсы s и t соответственно, V_s – множество вершин в компоненте K_s. Можно показать, что множество дуг $\Gamma P(D^+(V_s))$ есть множество прямых дуг в $W_{\text{пр}}$ в некотором простом минимальном сечении W в S с пропускной способностью $c(W) = c_{min}$, причем все прямые дуги этого простого сечения насыщены, а для всех обратных дуг $\overleftarrow{e}$ имеем $f(\overleftarrow{e}) = 0$. Тогда

$$M_f = \sum_{\vec{e} \in W_{\text{пр}}} f(\vec{e}) - \sum_{\overleftarrow{e} \in W_{\text{об}}} f(\overleftarrow{e}) = \sum_{\vec{e} \in W_{\text{пр}}} c(\vec{e}) = c_{min}.$$

Теорема (Форда-Фалкерсона о максимальном потоке). Любая транспортная сеть $S = (V,E,s,t,c,f)$ имеет максимальный поток и его величина равна c_{min}.

Доказательство. По теореме 1 строим полный поток в сети S. Тогда любой путь в S между s и t насыщен (имеет насыщенную дугу). По теореме 2 строим поток, в котором любая цепь между s и t, имеющая обратные дуги, не допускает увеличения потока. По теореме 3 построенный поток максимален и его величина равна c_{min} есть пропускная способность сети S.

25.6.1. Алгоритм вычисления максимального потока в транспортной сети

Пусть $S = (V,E,s,t,c)$ есть транспортная сеть с пропускной способностью дуг $c: E \to \mathbb{N}$, для которой требуется построить максимальный поток $f_{max}: E \to \mathbb{N}$.

1. Исходим из некоторого начального, например, из нулевого потока.

2. Перебираем все ориентированные пути в сети S от s до t

и увеличиваем исходный поток до насыщения одной из дуг рассматриваемого пути согласно теореме 1. Именно, пусть μ есть очередной путь между s и t в сети S с потоком f. Вычисляем $\delta = \min_{e \in \mu} (c(e) - f(e))$. Увеличивая на δ поток f в каждой дуге e из μ, придем к потоку $f' = f + \delta$. Дуга e, на которой $c(e) - f(e) = \delta$, оказывается насыщенной. В результате получим полный поток, всякий ориентированный путь которого содержит насыщенную дугу.

3. Перебираем все неориентированные пути (цепи) в сети S от s до t и увеличиваем исходный полный поток на каждой рассматриваемой цепи согласно теореме 2. Именно, пусть $\nu=[s,t]$ есть очередная цепь между s и t в сети S с потоком f. Пусть $\vec{e}$ есть прямая дуга в цепи ν, а дуга $\overleftarrow{e}$ обратная. Для дуги $e \in \nu$ положим $\delta(\vec{e}) = c(\vec{e}) - f(\vec{e})$; $\delta = \min(\delta(\vec{e}))$ по всем $\vec{e} \in \nu$; $\eta = \min(f(\overleftarrow{e}))$ по всем $\overleftarrow{e} \in \nu$; $\varepsilon = \min(\delta,\eta)$. Если $\varepsilon > 0$, то увеличивая на ε поток на каждой прямой дуге $\vec{e} \in \nu$ и уменьшая на ε поток на каждой обратной дуге $\overleftarrow{e} \in \nu$, получим новый поток $f' = f + \varepsilon$. В результате получим максимальный поток f_{max}.

Пример. Построить максимальный поток в транспортной сети $S = (V,E,s,t,c)$ заданной своими дугами, третьей координатой которых является пропускная способность c дуг:

$$e_1=(s,1,5);\ e_2=(s,2,7);\ e_3=(s,3,9);\ e_4=(1,2,1);$$

$$e_5=(1,4,4);\ e_6=(2,5,3);\ e_7=(3,5,1);\ e_8=(3,t,1);$$

$$e_9=(4,5,4);\ e_{10}=(4,t,2);\ e_{11}=(5,t,6).$$

1. Исходим из начального нулевого потока $f_0(e) \equiv 0$.

2. Перебираем все ориентированные пути между s и t, по которым возможно увеличение потока (рис.25.6).

$s \to 1 \to 2 \to 5 \to t$ – очередной путь μ между s и t;
5 1 3 6 – пропускная способность $c(e)$ дуг;
0 0 0 0 – старый поток $f_0(e)$;
5 1 3 6 – $\delta = \min_{e \in \mu} (c(e) - f_0(e)) = 1$;
1 1 1 1 – новый поток $f_1(e) = f_0(e)+\delta$.

Прямая дуга (1,2,1) насыщена.

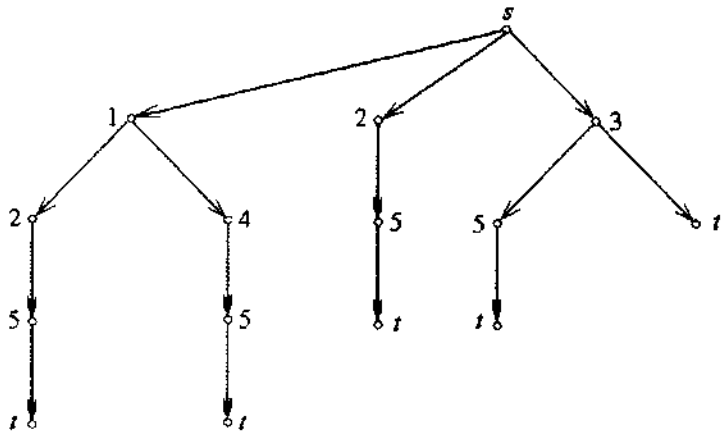


Рис.25.6

$s \to 1 \to 4 \to 5 \to t$ – очередной путь μ между s и t;

5 4 4 6 – пропускная способность $c(e)$ дуг;

1 0 0 1 – старый поток $f_1(e)$;

4 4 4 5 – $\delta = \min\limits_{e \in \mu} (c(e) - f_1(e)) = 4$;

5 4 4 5 – новый поток $f_2(e) = f_1(e) + \delta$.

Прямые дуги $(s,1,5)$; $(1,4,4)$; $(4,5,4)$ насыщены.

$s \to 2 \to 5 \to t$ – очередной путь μ между s и t;

7 3 6 – пропускная способность $c(e)$ дуг;

0 1 5 – старый поток $f_2(e)$;

7 2 1 – $\delta = \min\limits_{e \in \mu} (c(e) - f_2(e)) = 1$;

1 2 6 – новый поток $f_3(e) = f_2(e) + \delta$.

Прямая дуга $(5,t,6)$ насыщена.

$s \to 3 \to 5 \to t$ есть очередной путь μ между s и t.

Дуга $(5,t,6)$ в этом пути уже насыщена. Увеличение потока по этому пути невозможно.

$s \to 3 \to t$ – очередной путь μ между s и t;

9 1 – пропускная способность $c(e)$ дуг;

0 0 – старый поток $f_3(e)$;

9 1 – $\delta = \min\limits_{e \in \mu} (c(e) - f_3(e)) = 1$;

1 1 – новый поток $f_4(e) = f_3(e) + \delta$.

Прямая дуга $(3,t,1)$ насыщена.

3. Перебираем все неориентированные пути (цепи) между s и t, по которым возможно увеличение потока (рис.25.7).

$s \to 1$. Продолжение цепи от узла 1 не имеет смысла: прямая дуга $(s,1,5)$ насыщена.

$s \to 2 \leftarrow 1 \to 4$. Продолжение цепи от узла 4 не имеет смысла: прямая дуга $(1,4,4)$ насыщена.

$s \to 2 \to 5 \leftarrow 3$. Продолжение цепи от узла 3 не имеет смысла: обратная дуга $(3,5,1)$ нагружена нулем.

$s \to 2 \to 5 \leftarrow 4 \leftarrow 1$. Продолжение цепи без повторов вершин на этом пути невозможно.

$s \to 2 \to 5 \leftarrow 4 \to t$ – очередная цепь μ между s и t;

$\vec{e}_2$ $\vec{e}_6$ $\overleftarrow{e}_9$ $\vec{e}_{10}$ – направленность дуг в цепи μ;

7 3 4 2 – пропускная способность $c(e)$ дуг;

1 2 4 0 – старый поток $f_4(e)$;

6 1 – 2 – $\delta = \min\limits_{e \in \mu} (c(\vec{e}) - f_4(\vec{e})) = 1$;

– – 4 – – $\eta = \min\limits_{e \in \mu} (f_4(\overleftarrow{e})) = 4$; $\varepsilon = \min(\delta, \eta) = 1$;

2 3 3 1 – новый поток $f_5(e) = f_4(e)\begin{cases} +\varepsilon \text{ на } \vec{e} \\ -\varepsilon \text{ на } \overleftarrow{e} \end{cases}$.

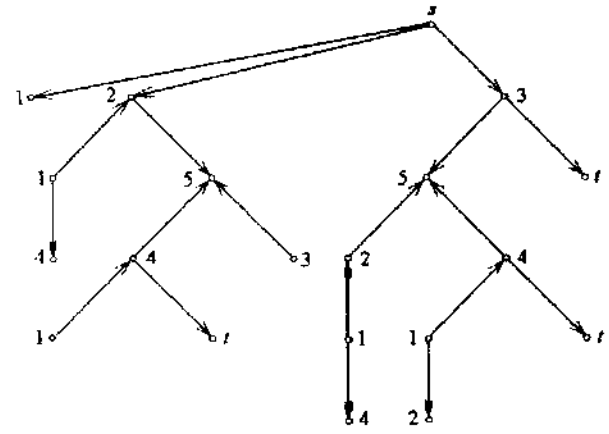


Рис.25.7

Прямая дуга (2,5,3) насыщена.

$s \to 3 \to 5 \leftarrow 2 \leftarrow 1 \to 4$. Продолжение цепи от узла 4 не имеет смысла: прямая дуга (1,4,4) насыщена.

$s \to 3 \to 5 \leftarrow 4 \leftarrow 1 \to 2$. Продолжение цепи от узла 2 не имеет смысла: прямая дуга (1,2,1) насыщена.

$s \to 3 \to 5 \leftarrow 4 \to t$ – очередная цепь μ между s и t;

$\vec{e}_3$ $\vec{e}_7$ $\overleftarrow{e}_9$ $\vec{e}_{10}$ – направленность дуг в цепи μ;

9 1 4 2 – пропускная способность $c(e)$ дуг;

1 0 3 1 – старый поток $f_5(e)$;

8 1 – 1 – $\delta = \min\limits_{e \in \mu} (c(\vec{e}) - f_5(\vec{e})) = 1$;

– – 3 – – $\eta = \min\limits_{e \in \mu} (f_5(\overleftarrow{e})) = 3$; $\varepsilon = \min(\delta, \eta) = 1$;

2 1 2 2 – новый поток $f_6(e) = f_5(e)\begin{cases} +\varepsilon \text{ на } \vec{e} \\ -\varepsilon \text{ на } \overleftarrow{e} \end{cases}$.

Прямые дуги (3,5,1) и (4,t,2) насыщены.

$s \to 3 \to t$.

Увеличение потока невозможно: прямая дуга (3,t,1) насыщена.

В табл.15.1 приведены значения промежуточных потоков; насыщенные дуги подчеркнуты.

Максимально возможная величина потока (нагружающая дуги истока, равно как и дуги стока) $M_{f_{max}} = 9$.

25.6.2. Помечивающий алгоритм Дейкстры вычисления максимального потока в транспортной сети

Пусть $S = (V,E,s,t,c)$ есть транспортная сеть и $v_1, v_2, \ldots, v_n$ есть внутренние вершины сети.

1. Задать начальный поток f, например, нулевой.
2. Вершину s пометим знаком s.
3. Присвоим всем вершинам сети пометки:

если v_i помеченная вершина, то помечаем

а) знаком $+i$ все непомеченные вершины v_j, для которых в дуге $e=(v_i,v_j)$, исходящей из v_i, имеем $f(e)<c(e)$;

б) знаком $-i$ все непомеченные вершины v_j, для которых в дуге $e=(v_j,v_i)$, заходящей в v_i, имеем $f(e)>0$.

4. Если полюс t получил пометку, то между s и t существует неориентированный путь (его следует строить от t), вершины которого помечены номерами предыдущих вершин (со знаком плюс или минус) и который допускает увеличение потока до потока f' по правилу построения потока для найденного пути. Стираем все пометки вершин и переходим к пункту 2 с новым потоком. Если полюс t пометки не получил, то последний построенный поток максимален.

Таблица 25.1

Дуга	Последовательный поток							
	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_{max}
e_1=(s,1,5)	0	1	5					5
e_2=(s,2,7)	0			1		2		2
e_3=(s,3,9)	0				1		2	2
e_4=(1,2,1)	0	1						1
e_5=(1,4,4)	0		4					4
e_6=(2,5,3)	0	1		2		3		3
e_7=(3,5,1)	0						1	1
e_8=(3,t,1)	0				1			1
e_9=(4,5,4)	0		4			3	2	2
e_{10}=(4,t,2)	0					1	2	2
e_{11}=(5,t,6)	0	1	5	6				6

Пример. Пусть

$S = (V,E,s,t,c)$ есть транспортная сеть, где

$V = \{s,1,2,3,4,5,t\}$,

$E = \{e_1=(s,1,5);\ e_2=(s,2,7);\ e_3=(s,3,9);\ e_4=(1,2,1);$

e_5=(1,4,4); e_6=(2,5,4); e_7=(3,5,1); e_8=(3,t,1);

e_9=(4,5,4); e_{10}=(4,t,2); e_{11}=(5,t,6)}.

Пропускная способность дуги $e = (i,j,c)$ есть ее третья координата c.

Построить в сети S максимальный поток $f_{max}: E \to \mathbb{N}$ и минимальное сечение (разрез).

Решение. Граф-схема транспортной сети S приведена на рис. 25.14. В скобках приведены пропускные способности ребер.

Положим начальный поток f_0 нулевым. Пометим вершины сети (рис.25.15).

Ход от вершины t до вершины s: t,5,4,1,s.

$s \to 1 \to 4 \to 5 \to t$ – очередная цепь μ между s и t;

$\vec{e}_1$ $\vec{e}_5$ $\vec{e}_9$ $\vec{e}_{11}$ – направленность дуг в цепи μ;

5 4 4 6 – пропускная способность $c(e)$ дуг;

0 0 0 0 – старый поток $f_0(e)$;

5 4 4 6 – $\delta = \min\limits_{e \in \mu} (c(\vec{e}) - f_0(\vec{e})) = 4$;

4 4 4 4 – новый поток $f_1(e) = f_0(e)+\delta = f_0(e)+4$.

Новый поток f_1 и новая разметка вершин сети приведена на рис.25.16.

Ход от вершины t до вершины s: t,5,2,s.

$s \to 2 \to 5 \to t$ – очередная цепь μ между s и t;

$\vec{e}_2$ $\vec{e}_6$ $\vec{e}_{11}$ – направленность дуг в цепи μ;

7 4 6 – пропускная способность $c(e)$ дуг;

0 0 4 – старый поток $f_1(e)$;

7 4 2 – $\delta = \min\limits_{e \in \mu} (c(e) - f_1(e)) = 2$;

2 2 6 – новый поток $f_2(e) = f_1(e)+\delta = f_1(e)+2$.

Новый поток f_2 и новая разметка вершин сети приведена на рис.25.17.

Ход от вершины t до вершины s: t,3,s.

$s \to 3 \to t$ – очередная цепь μ между s и t;

$\vec{e}_3$ $\vec{e}_8$ – направленность дуг в цепи μ;

9 1 – пропускная способность $c(e)$ дуг;

0 0 – старый поток $f_2(e)$;

9 1 – $\delta = \min\limits_{e \in \mu} (c(e) - f_2(e)) = 1$;

1 1 – новый поток $f_3(e) = f_2(e)+\delta = f_2(e)+1$.

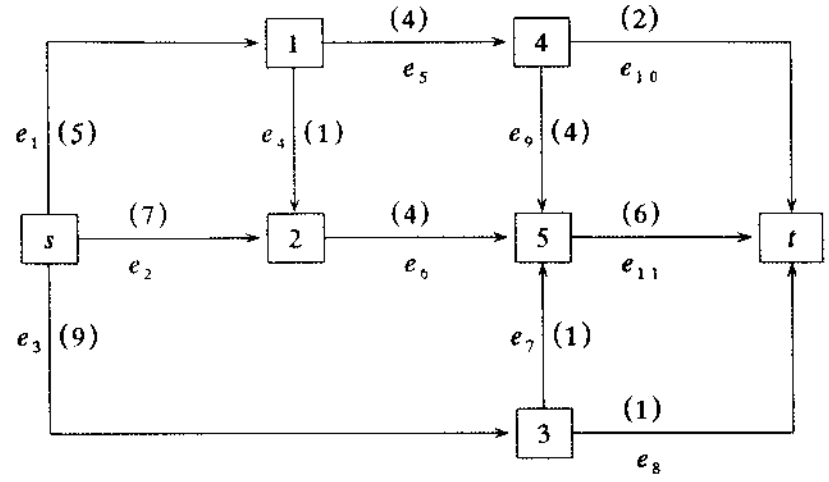
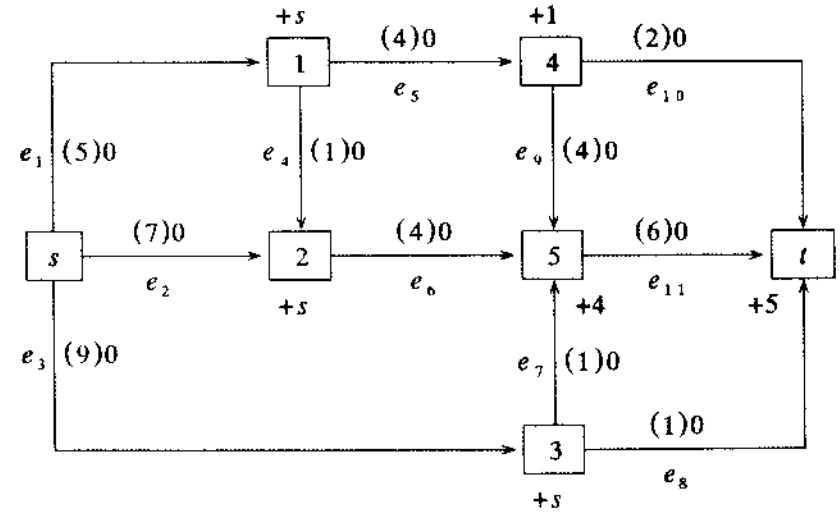


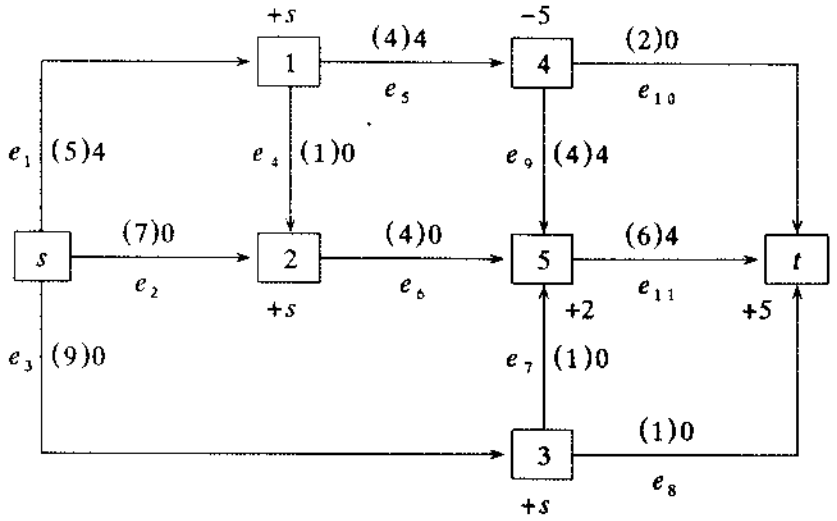
Рис.25.14



Поток f_0

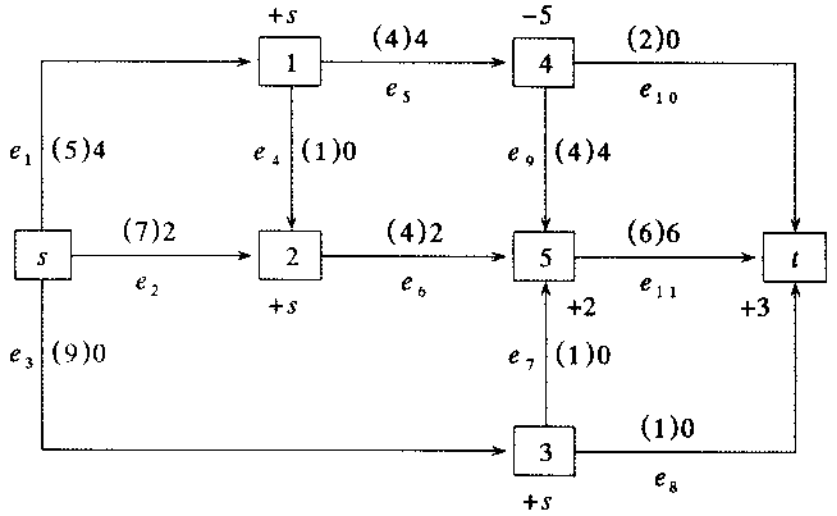
Рис.25.15

Новый поток f_3 и новая разметка вершин сети приведена на рис.25.18.



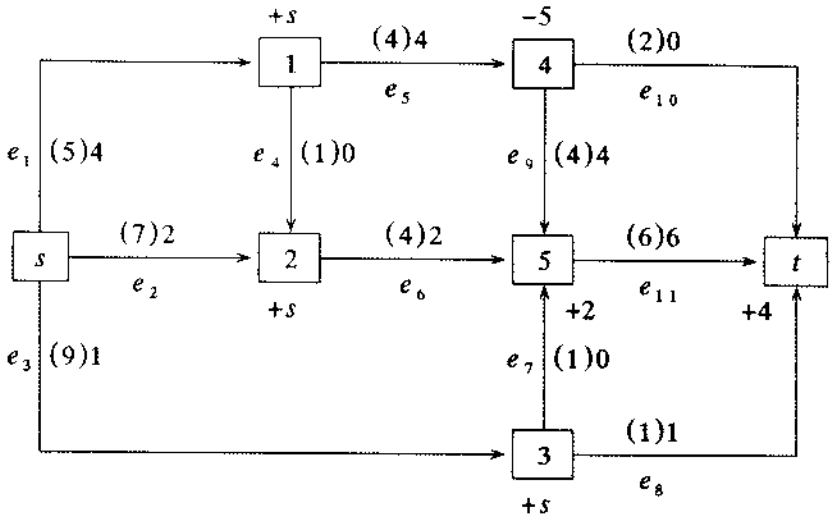
Поток f_1

Рис.25.16



Поток f_2

Рис.25.17



Поток f_3

Рис.25.18

Ход от вершины t до вершины s: $t,4,5,2,s$.

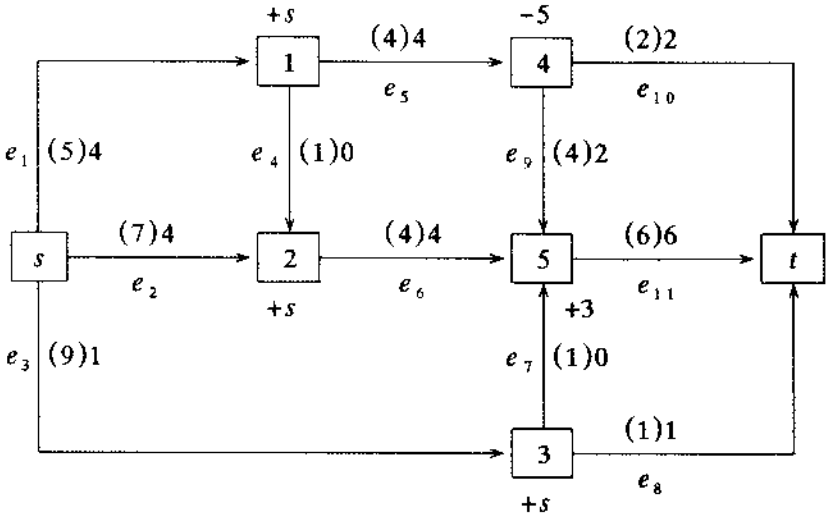
$s \to 2 \to 5 \leftarrow 4 \to t$ – очередная цепь μ между s и t;

$\vec{e}_2$	$\vec{e}_6$	$\overleftarrow{e}_9$	$\vec{e}_{10}$	
7	4	4	2	– пропускная способность $c(e)$ дуг;
2	2	4	0	– старый поток $f_3(e)$;
5	2	–	2	– $\delta = \min\limits_{e \in \mu} (c(\vec{e}) - f_3(\vec{e})) = 2$;
–	–	4	–	– $\eta = \min\limits_{e \in \mu} (f_3(\overleftarrow{e})) = 4$; $\varepsilon = \min(\delta,\eta) = 2$;
4	4	2	2	– новый поток $f_4(e) = f_3(e)\begin{cases}+\varepsilon \text{ на } \vec{e} \\ -\varepsilon \text{ на } \overleftarrow{e}\end{cases}$.

Новый поток f_4 и новая разметка вершин сети приведена на рис.25.19.

Вершина t пометки не получила. От s до t новой цепи построить не удается. Последний поток f_4 есть максимальный и величина потока $M_{f_4} = 9$.

Максимально возможная величина потока (нагружающая дуги истока, равно как и дуги стока) $M_{f\,max} = 9$.



Поток f_4

Рис.25.19

Минимальный разрез MS есть множество дуг для f_4 на рис. 25.19, заходящих в непомеченные вершины из помеченных вершин. $MS = \{e_8,e_{10},e_{11}\} = \{(3,t,1),(4,t,2),(5,t,6)\}$. Пропускная способ-ность минимального разреза $c_{\min} = M_{f\,max} = f_4(e_8) + f_4(e_{11}) + f_4(e_{12}) = 1+2+6 = 9$.

Ответ. $MS = \{e_8,e_{10},e_{11}\}$,
$M_{f\,max} = f_4(e_1)+f_4(e_2)+f_4(e_3) = 4+4+1 = 9$.

26. ПЕРЕЧИСЛЕНИЕ ГРАФОВ

26.1. Число помеченных графов

Определение. (p,q)-граф $G = (V,E)$ с множеством $V=\{v_1,\ldots,v_p\}$ из p вершин и множеством $E = \{e_1=(v_{i_1},v_{j_1}),\ldots,e_q=(v_{i_q},v_{j_q})\}$ из q ребер *помечен*, если взаимно однозначная функция $f: V\to P$, где $P=\{1,2,\ldots,p\}$, каждой вершине v из G приписывает пометку $f(v)$. Пометки f_1,f_2 графа G одинаковы, если существует изоморфизм $\varphi: G\to G$, сохраняющий пометки вершин: $f_1(v)=f_2(\varphi(v))$.

На рис.26.1 изображены все различные пометки графа $G=(V=\{v_1,v_2,v_3,v_4\},E=\{(1,2),(2,3),(3,4),(4,1),(1,3)\},P=\{1,2,3,4\}$.

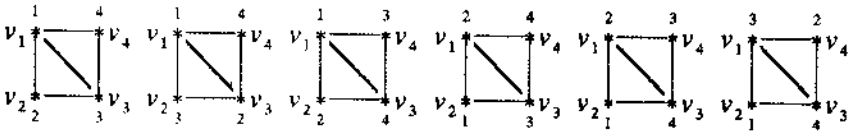


Рис.26.1

Теорема. Производящая функция $G_p(x)$ для числа помеченных графов с p вершинами задается соотношением

$$G_p(x) = (1+x)^m = \sum_{k=0}^{m}\binom{m}{k}x^k, \text{ где } m = \binom{p}{2}.$$

Доказательство. Пусть $P=\{1,\ldots,p\}$ есть множество из p пометок для p вершин. Каждые две вершины можно соединить ребром и пометить следующим образом: $1\!*\!-\!-\!*2$, $1\!*\!-\!-\!*3$, …, $1\!*\!-\!-\!*p$, $2\!*\!-\!-\!*3$, $2\!*\!-\!-\!*4,\ldots,$ $2\!*\!-\!-\!*p,\ldots,$ $p-1\!*\!-\!-\!*p$. Число ребер $m = \binom{p}{2}$. Из этого набора m ребер можно выбрать k ребер (образующих (p,k)-граф) $\binom{m}{k}$ способами. Тогда $G_p(x) = \sum_{k=0}^{m}\binom{m}{k}x^k = (1+x)^m$ есть производящая функция для числа помеченных графов с p вершинами. Коэффициент при x^k равен числу помеченных графов с p вершинами и k ребрами.

Замечание. $G_p(1) = (1+1)^m = 2^m$ есть число помеченных графов с p вершинами.

26.2. Число помеченных деревьев

Дерево есть связный граф, не имеющий циклов. Индукцией по числу p вершин можно показать, что (p,q)-дерево имеет $q=p-1$ ребер.

Теорема (Кэли). Число помеченных деревьев с p вершинами равно $t_p = p^{p-2}$.

Доказательство (Прюфер). Пусть T есть помеченное дерево с p пометками $P=\{1,2,\ldots,p\}$. Сопоставим дереву T единственный набор $(a_1,\ldots,a_{p-2})$ из P^{p-2} следующим образом. Выберем в T висячую вершину v с наименьшей пометкой a_1. Аналогично найдем наименьшую пометку a_2 для висячей вершины в дереве $T-v$ (T без v вместе с принадлежащим v ребром). И так далее. На шаге $p-2$ выберем пометку a_{p-2}. От дерева T останется одно

ребро. Набор $\mathbf{a}=(a_1,\ldots,a_{p-2})$ для дерева T построен. Так как каждое a_i было найдено единственным образом, то набор **a** единственнен. Так как каждому помеченному дереву с p вершинами соответствует единственный набор длины $p-2$ из p элементов и так как число таких наборов равно p^{p-2}, то число деревьев $t_p \leqslant p^{p-2}$.

Утверждение. Каждому набору длины $p-2$ из $1,2,\ldots,p$ можно единственным образом сопоставить помеченное дерево с p вершинами $1,2,\ldots,p$.

Доказательство. Индукция по p.

Базис. $p=2$, $P=\{1,2\}$, $p^{p-2} = 2^{2-2} = 2^0 = 1$, набор **a** пуст. Для пустого набора существует только одно помеченное дерево $1*\!\!-\!\!-\!\!*2$.

Предположение индукции. Предположим, что каждому набору длины $p-3$ из $1,2,\ldots,p-1$ можно единственным образом сопоставить помеченное дерево с $p-1$ вершинами $1,2,\ldots,p-1$.

Шаг индукции. Пусть набор $\mathbf{a}=(a_1,\ldots,a_{p-2})$ составлен из $1,2,\ldots,p$. Пусть b_1 есть наименьшее натуральное число, не встречающееся в наборе **a**, и пусть $\mathbf{c}=(c_2,\ldots,c_{p-2})$ есть набор длины $p-3$, получающийся из **a** уменьшением всех его координат, больших чем b_1, на 1. Тогда набор **c** состоит из чисел, принадлежащих множеству $P_{p-1} = \{1,2,\ldots,p-1\}$. Так как набор **c** состоит из $p-3$ элементов, то по предположению индукции набору **c** соответствует единственное помеченное дерево T_{p-1} с $p-1$ вершинами из P_{p-1}. Прибавим в дереве T_{p-1} по единице к каждой пометке вершины, превосходящей b_1-1. Затем введем вершину p с пометкой b_1 и соединим ее с вершиной, помеченной числом a_1 в дереве T_{p-1}. Таким образом, однозначно получено дерево T с p вершинами, соответствующее набору **a**.

Шаг индукции установлен, утверждение доказано.

Продолжим доказательство теоремы. Так как каждому из p^{p-2} наборов длины $p-2$ однозначно соответствует помеченное дерево с p вершинами, то $p^{p-2} \leqslant t_p$. Вместе с ранее полученным $p^{p-2} \geqslant t_p$ получаем $t_p = p^{p-2}$. Теорема доказана.

Замечание. Пусть граф $G = (V,E)$ имеет множество вершин $V = \{v_1,\ldots,v_p\}$ и ребер E. Пусть

A есть матрица смежности (соседства вершин) графа G,

M есть матрица, полученная из матрицы $-A$ заменой элемента i главной диагонали на степень вершины v_i, то есть на число ребер, принадлежащих вершине v_i.

Стягивающее дерево графа G есть наименьшее по числу ребер

подграф-дерево графа G, соединяющее все вершины в G.

Матричная теорема о деревьях (Кирхгоф). Все алгебраические дополнения матрицы M равны между собой и их общее значение равно числу стягивающих деревьев (каркасов) графа G.

На рис.26.2 представлен граф G и три его стягивающих дерева.

Для графа G вычисления дают следующее. Матрицы

$$A=\begin{bmatrix}0&1&1&0\\1&0&1&0\\1&1&0&1\\0&0&1&0\end{bmatrix},\quad M=\begin{bmatrix}2&-1&-1&0\\-1&2&-1&0\\-1&-1&3&-1\\0&0&-1&-1\end{bmatrix},\quad A_{14}=(-1)^{1+4}\begin{vmatrix}-1&2&-1\\-1&-1&3\\0&0&-1\end{vmatrix}=3.$$

26.3. Графы и группы подстановок

26.3.1. Группы подстановок и лемма Бернсайда

Напомним: группа есть множество G с одной двуместной операцией $x*y$: $G\times G\to G$, удовлетворяющая следующим аксиомам: $\forall x,y,z\in G$

1) $x*(y*z) = (x*y)*z$;

2) $\exists e\in G\ e*x = x$;

3) $\forall x\in G\ \exists x^{-1}\in G\ x*x^{-1} = x^{-1}*x = e$.

Группа G коммутативна (абелева), если

4) $x*y = y*x$.

Замечание. 1. Знак "*" называют также знаком умножения.

2. Знак "*" иногда заменяют точкой "$\cdot$" или знаком суммы "+" и тогда группу называют мультипликативной или аддитивной соответственно. Знаки "*" и "$\cdot$" часто опускают.

3. Можно показать, что для любого элемента h из G произведение $h\cdot G=(h\cdot g:\ g\in G\}$ исчерпывает все элементы группы G.

4. Подмножество H группы G есть подгруппа группы G, если H по отношению к операции умножения является группой.

5. Подгруппа $\{e\}$ состоящая только из единичного элемента и сама группа G называются несобственными подгруппами группы G. Все остальные подгруппы называются собственными.

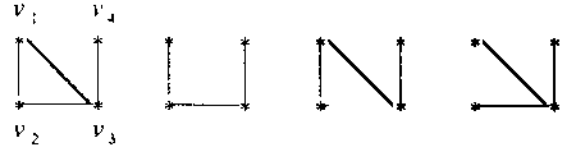


Рис.26.2

6. Везде далее будем рассматривать конечные группы (имеющие конечное число элементов).

7. Порядок конечной группы есть число ее элементов.

8. (Теорема Лагранжа). В конечной группе порядок ее подгруппы есть делитель порядка группы.

9. Индекс группы G по подгруппе H есть число $|G|/|H|$.

Определение. Подстановка элементов множества X есть взаимно однозначная функция $p: X \to X$.

Пример. $X=\{1,2,3,4\}$, $p = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$, $p(1)=3$, $p(2)=1$, $p(3)=4$, $p(4)=2$.

Пусть $X=\{1,2,\ldots,n\}$. Введем умножение подстановок:

$$p \cdot q = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} \cdot \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix},$$

то есть положив $(p \cdot q)(i)=q(p(i))$ $\forall i=1,2,\ldots,n$.

Определим единицу $e=\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}$ как тождественную подстановку и обратный элемент $p^{-1}=\begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ 1 & 2 & \ldots & n \end{pmatrix}$.

Можно проверить, что множество всех подстановок множества X есть (некоммутативная) группа.

Группа подстановок множества, состоящего из n элементов, называется *симметрической группой* и обозначается через S_n.

Подгруппы симметрической группы S_n называются группами подстановок множества $X=\{1,2,\ldots,n\}$.

Утверждение. Всякая конечная группа порядка n изоморфна некоторой подгруппе симметрической группы S_n.

Пусть $G=\{p_0=e,p_1,\ldots,p_{r-1}\}$ есть некоторая группа подстановок на множестве $X=\{1,2,\ldots,n\}$ с тождественной подстановкой p_0. Введем отношение $x \sim y$ на X следующим образом.

$x \sim y \longleftrightarrow \exists p \in G\ (p(x) = y)$.

Утверждение. Отношение $x \sim y$ есть отношение эквивалентности.

Доказательство. Покажем, что отношение $x \sim y$ удовлетворяет трем аксиомам эквивалентности:

1) $x \sim x$, 2) $x \sim y \to y \sim x$, 3) $x \sim y \,\&\, y \sim z \to x \sim z$.

В самом деле,

1) $x \sim x$, ибо $\exists p=p_0 \in G\ (p_0(x) = x)$.

2) пусть $x \sim y$. Тогда $\exists p \in G\ (p(x)=y)$ и тогда $\exists q=p^{-1}$, для которой $q(y) = p^{-1}(y) = x$. Следовательно, $y \sim x$. Получили $x \sim y \to y \sim x$.

3) пусть $x \sim y$ и $y \sim z$. Тогда

$\exists p \in G\ (p(x)=y)$, $\exists q \in G\ (q(y)=z)$,

$\exists r=pq \in G\ (r(x) = (pq)(x) = q(p(x)) = q(y) = z)$,

$\exists r \in G\ (r(x)=z)$, $x \sim z$.

Получили $x \sim y \,\&\, y \sim z \to x \sim z$.

Справедливость трех аксиом эквивалентности показана. Утверждение доказано.

Замечание. Отношение эквивалентности $x \sim y$ индуцирует разбиение множества X на попарно непересекающиеся классы эквивалентных между собой элементов.

Определение. Класс эквивалентности в множестве X по отношению $x \sim y$ называется орбитой группы подстановок G. Длина орбиты есть число ее элементов.

Замечание. $\forall a \in X$ множество $A(a) = \{p_0(a)=a,\ p_1(a),\ldots,p_{r-1}(a)\}$ перечисляет все элементы, эквивалентные a, и потому $A(a)$ есть орбита группы G.

Замечание. Орбита O замкнута относительно любой функции $p(x)$ из G, то есть $\forall a \in O\ \forall p \in G\ (p(a) \in O)$, ибо если $a \in O$ и подстановка $p \in G$, то $a \sim p(a)$ и потому $p(a) \in O$. Всякие два элемента из O можно перевести друг в друга некоторой подстановкой из G, ибо если $a,b \in O$, то $a \sim b$ и потому $p(a)=b$ для некоторой $p \in G$.

Утверждение. Для всякой орбиты $O=\{i_1,\ldots,i_u\}$ группы подстановок G, для всякого элемента $c \in O$ орбита $O = A(c)=\{p_0(c)=c,p_1(c),\ldots,p_{r-1}(c)\}$.

Доказательство. Так как всякие два элемента из O можно перевести друг в друга некоторой подстановкой из G, то, зафиксировав любой элемент c в O, получим $i_1=p_{i1}(c),\ldots,i_u = p_{iu}(c)$ для некоторых подстановок $p_{i1},\ldots,p_{iu}$ из G, и тогда $O=\{p_{i1}(c),\ldots,p_{iu}(c)\} \subseteq \{p_0(c)=c,p_1(c),\ldots,p_{i1}(c),\ldots,p_{iu}(c),\ldots,p_{r-1}(c)\}=A(c)$. Отсюда получаем $O \subseteq A(c)$.

Так как $\forall a \in O\ \forall p \in G\ (p(a) \in O)$, то множество $A(c) = \{p_0(c) = c,\ p_1(c),\ldots,p_{r-1}(c)\} \subseteq O$. Тогда $A(c) \subseteq O$ и потому $O = A(c)$.

Следствие. Все орбиты группы подстановок G исчерпываются орбитами вида $A(a)$.

Замечание. Будучи классами эквивалентности, орбиты $A(a)$ и $A(b)$ либо не пересекаются, либо пересекаются и тогда совпа-

дают.

Замечание. Множество X распадается в сумму (объединение) попарно непересекающихся между собой орбит группы G.

Вычислим длину орбиты и число орбит данной группы подстановок.

Определение. Подстановка p сохраняет элемент a (оставляет a неподвижным), или a есть неподвижная точка подстановки p, если $p(a)=a$.

Определение. Стабилизатор элемента a из X для группы G есть множество $G_a = \{p\in G: p(a)=a\}$ подстановок, сохраняющих элемент a.

Замечание. Стабилизатор G_a есть группа.

Утверждение. Длина орбиты $A(a)$ равна индексу стабилизатора G_a в группе G, то есть $|A(a)| = |G|/|G_a|$.

Доказательство. Пусть $G=\{p_0=e,p_1,\dots,p_{r-1}\}$ и $G_a=\{e=q_0,q_1,\dots,q_{s-1}\}$ и орбита $A(a)=\{p_0(a)=a,p_1(a),\dots,p_{r-1}(a)\}$. По теореме Лагранжа в G существуют подстановки $v_0=e,v_1,\dots,v_{t-1}$, для которых все подстановки таблицы R

$$G_a\cdot v_0=\{p_0=q_0v_0=v_0=e,\ p_1=q_1v_0,\ p_2=q_2v_0,\dots,p_{s-1}=q_{s-1}v_0\},$$
$$G_a\cdot v_1=\{p_s=q_0v_1=v_1,\ p_{s+1}=q_1v_1,\ p_{s+2}=q_2v_1,\dots,p_{2s-1}=q_{s-1}v_1\},$$
$$\dots$$
$$G_a\cdot v_{t-1}=\{p_{(t-1)s}=q_0v_{t-1}=v_{t-1},\ \dots,\ p_{rs-1}=q_{s-1}v_{t-1}\}$$

попарно различны и исчерпывают всю группу G (разложение группы G по подгруппе G_a на правые смежные классы $G_a\cdot v_0,\dots,G_a\cdot v_{t-1}$, которые попарно не пересекаются и на которые распадается группа G). Значения подстановок ряда R на элементе a дают орбиту $A(a)$. Тогда $r = |G| = s\cdot t = |G_a|\cdot t$.

Для любой строки $i=0,1,\dots,t-1$

$$p_{is}=q_0v_i,\ p_{is+1}=q_1v_i,\ p_{is+2}=q_2v_i,\dots,p_{(i+1)s-1}=q_{s-1}v_i,$$

для любого члена $j=0,1,\dots,s-1$ в ней и для элемента a имеем

$$p_{is+j}(a) = (q_j\cdot v_i)(a) = v_i(q_j(a)) = v_i(a),$$

то есть все подстановки строки i переводят элемент a в один и тот же элемент $v_i(a)$.

Покажем, что все t элементов $v_i(a)$ попарно различны. Допустим противное: $v_i(a)=v_j(a)$ для некоторых $i\neq j$. Тогда $(v_j\cdot v_i^{-1})(a)= (v_i\cdot v_i^{-1})(a)=e(a)=a$, то есть подстановка $v_j\cdot v_i^{-1}\in G_a$. Аналогично можно получить $v_i\cdot v_j^{-1}\in G_a$. Так как единица $e\in G_a$ и $v_j\cdot v_i^{-1}\in G_a$, то $v_i=e\cdot v_i\in G_a\cdot v_i$ и $v_j=(v_j\cdot v_i^{-1})\cdot v_i\in G_a\cdot v_i$. То есть подстановки v_i,v_j лежат в $G_a\cdot v_i$. Аналогично показываем, что подстановки v_i,v_j лежат в $G_a\cdot v_j$. Тогда $G_a\cdot v_i = G_a\cdot v_j$. Противоречие, ибо эти смежные классы при разных v_i,v_j

различны.

Из таблицы R построим таблицу R_a, взяв в качестве ее значений значения подстановок в R на элементе a. Таблица R_a имеет t строк, s столбцов и число элементов ней равно $s\cdot t$.

Элементы таблицы R_a составляют орбиту $A(a)$, причем каждый ряд i длины s в R_a состоит только из $v_i(a)$. Поэтому орбита $A(a) = \{v_0(a),v_1(a),\dots v_{t-1}(a)\}$, ее длина равна t, причем $r = |G| = s\cdot t = |G_a|\cdot|A(a)|$, откуда $|A(a)| = |G|/|G_a|$.

Утверждение доказано.

Лемма Бернсайда. Если

1) $X = \{1,2,\dots,n\}$,

2) $G = \{p_0=e,p_1,\dots,p_{r-1}\}$ есть группа подстановок множества X,

3) $\chi(p)$ есть число неподвижных точек подстановки p,

то число орбит группы подстановок G есть

$$f(G) = |G|^{-1}\sum_{p\in G}\chi(p).$$

Доказательство. Изобразим отношение "подстановка p сохраняет неподвижным элемент m" на графике (рис.26.3).

На каждой вертикали p_i графика отмечаются все (неподвижные) точки (звезды), сохраняемые подстановкой p_i. Их число равно $\chi(p_i)$. При подсчете по вертикали число всех звезд графика равно $\chi(p_0) + \chi(p_1) +\dots+ \chi(p_{r-1}) = \sum_{p\in G}\chi(p)$.

На каждой горизонтали m отмечаются все подстановки, сохраняющие элемент m. Эти подстановки образуют стабилизатор G_m, являющийся группой, и по предыдущей теореме

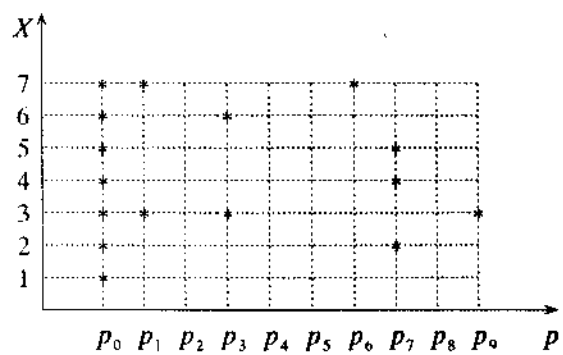


Рис.26.3

$|A(m)| = |G|/|G_m|$, откуда
$|G_m| = |G|/|A(m)|$.

При подсчете по горизонтали число звезд графика равно

$$|G_1| + |G_2| + \ldots + |G_n| = \sum_{m \in X} |G_m|.$$

Множество X распадается в сумму (объединение) попарно непересекающихся орбит: $X = O_1 \cup O_2 \cup \ldots \cup O_t$, где $t = f(G)$ есть их число. Перегруппируем слагаемые предыдущей суммы относительно элементов орбит и получим

$$\sum_{m \in X} |G_m| = \sum_{m \in O_1} |G_m| + \sum_{m \in O_2} |G_m| + \ldots + \sum_{m \in O_t} |G_m|.$$

Если для орбиты O справедливо $m \in O$, то $O = A(m)$. Тогда каждое слагаемое этой суммы

$$\sum_{m \in O} |G_m| = \sum_{m \in O} \frac{|G|}{|A(m)|} = \sum_{m \in O} \frac{|G|}{|O|} = \frac{|G|}{|O|} \sum_{m \in O} 1 = \frac{|G|}{|O|} |O| = |G|,$$

и тогда вся сумма

$$\sum_{m \in X} |G_m| = |G| + |G| + \ldots + |G| = t \cdot |G| = f(G) \cdot |G|.$$

Подсчеты звезд по вертикали и по горизонтали совпадают, поэтому $\sum_{p \in G} \chi(p) = f(G) \cdot |G|$, откуда следует требуемое соотношение.

26.3.2. Теорема Пойа

Пусть G есть группа подстановок множества $X = \{1,2,\ldots,n\}$ и конечное множество $Y = \{1,2,\ldots,m\}$ содержит не менее двух элементов. Пусть $M = Y^X$ есть множество всех функций $X \to Y$ вида $f = \begin{pmatrix} 1 & 2 & \ldots & n \\ y_1 & y_2 & \ldots & y_n \end{pmatrix}$. Число таких функций равно числу $|Y|^{|X|}$ всех наборов из $m = |Y|$ элементов множества Y по $n = |X|$ элементов в каждом наборе. Пусть подстановка p из G действует на функцию f из M и выдает функцию $g = p(f) = pf$ согласно равенству $(pf)(x) = g(x) = f(p(x))$. Каждая подстановка p из G определяет подстановку $p_M = \begin{pmatrix} f_1 & f_2 & \ldots & f_{m^n} \\ p(f_1) & p(f_2) & \ldots & p(f_{m^n}) \end{pmatrix}$ на множестве функций Y^X так, что $p_M(f_i) = p_M f_i = p(f_i)$. Обозначим множество таких подстановок через G_M.

Например, $X = \{1,2,3,4\}$, $Y = \{1,2,3\}$,

$p = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$, $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 2 & 3 \end{pmatrix}$, $g(x) = f(p(x)) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 1 \end{pmatrix}$, ибо

$(pf)(1) = f(p(1)) = f(3) = 2$,
$(pf)(2) = f(p(2)) = f(2) = 2$,
$(pf)(3) = f(p(3)) = f(4) = 3$,
$(pf)(4) = f(p(4)) = f(1) = 1$.

Действие подстановки p_M на функцию f сводится к перестановке значений функции f в соответсвии с подстановкой p.

Произведение подстановок из G_M определим обычным образом.

Тождественная подстановка является единицей множества G_M.

Обратная подстановка $(p_M^{-1}g)(x) = g(p^{-1}(x)) = f(p(p^{-1}(x))) = f(x)$. Следовательно, множество G_M есть группа. Она называется степенной группой, порожденной группой G.

Пусть G есть группа подстановок на множестве объектов $X = \{1,2,\ldots,n\}$. Каждая подстановка представима в виде произведения непересекающихся циклов.

Пусть выражение s_k^r означает: подстановка имеет r циклов длины k. Обозначим через $j_k(p)$ число циклов длины k в подстановке p. Многочлен циклов, или цикловый индекс, $Z(G)$ группы G есть многочлен от n переменных $s_1, s_2, \ldots, s_n$, определяемый формулой

$$Z(G) = Z(G, s_1, \ldots, s_n) = |G|^{-1} \sum_{p \in G} \prod_{k=1}^{n} s_k^{j_k(p)}.$$

Наибольшую длину цикла имеет подстановка $(1,2,\ldots,n)$, состоящая из единственного цикла, его длина равна $n = |X|$. Поэтому многочлен циклов имеет n переменных $s_1, \ldots, s_n$.

Например, подстановке

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 2 & 1 & 5 & 6 & 8 & 7 \end{pmatrix} = (1324)(5)(6)(78)$$

соответствует произведение $s_1^2 s_2 s_4$. Для группы подстановок G

(1)(2)(3), (1)(23), (2)(13), (3)(12), (123), (132)

множества $X = \{1,2,3\}$ многочлен циклов (цикловый индекс)

$$Z(G) = Z(G, s_1, s_2, s_3) = (1/3!)(s_1^3 + 3s_1 s_2 + 2s_3).$$

Теорема (перечисления Пойа). Пусть G есть группа подстановок порядка r на множестве $X = \{1,2,\ldots,n\}$ и множество

$Y=\{1,2,\ldots,m\}$. Число орбит степенной группы G_M (она тоже порядка r), действующей на множестве функций $X \to Y$ получается подстановкой числа m вместо каждой переменной в многочлене циклов $Z(G)$:

$$N(G)=Z(G,s_1,\ldots,s_n)\Big|_{s_1=\ldots=s_n=m}=|G|^{-1}\sum_{p\in G}\prod_{k=1}^{n}s_k^{j_k(p)}\Big|_{s_i=m}.$$

Доказательство. Каждая подстановка p из G переставляет значения функции f и дает функцию $g=p_M(f)$. Группа G определяет группу G_M подстановок множества функций M как:

$$G_M=\left\{p_M=\begin{pmatrix} f_1 & f_2 & \ldots & f_{m^n}\\ p(f_1) & p(f_2) & \ldots & p(f_{m^n})\end{pmatrix}:\ p\in G\right\}.$$

Точку f (функцию f) назовем неподвижной, если $p(f)=f$.

Образы $p(f)$ всех m^n функций из M можно разместить в таблице следующим образом.

$$\begin{array}{ccccc} p_0(f_1) & p_0(f_2) & & & p_0(f_{m^n})\\ p_1(f_1) & p_1(f_2) & & & p_1(f_{m^n})\\ \ldots & \ldots & , & \ldots, & \ldots\\ p_{r-1}(f_1) & p_{r-1}(f_2) & & & p_{r-1}(f_{m^n})\end{array}$$

Столбцы таблицы это образ одной и той же функции, значения которой переставляются подстановками из G. Но столбцы таблицы это и орбиты группы G. Согласно лемме Бернсайда это число равно числу неподвижных точек в подстановках p из G: $N(G)=|G|^{-1}\sum_{p\in G}\chi(p)$.

Для неподвижной функции f равенство $p(f)=f$ означает, что подстановка p в пределах одного своего цикла переставляет равные между собой значения функции f. Например, для подстановки $p=(a_1,\ldots,a_{l_1})(b_1,\ldots,b_{l_2})\ldots(c_1,\ldots,c_{l_u})$ из G, состоящей из u циклов длин $l_1,l_2,\ldots,l_u$ соответственно, неподвижными будут функции, приведенные в табл.26.1. (Указаны значения функции, общие всем элементам соответствующего цикла).

Число неподвижных функций, соответствующих подстановке p, равно $\chi(p)=m^u$, то есть числу наборов длины u из $\{1,2,\ldots,m\}$.

Таблица 26.1

$(a_1,\ldots,a_{l_1})$	$(b_1,\ldots,b_{l_2})$	$\ldots$	$(c_1,\ldots,c_{l_u})$
1	1	$\ldots$	1
1	1	$\ldots$	2
	$\ldots$	$\ldots$	
y_1	y_2	$\ldots$	y_u
	$\ldots$		
m	m	$\ldots$	m

Тип подстановки p, представленной как произведение u циклов, есть набор $<l_1,\ldots,l_u>$, где l_k есть длина цикла k, $k=1,2,\ldots,u$. В многочлене циклов этой подстановке будет соответствовать слагаемое $s_{l_1}s_{l_2}\ldots s_{l_u}=\prod_{k=1}^{u}s_{l_k}$. Если в произведении $s_{l_1}s_{l_2}\ldots s_{l_u}$ для p переменная s_k (то есть цикл длины k) встречается $j_k(p)$ раз, то $s_{l_1}s_{l_2}\ldots s_{l_u}=\prod_{k=1}^{n}s_k^{j_k(p)}$. Верхний индекс мы взяли равным $n\geq u$. Переменные s, которых среди $s_{l_1},\ldots,s_{l_u}$ нет, входят в произведение в нулевой степени $s^0=1$.

Каждое s_{l_k} может принимать m значений. Тогда число неподвижных функций для подстановки p есть $\chi(p)=$

$$s_{l_1}s_{l_2}\ldots s_{l_u}\Big|_{s_{l_1}=\ldots=s_{l_u}=m}=\prod_{k=1}^{n}s_{l_k}\Big|_{s_{l_k}=m}=\prod_{k=1}^{n}s_k^{j_k(p)}\Big|_{s_k=m}=m^u.$$

Например, подстановка p имеет тип <2,5,5,5,6> означает, что p имеет 5 циклов, из которых один цикл длины 2, три цикла длины 5, один цикл длины 6. В многочлене циклов подстановке p будет соответствовать слагаемое $s_2s_5s_5s_5s_6=s_2s_5^3s_6$, каждая переменная которого может принимать m значений. Тогда число неподвижных функций для p будет равно

$$s_2s_5^3s_6\Big|_{s_2=s_5=s_6=m}=m\cdot m^3\cdot m=m^5.$$

Число всех орбит

$$N(G) = |G|^{-1} \sum_{p \in G} \chi(p) = |G|^{-1} \sum_{p \in G} \prod_{k=1}^{n} s_k^{j_k(p)} \Big|_{s_k = m}.$$

26.3.3. Раскраска вершин куба

Вычислим, сколькими способами можно раскрасить вершины куба в не более, чем три цвета. Две раскраски считаются одинаковыми, если вращением куба в пространстве их раскраски можно совместить. Восемь вершин куба не более чем тремя красками, например, синей, зеленой, красной ($с$,$з$,$к$) можно раскрасить 3^8=6561 способами. Многие раскраски окажутся одинаковыми. Раскрашенный куб можно представлять себе функцией

$K: X \to Y$, где $X = \{1,2,\ldots,8\}$, $Y = \{с,з,к\}$.

Возможны, например, раскраски:

$$K_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ с & к & з & з & с & к & с & к \end{pmatrix}, \quad K_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ с & к & к & с & к & с & з & з \end{pmatrix}.$$

Пусть M есть множество кубов, раскрашенных не более, чем тремя красками. Мощность $|M| = 3^8$.

Пусть G есть группа вращений куба (рис.26.4), состоящая из 24 следующих подстановок (укажем только вторую строку подстановки).

1. Вокруг каждой из трех осей, соединяющей центры противоположных граней.

(1,5,8,4)(2,6,7,3), (1,4,3,2)(5,8,7,6),
(1,8)(2,7)(3,6)(4,5), (1,3),(2,4)(5,7)(6,8),
(1,4,8,5)(2,3,7,6), (1,2,3,4)(5,6,7,8),
(1,5,6,2)(3,4,8,7), (1,6)(2,5)(3,8)(4,7),
(1,2,6,5)(3,7,8,4).

2. Вокруг каждой из четырех диагоналей куба.

(1)(2,5,4)(3,6,8)(7), (2)(1,3,6)(4,7,5)(8),
(3)(1,6,8)(2,7,4)(5), (4)(1,3,8)(2,7,5)(6),
(1)(2,4,5)(3,8,6)(7), (2)(1,6,3)(4,5,7)(8),
(3)(1,8,6)(2,4,7)(5), (4)(1,8,3)(2,5,7)(6).

3. Вокруг каждой из шести осей, соединяющих середины противоположных ребер.

(1,5)(2,8)(3,7)(4,6), (1,2)(3,5)(4,6)(7,8),
(1,7)(2,3)(4,6)(5,8), (1,7)(2,6)(3,5)(4,8),

(1,7)(2,8)(3,4)(5,6), (1,4)(2,8)(3,5)(6,7).

Вместе с тождественной (1)(2)(3)(4)(5)(6)(7)(8) это составляет 24 подстановки группы G.

Каждая подстановка p из G определяет поворот куба K как $p(K)$. Группа G определяет группу G_M подстановок множества раскрашенных кубов M как:

$$G_M = \left\{ p_M = \begin{pmatrix} K_1 & K_2 & \ldots & K_{3^8} \\ p(K_1) & p(K_2) & \ldots & p(K_{3^8}) \end{pmatrix} : p \in G \right\},$$ причем для куба K

куб $p(K)$ есть функция $p(K) = \begin{pmatrix} 1 & 2 & \ldots & 8 \\ K(p(1)) & K(p(2)) & \ldots & K(p(8)) \end{pmatrix}$.

Все повороты всех 3^8 раскрашенных помеченных кубов из M можно разместить в таблице следующим образом.

$$\begin{matrix} p_0(K_1) & p_0(K_2) & & p_0(K_{3^8}) \\ p_1(K_1) & p_1(K_2) & \ldots, & p_1(K_{3^8}) \\ \ldots & \ldots & & \ldots \\ p_{23}(K_1) & p_{23}(K_{3^8}) & & p_{23}(K_{3^8}) \end{matrix}$$

Каждый столбец таблицы есть вращаемый 24 раза один и тот же непомеченный раскрашенный куб. Каждый столбец таблицы также есть орбита группы G_M. Поэтому число непомеченных раскрашенных кубов равно числу орбит группы G_M. Согласно лемме Бернсайда число орбит группы G_M равно числу неподвижных точек в подстановках p_M из G_M, совпадающим с числом неподвижных точек для p на M и равно $N(G_M) = N(G) = |G|^{-1} \sum_{p \in G} \chi(p)$.

Для неподвижной точки (куба) справедливо равенство $p(K) = K$, а это означает, что подстановка p переставляет вершины в пределах одной раскраски.

Например, для подстановки p = (2)(1,3,6)(4,7,5)(8), состоящей из четырех циклов, $K = p(K)$ тогда и только тогда, когда вершины каждого цикла окрашены, например, последовательно в $с$,$к$,$с$,$з$ цвета. Возможно 3^4 таких раскрасок (или неподвижных точек).

В группе вращений G куба:

1 подстановка типа <1,1,1,1,1,1,1,1> из 8 циклов дает 3^8 неподвижных точек и соответствует слагаемому s_1^8 многочлена циклов;

6 подстановок типа <4,4> (это 2 цикла длины 4) дают $6 \cdot 3^2$

неподвижных точек и соответствуют слагаемому $6s_4^2$ многочлена циклов;

9 подстановок типа <2,2,2,2> (это 4 цикла длины 2) дают $9\cdot 3^4$ неподвижных точек и соответствуют слагаемому $9s_2^4$ многочлена циклов;

8 подстановок типа <1,1,3,3> (это 4 цикла, из которых 2 длины 1 и других 2 длины 3) дают $8\cdot 3^4$ неподвижных точек и соответствуют слагаемому $8s_1^2s_3^2$ многочлена циклов.

По теореме Пойа число орбит

$$N(G_M) = |G|^{-1}(s_1^8 + 6s_4^2 + 9s_2^4 + 8s_1^2s_3^2)\Big|_{s_1=\ldots=s_8=3} =$$
$$(1/24)\cdot(3^8 + 6\cdot 3^2 + 9\cdot 3^4 + 8\cdot 3^4) = 333.$$

Вершины куба можно раскрасить не более чем тремя красками 333 различными способами.

26.3.4. Составление ожерелий

Ожерелье типа (n,k) есть правильный n-угольник, вершины которого раскрашены в не более чем k цветов.

Два ожерелья неотличимы (одинаковы), если одно можно получить из другого, поворачивая его относительно точки симметрии или симметрично отражая относительно одной из осей симметрии.

Для подсчета числа ожерелий типа (n,k) нужно найти группу G вращений и симметрий правильного n–угольника, которая есть некоторая группа подстановок на множестве $X = \{1,2,\ldots,n\}$, потом составить многочлен циклов, а затем применить теорему Пойа.

Подсчитаем число ожерелий, которые можно составить из шести бусин (рис.26.5) не более чем двух цветов, синего и красного.

Для перечисленных операций соответствующая группа G состоит из 12 следующих подстановок, которые распределены по типам следующим образом.

Повороты.

$$p_0 = \begin{pmatrix}1&2&3&4&5&6\\1&2&3&4&5&6\end{pmatrix} = (1)(2)(3)(4)(5)(6),\ <1,1,1,1,1,1>.$$

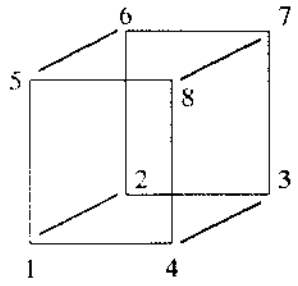


Рис.26.4

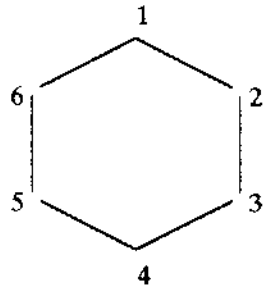


Рис.26.5

$$p_1 = \begin{pmatrix}1&2&3&4&5&6\\2&3&4&5&6&1\end{pmatrix} = (123456),\ <6>.$$

$$p_2 = \begin{pmatrix}1&2&3&4&5&6\\3&4&5&6&1&2\end{pmatrix} = (135)(246),\ <3,3>.$$

$$p_3 = \begin{pmatrix}1&2&3&4&5&6\\4&5&6&1&2&3\end{pmatrix} = (14)(25)(36),\ <2,2,2>.$$

$$p_4 = \begin{pmatrix}1&2&3&4&5&6\\5&6&1&2&3&4\end{pmatrix} = (153)(264),\ <3,3>.$$

$$p_5 = \begin{pmatrix}1&2&3&4&5&6\\6&1&2&3&4&5\end{pmatrix} = (165432),\ <6>.$$

Симметрия относительно диагоналей.

$$p_6 = \begin{pmatrix}1&2&3&4&5&6\\1&6&5&4&3&2\end{pmatrix} = (1)(26)(35)(4),\ <1,1,2,2>.$$

$$p_7 = \begin{pmatrix}1&2&3&4&5&6\\3&2&1&6&5&4\end{pmatrix} = (13)(2)(46)(5),\ <1,1,2,2>.$$

$$p_8 = \begin{pmatrix}1&2&3&4&5&6\\5&4&3&2&1&6\end{pmatrix} = (15)(24)(3)(6),\ <1,1,2,2>.$$

Симметрия относительно прямых через середины сторон.

$$p_9 = \begin{pmatrix}1&2&3&4&5&6\\2&1&6&5&4&3\end{pmatrix} = (12)(36)(45),\ <2,2,2>.$$

$$p_{10} = \begin{pmatrix}1&2&3&4&5&6\\4&3&2&1&6&5\end{pmatrix} = (14)(23)(56),\ <2,2,2>$$

$$p_{11} = \begin{pmatrix}1&2&3&4&5&6\\6&5&4&3&2&1\end{pmatrix} = (16)(25)(34),\ <2,2,2>.$$

Мы получили следующее.

1 подстановка типа <1,1,1,1,1,1> дает слагаемое s_1^6 и m^6 неподвижных точек.

2 подстановки типа <6> дают слагаемое $2s_6$ и $2m$ неподвиж-

ных точек.

2 подстановки типа <3,3> дают слагаемое $2s_3^2$ и $2m^2$ неподвижных точек.

4 подстановки типа <2,2,2> дают слагаемое $4s_2^3$ и $4m^3$ неподвижных точек.

3 подстановки типа <1,1,2,2> дают слагаемое $3s_1^2s_2^2$ и $3m^2m^2$ неподвижных точек.

По теореме Пойа число орбит

$$N(G) = |G|^{-1} \sum_{p \in G} \prod_{k=1}^{n} s_k^{j_k(p)} \Big|_{s_k = m}.$$

Из 6 бусин не более чем двух цветов можно составить

$$N(G) = |G|^{-1} \sum_{p \in G} \prod_{k=1}^{n} s_k^{j_k(p)} \Big|_{s_k = m=2} =$$

$$\frac{1}{12}(s_1^6 + 2s_6 + 2s_3^2 + 4s_2^3 + 3s_1^2s_2^2)\Big|_{s_1 = \ldots = s_6 = m=2} =$$

$(m^6 + 2 \cdot m + 2 \cdot m^2 + 4 \cdot m^3 + 3 \cdot m^2 \cdot m^2)/12 =$

$(2^6 + 2 \cdot 2 + 2 \cdot 2^2 + 4 \cdot 2^3 + 3 \cdot 2^2 \cdot 2^2)/12 =$

$(64 + 4 + 8 + 32 + 48)/12 = 156/12 = 13$ ожерелий.

Часть 5

МОНАДИЧЕСКАЯ ЛОГИКА И КОНЕЧНЫЕ АВТОМАТЫ

27. КОНЕЧНЫЕ АВТОМАТЫ

27.1. Автоматы Мили и Мура

Алфавит X есть любое непустое конечное множество символов (букв). Слово в алфавите X есть конечная последовательность символов алфавита X. Пусть * есть символ пустого слова; длина слова есть число входящих в него символов; длина пустого слова равна нулю. Два слова (графически) *равны*, если они представляют одну и ту же последовательность символов алфавита. *Конкатенация* xy двух слов x и y есть результат приписывания к слову x слова y. Пусть X^* есть множество всех слов в алфавите X. *Язык* есть произвольное подмножество слов в множестве X^*.

Автомат Мили есть система объектов $A = (X,Y,Q,q_0,T,B)$, где X есть входной алфавит; Y есть выходной алфавит; Q есть алфавит (внутренних) состояний; $q_0 \in Q$ есть начальное состояние; $T: Q \times X \to Q$ есть функция переходов; $B: Q \times X \to Y$ есть функция выходов.

Пример. Алфавиты $X = \{0,1\}$; $Y = \{0,1,2\}$; $Q = \{q_0,q_1,q_2\}$; функции переходов и выходов задаются в табл.27.1 или граф-схемой (рис.27.1).

Автомат A работает следующим образом. На вход автомата в дискретные моменты времени поступает входная последовательность $x = x(0)x(1)x(2)\ldots x(k)$ символов алфавита X. Автомат A

Таблица 27.1

		состояния		
		q_0	q_1	q_2
вход	0	$q_2,1$	$q_2,0$	$q_1,1$
	1	$q_1,2$	$q_2,2$	$q_2,2$

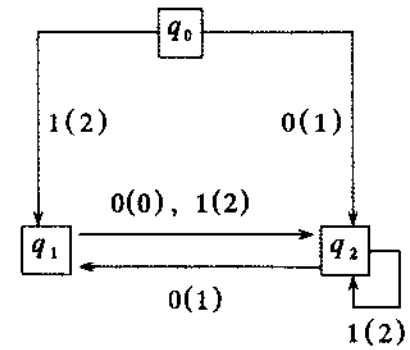


Рис.27.1

в ответ проходит последовательность состояний $q = q(0)q(1)q(2)\ldots q(k)q(k+1)$ (она называется ходом автомата A на x; множество всех таких последовательностей обозначается через $Rn(A,x)$), для которой

$$q(0) = q_0,$$
$$q(t+1) = T(q(t),x(t)),$$

и выдает выходную последовательность $y = y(0)y(1)y(2)\ldots y(k)$, для которой $y(t) = B(q(t),x(t))$, $t = 0,1,\ldots,k$.

В примере автомата из табл.16.1

$$x = 0\ \ 0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0,$$
$$q = q_0\ \ q_2\ \ q_1\ \ q_2\ \ q_1\ \ q_2\ \ q_2\ \ q_1\ \ q_2\ \ q_1,$$
$$y = 1\ \ 1\ \ 2\ \ 1\ \ 2\ \ 2\ \ 1\ \ 0\ \ 1.$$

Автомат Мили можно задать каноническими уравнениями

$$q(0) = q_0;$$
$$q(t+1) = T(q(t),x(t));$$
$$y(t) = B(q(t),x(t)).$$

Автомат Мура есть система объектов $A = (X,Y,Q,q_0,T,B)$, где X,Y,Q,q_0,T задаются как для автомата Мили, а $B: Q \to Y$ есть функция выходов. Канонические уравнения автомата Мура имеют вид

$$q(0) = q_0;$$
$$q(t+1) = T(q(t),x(t));$$
$$y(t) = B(q(t)).$$

Пример. $X = \{0,1\}$; $Y = \{0,1,2\}$; функции переходов и выходов задаются в табл.27.2 и граф-схемой на рис.27.2.

Ниже приведена входная последовательность x, а также соответствующие ей последовательность состояний q и выходная последовательность y:

$$x = 0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1;$$
$$q = q_0\ \ q_2\ \ q_0\ \ q_1\ \ q_2\ \ q_0\ \ q_2\ \ q_2;$$
$$y = 1\ \ 0\ \ 1\ \ 2\ \ 0\ \ 1\ \ 0\ \ 0.$$

Автомат Мура есть частный случай автомата Мили. Автомат Мура проще и потому в некоторых случаях предпочтительнее автомата Мили.

Таблица 27.2

	1	2	0
	q_0	q_1	q_2
0	q_2	q_2	q_0
1	q_1	q_2	q_2

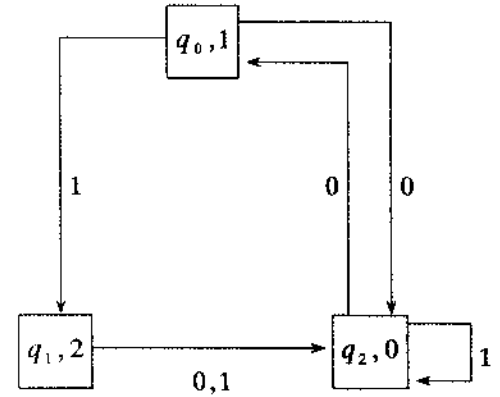


Рис.27.2

Распространим функцию переходов T автомата на множество $Q \times X^*$, положив $T(q,xa) = T(T(q,x),a)$; $q \in Q$, $x \in X^*$, $a \in X$. Значение функции $T(q,x)$ есть то состояние, в которое перейдет автомат A из состояния q под воздействием входного слова x.

Аналогичным образом распространим функцию выходов B на множество $Q \times X^*$, положив $B(q,xa) = B(T(q,x),a)$; $q \in Q$, $x \in X^*$, $a \in X$. Значение функции $B(q,x)$ есть та выходная буква y, которую выдает автомат A, находящийся в состоянии q, если на его вход подать слово x.

Пусть $Q' \subseteq Q$ и $T(Q',x) = \{q \in Q : \exists q' \in Q'\ T(q',x)=q\}$ есть множество всех состояний, в которые перейдет автомат из состояний множества Q' в качестве начальных состояний под воздействием входного слова x. Определим для автомата $A = (X,Y,Q,q_0,T,B)$ оператор $\mathbb{O}: X^* \to Y^*$, положив $\mathbb{O}(a) = B(q_0,a)$, $a \in X$; $\mathbb{O}(xa) = \mathbb{O}(x) \cdot B(T(q_0,x), a)$, $x \in X^*$, $a \in X$.

Оператор $\mathbb{O}$ по входному слову x выдает соответствующее ему выходное слово y. Определим оператор $\mathbb{O}_Q: X^* \to Q^*$, положив $\mathbb{O}_Q(a) = q_0 \cdot T(q_0,a)$, $a \in X$; $\mathbb{O}_Q(xa) = \mathbb{O}_Q(x) \cdot T(T(q_0,x),a)$; $x \in X^*$, $a \in X$.

Оператор $\mathbb{O}_Q$ по входному слову x выдает соответствующее ему слово в алфавите состояний автомата A. Если автомат A начинает работу над словом x из некоторого состояния q, то будем писать $\mathbb{O}(q,x)$, $\mathbb{O}_Q(q,x)$.

Два автомата с выходом *эквивалентны*, если они реализуют один и тот же оператор $\mathbb{O}: X^* \to Y^*$.

Теорема. По любому автомату Мили можно построить эквивалентный ему автомат Мура.

Доказательство. Пусть $A = (X,Y,Q,q_0,T,B)$ есть автомат Мили. Построим автомат Мура $A' = (X,Y,Q',q'_0,T',B')$, положив

$$q'_0 = q_0;$$
$$Q' = \{q_0\} \cup \{(q,a) : q \in Q,\ a \in X\};$$
$$T'((q,a),b) = (T(q,a),b),\ q \in Q,\ \{a,b\} \subseteq X;$$
$$B'((q,a)) = B(q,a),\ q \in Q,\ a \in X.$$

Построенный автомат Мура эквивалентен автомату Мили; однако для этого необходимо пренебречь выходом автомата Мура в начальный момент времени.

Пример. Для автомата Мили из табл.27.1 функции переходов и выходов эквивалентного ему автомата Мура приведены в табл. 27.3. Выход $B'(q_0)$ произволен.

Автомат без выхода есть система объектов $A = (X,Q,q_0,T,F)$, где (X,Q,q_0,T) задаются как у автомата с выходом, а $F \subseteq Q$ есть множество выделенных (отмеченных, заключительных, финальных) состояний. Так определенный автомат называется также детерминированным автоматом. Автомат без выхода можно рассматривать как автомат Мура с выходом, полагая его выход

$$y(t) = B(q(t)) = \begin{cases} 1, & \text{если } q(t) \in F, \\ 0, & \text{если } q(t) \notin F. \end{cases}$$

Слово $x \in X^*$ *допустимо* (определимо) автоматом $A = (X,Q,q_0,T,F)$, если $T(q_0,x) \in F$. Если $T(q_0,x) \notin F$, то слово x не допускается (отвергается) автоматом A.

Поведение $Beh(A)$ автомата A есть множество всех слов в алфавите X, допустимых автоматом A. *Язык* $L \subseteq X^*$ (автоматно) *определим* (допустим, представим), если существует автомат A без выхода, для которого $L = Beh(A)$.

Иногда автомат $A = (X,Q,q_0,T)$ задается без выхода и без множества выделенных состояний, иногда и без указания начального состояния $A = (X,Q,T)$. Два автомата без выхода эквивалентны, если они представляют один и тот же входной язык.

Таблица 27.3

	q_0	1 $(q_0,0)$	2 $(q_0,1)$	0 $(q_1,0)$	2 $(q_1,1)$	1 $(q_2,0)$	2 $(q_2,1)$
0		(q_0,q)	$(q_2,0)$	$(q_2,0)$	$(q_2,0)$	$(q_2,0)$	$(q_1,0)$	$(q_1,0)$
1		$(q_0,1)$	$(q_1,1)$	$(q_1,1)$	$(q_2,1)$	$(q_2,1)$	$(q_2,1)$	$(q_2,1)$

Прямое (декартово) произведение автоматов $A' = (X,Q',q'_0,T')$ и $A'' = (X,Q'',q''_0,T'')$ есть автомат $A' \times A'' = (X,Q' \times Q'',(q'_0,q''_0),T)$, где $T((q',q''),a) = (T'(q',a),T''(q'',a))$, $q' \in Q'$, $q'' \in Q''$, $a \in X$.

Пример. Пусть граф-схемы автоматов A' и A'' (без выходов) приведены на рис.27.3 и 27.4. Граф-схема автомата $A' \times A''$ изображена на рис.27.5.

Теорема. Класс автоматно представимых языков замкнут относительно булевых операций (объединения, пересечения, дополнения).

Доказательство. Пусть автоматы $A' = (X,Q',q'_0,T',F')$ и $A'' = (X,Q'',q''_0,T'',F'')$ определяют языки $Beh(A')$ и $Beh(A'')$ соответственно.

Дополнение $X^* - Beh(A')$ определимо автоматом $(X, Q', q'_0, T', Q'-F')$.

Пересечение $Beh(A') \cap Beh(A'')$ определимо декартовым произведением $A' \times A''$ с множеством выделенных состояний $F' \times F''$.

Объединение $Beh(A') \cup Beh(A'')$ определимо декартовым произведением $A' \times A''$ с множеством выделенных состояний $F' \times Q'' \cup Q' \times F''$.

27.2. Источники

Источник есть объект $S = (X,Q,Q_0,D,F)$, где X есть входной алфавит; Q есть алфавит состояний, $Q_0 \subseteq Q$ есть множество начальных состояний, $D \subseteq Q \times X \times Q$ есть (недетерминированная) таблица переходов (здесь в качестве входного сигнала допускается пустой символ, обозначаемый $*$), $F \subseteq Q$ есть множество выделенных состояний.

Тройка (q,a,q') из D называется переходом источника.

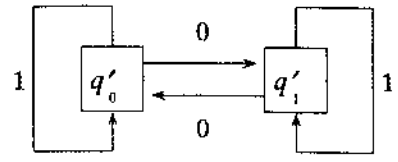


Рис.27.3

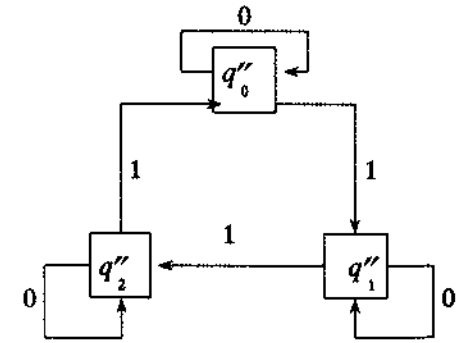


Рис.27.4

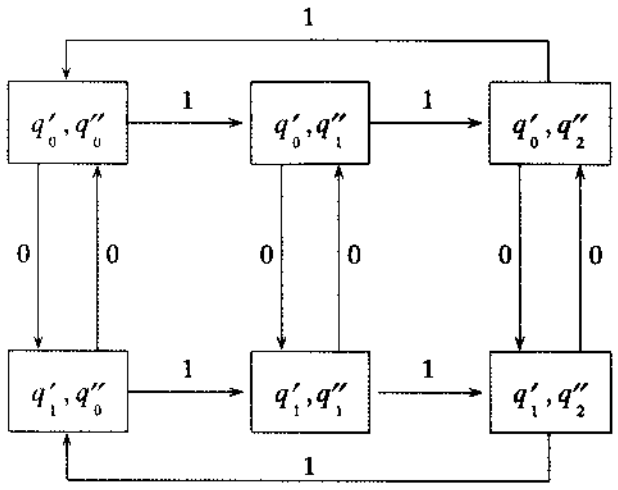


Рис.27.5

Пример. $X = \{0,1,2\}$; $Q = \{q_0,q_1,q_2,q_3\}$; $Q_0 = \{q_0,q_1\}$; $F = \{q_2,q_3\}$; $D = \{(q_0,0,q_1)$; $(q_0,0,q_3)$; $(q_0,1,q_1)$; $(q_0,*,q_0)$; $(q_0,*,q_1)$; $(q_0,*,q_3)$; $(q_1,*,q_1)$; $(q_1,2,q_2)$; $(q_2,0,q_2)$; $(q_2,0,q_3)$; $(q_2,1,q_2)$; $(q_2,2,q_2)$; $(q_3,1,q_2)$; $(q_3,1,q_3)$; $(q_3,*,q_0)\}$.

Граф-схема этого источника изображена на рис.27.6.

Заметим, что в случае источника некоторые дуги (стрелки) в его граф-схеме могут быть помечены пустым символом (т.е. дуги ничем не помечены).

Недетерминированный автомат есть частный случай источника, в котором нет дуг, помеченных пустым символом. Детерминированный автомат есть частный случай недетерминированного автомата, а потому и частный случай источника.

Входное слово $x = x(0)x(1)\ldots x(k)$ *допустимо* источником S, если существует последовательность $Q(0),Q(1),\ldots,Q(i),\ldots,Q(k),Q(k+1)$ (она называется *ходом* источника S на входном слове x; множество всех таких последовательностей обозначается через $Rn(S,x)$), где каждое $Q(i)$ есть конечная последовательность состояний $q(i,0)$, $q(i,1),\ldots,q(i,n_i)$, для которой:

$$q(0,0) \in Q_0;$$
$$(q(i,j),*,q(i,j+1)) \in D;\ i=0,1,\ldots,k;\ j=0,1,\ldots,n_i-1;$$
$$(q(i,n_i),x(i),q(i+1,0)) \in D;\ i=0,1,\ldots,k;\ q(k+1,n_{k+1}) \in F.$$

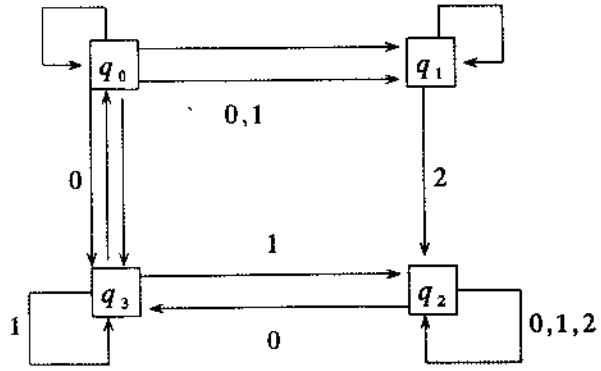


Рис.27.6

В противном случае слово x источник S не допускает (отвергает). Для примера источника из рис.27.6 на входном слове $x = 1\ 2\ 2\ 0\ 1\ 1\ 1\ 2$ возможна следующая последовательность состояний:

$x=x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$	$x(5)$	$x(6)$	$x(7)$	
1	2	2	0	1	1	1	2	
q_0	q_1	q_2	q_2	q_3	q_3	q_3	q_2	q_2
q_0	q_1			q_0		q_0		
q_3				q_0		q_3		
q_0				q_3				
				q_0				
				q_3				

$Q(0)$	$Q(1)$	$Q(2)$	$Q(3)$	$Q(4)$	$Q(5)$	$Q(6)$	$Q(7)$	$Q(8)$
$q(0,0)$	$q(1,0)$	$q(2,0)$	$q(3,0)$	$q(4,0)$	$q(5,0)$	$q(6,0)$	$q(7,0)$	$q(8,0)$
$q(0,1)$	$q(1,1)$			$q(4,1)$		$q(6,1)$		
$q(0,2)$				$q(4,2)$		$q(6,2)$		
$q(0,3)$				$q(4,3)$				
				$q(4,4)$				
				$q(4,5)$				

Так как $q_2 \in F$, то слово x допустимо источником S.

Поведение $Beh(S)$ источника S есть множество всех слов, допустимых источником S. Два источника *эквивалентны*, если они имеют одинаковые поведения (т.е. если они допускают одно и то же множество входных слов). Язык $L \subseteq X^*$ *представим* (доустим, определим) источником, если существует источник S, для которого $L=Beh(S)$. Множество $Q' \subseteq Q$ состояний источника $S = (X,Q,Q_0,D,F)$ *замкнуто*, если $\forall q \in Q\ \forall q' \in Q'((q',*,q) \in D \rightarrow$

$q \in Q'$). *Замыкание* $[Q']$ множества Q' есть наименьшее замкнутое множество состояний, содержащее Q'.

Для источника из рис.16.6 $[\{q_0\}] = \{q_0,q_1,q_3\}$; $[\{q_1\}] = \{q_1\}$; $[\{q_2,q_3\}] = \{q_0,q_1,q_2,q_3\}$; $[\{q_2\}] = \{q_2\}$.

Теорема (о детерминизации источника). Пусть язык $L \subseteq X^*$. Следующие утверждения равносильны.

1. Язык L представим конечным автоматом.
2. Язык L представим источником.

Доказательство. Следование 1 → 2 тривиально, ибо всякий автомат есть частный случай источника.

Покажем, что 2 → 1. Пусть язык L представим источником $S = (X,Q,Q_0,D,F)$. Построим автомат $A = (X,S,s_0,T,H)$, где $S \subseteq P(Q)$ есть множество всех замкнутых подмножеств множества Q (включая пустое подмножество); начальное состояние $s_0=[Q_0]$; функция переходов $T: S \times X \to S$ такова, что $T(s,a) = [\{q \in Q: \exists q' \in s\ (q',a,q) \in D\}]$; множество выделенных состояний $U = \{s \in S: \exists q \in s (q \in F)\}$.

Утверждение. $Beh(S) = Beh(A)$.

Доказательство. Пусть слово $x = x(0)x(1)\ldots x(k) \in Beh(A)$. Тогда существует последовательность состояний $s(0),s(1),\ldots,s(k),s(k+1)$ автомата A (рис.27.7), для которой

$s(0) \in s_0$;
$T(s(j),x(j)) = s(j+1)$, $j=0,1,\ldots,k$;
$s(k+1) \in U$.

Состояние $s(k+1) = [\{q \in Q: \exists q' \in s(k)((q',a,q) \in D\}] = s'$.

Так как $s(k+1) \in U$, то существует состояние $q' \in s(k+1)$, причем $q' \in F$. Возможны следующие случаи:

$q' \in s'$. Тогда $\exists q \in s(k)\ ((q,x(k),q') \in D)$;

$q' \in [s']$, $q' \notin s'$, т.е. $q' \in [s']-s'$.

Тогда существует последовательность состояний $Q(k+1)$, начинающаяся с некоторого состояния $q(k+1,0) \in s'$ и заканчивающаяся состоянием $q(k+1,n) \in [s']$, причем $(q(k+1,j),*,q(k+1,j+1)) \in D$, $j = 0,1,\ldots,n_{k+1}-1$. Найдется состояние $q \in s(k)$, для которого $(q,x(k),q(k+1,0)) \in D$. Продвигаясь к началу последовательности состояний автомата A, построим последовательность состояний источника

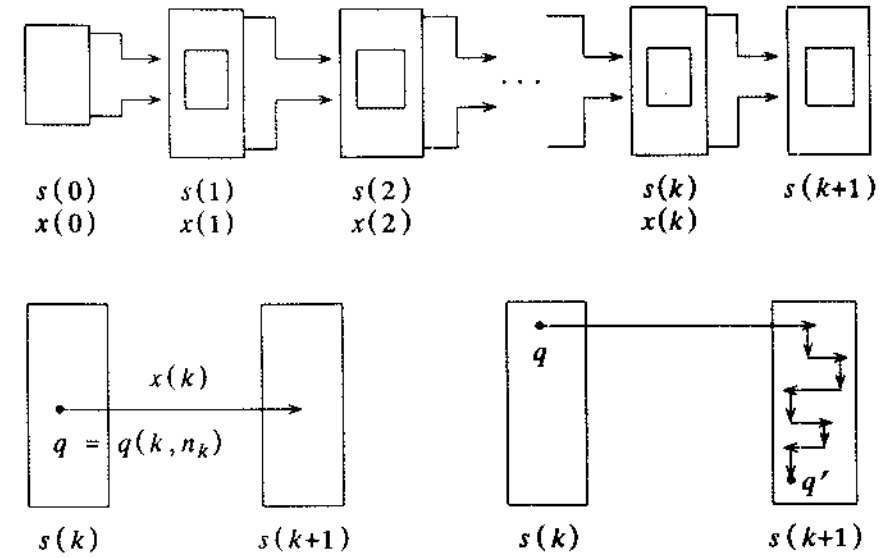


Рис.27.7

$$Q(0),Q(1),\ldots,Q(k),Q(k+1), \qquad (27.1)$$

для которой:

$$\begin{gathered} q(0,0) \in Q_0; \\ (q(i,n\ ),x(i),q(i+1,0)) \in D,\ i=0,1,\ldots,k; \\ (q(i,j),*,q(i,j+1)) \in D,\ i=0,1,\ldots,k+1; \\ j=0,1,\ldots,n_i-1; \\ q(k+1,n_{k+1}) \in F. \end{gathered} \qquad (27.2)$$

Поэтому слово $x \in Beh(S)$ и, следовательно, $Beh(A) = Beh(S)$.

Пусть теперь слово $x = x(0)x(1)\ldots x(k) \in Beh(S)$. Тогда существует последовательность состояний (27.1), для которой выполняются условия (27.2). Для соответствующей последовательности $s(0),s(1),\ldots,s(k+1)$ автомата A состояние $q(k+1, n_{k+1}) \in s(k+1)$, и потому $s(k+1) \in U$. Следовательно, слово $x \in Beh(A)$. Так что $Beh(S) \subseteq Beh(A)$, что вместе с $Beh(A) \subseteq Beh(S)$ дает $Beh(A) = Beh(S)$.

Утверждение установлено. Теорема доказана.

Следствие. Класс множеств, представимых источниками, замкнут относительно булевых операций.

Заметим, что если язык L в алфавите X автоматно предста-

вим, то язык L автоматно представим над любым расширением X' алфавита X, ибо автомат, представляющий язык L над X, можно рассматривать как источник, представляющий язык L над X'.

27.2.1. Алгоритм детерминизации источника

Пусть источник $S = (X,Q,Q_0,D,F)$. Возьмем $Q' \subseteq Q$, $a \in X$. Пусть $S(Q',a) = \{q \in Q: \exists q' \in Q'\ ((q',a,q) \in D\}$ есть множество всех состояний, в которые источник S переходит из состояний множества Q' под воздействием входной непустой буквы a из X. Автомат A с тем же поведением, что и источник S, строим следующим образом.

1. Формируем замыкание множества начальных состояний источника и объявляем это замыкание начальным состоянием конструируемого автомата.

2. Если состояние $s = Q' \subseteq Q$ автомата A уже построено, то $T(s,a) = [S(Q',a)]$ есть состояние, в которое перейдет автомат A из состояния s под воздействием буквы a.

3. Применяем п.2 алгоритма до тех пор, пока его применение порождает новые состояния автомата A.

4. Объявляем выделенными те состояния $s=Q' \subseteq Q$ автомата A, которые содержат в себе выделенные состояния источника S.

Пример. $X = \{0,1\}$; $Q = \{q_0,q_1,q_2\}$; $Q_0 = \{q_0\}$; $F = \{q_0,q_2\}$. Таблица переходов D источника $S = (X,Q,\{q_0\},D,F)$ изображена на рис.27.8. Таблица переходов детерминированного автомата A, эквивалентного источнику S, приведена в табл.27.4. Выделенные состояния автомата A помечены звездочками.

Замечание. Иногда будем задавать источник в виде $S = (X, Q,D)$ без указания множеств начальных и выделенных состояний. Пусть Q',Q'',Q''' есть подмножества из Q; $L(S,Q',Q''')$ есть поведение источника $S = (X,Q,Q',D,Q''')$, а язык $L_1 = L(S,Q',Q'', Q''')$ есть множество всех входных слов x, для которых существует ход из $Rn(S,x)$, в котором все состояния лежат в множестве Q''. Язык L_1 представим источником, который получается из источника S удалением всех ребер, инцидентных на граф-схеме источника S, состояниям вне Q''.

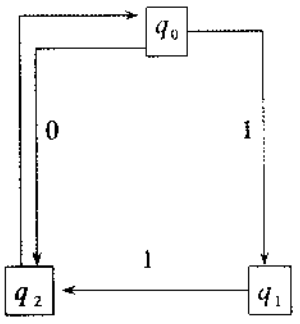


Рис.27.8

Таблица 27.4

	$\{q_0\}^*$	$\{q_1\}$	$\{q_0,q_2\}^*$	$\varnothing$
0	$\{q_0,q_2\}$	$\varnothing$	$\{q_0,q_2\}$	$\varnothing$
1	$\{q_1\}$	$\{q_0,q_2\}$	$\{q_1\}$	$\varnothing$

27.3. Регулярные языки

Пусть X есть произвольный конечный алфавит.

Конкатенация (умножение) двух языков L_1 и L_2 из X^* есть множество $L_1 \cdot L_2 = \{z \in X^*: \exists x,y \in X^*(x \in L_1, y \in L_2, z = x \cdot y)\}$. *Степень* k языка L есть множество $L^k = L \cdot L \cdot \ldots \cdot L$, k раз. Примем по определению, что $L^0 = \{*\}$. Справедливо равенство $L \cdot \varnothing = \varnothing \cdot L = \varnothing$. *Итерация* языка L из X^* есть множество $L^* = L^1 \cup L^2 \cup L^3 \cup L^4 \cup \ldots$ *Операции Клини* есть операции конкатенации, объединения, итерации.

Примеры. 1. $\{0,011.00\} \cdot \{11,101\} = \{011,0101,01111, 011101,0011,00101\}$.

2. $\{0\}^* = \{0,00,000,\ldots\}$.

3. $\{0,1\}^*$ есть множество всех непустых слов в алфавите $\{0,1\}$.

4. $\{0,1\}^k$ есть множество всех слов длины k в алфавите $\{0,1\}$.

Определение (регулярного языка). Одноэлементные языки из X регулярны и имеют глубину построения 1. Если языки L,L_1,L_2 глубины построения k,k_1,k_2 соответственно регулярны, то языки $(L_1 \cup L_2)$, $(L_1 \cdot L_2)$, (L^*) регулярны и имеют глубину построения $\max(k_1,k_2)+1$, $\max(k_1,k_2)+1$, $k+1$ соответственно.

Будем опускать скобки при записи регулярного языка, принимая следующий приоритет операций: $*$, $\cdot$, $\cup$; для одноэлементного языка вместо $\{a\}$ писать и просто a; вместо $\{a,b,\ldots, c\}$ писать иногда $a \vee b \vee \ldots \vee c$; вместо объединения $\cup$ при записи регулярных языков использовать также знак $\vee$.

Теорема. Класс языков, представимых источниками, замкнут относительно операций Клини.

Доказательство. Пусть языки L_1 и L_2 представимы источниками $S_1 = (X,Q',Q'_0,D',F')$ и $S_2 = (X,Q'',Q''_0,D'',F'')$ соответственно.

Объединение $L_1 \cup L_2$ представимо источником $S = (X, Q' \cup Q'', Q'_0 \cup Q''_0, D' \cup D'', F' \cup F'')$.

Конкатенация $L_1 \cdot L_2$ представима источником $S = (X, Q' \cup Q'',$

Q_0', D, F''), где $D = D' \cup D'' \cup \{(q', *, q'') : q' \in F', q'' \in Q_0''\}$.

Итерация L_1^* представима источником $S = (X, Q', Q_0, D, F')$, где $D = D' \cup \{(q, *, q') : q \in F', q' \in Q_0'\}$.

Граф-схемы источников S_1 и S_2 при объединении объединяются. Начальные состояния для S_1 и S_2 становятся начальными состояниями для S, а выделенные состояния для S_1 и S_2 являются выделенными состояниями для S.

Граф-схемы источников S_1 и S_2 при конкатенации объединяются. Добавляются пустые (т.е. ничем не помеченные) стрелки, ведущие из выделенных состояний для S_1 в начальные состояния для S_2. Начальные состояния для S_1 являются начальными состояниями для S. Выделенные состояния для S_2 являются выделенными состояниями для S.

При итерации в граф-схеме для S_1 добавляются пустые стрелки из выделенных состояний для S_1 в его начальные состояния.

Следствие. Класс языков, представимых автоматами, замкнут относительно операций Клини.

Теорема (синтеза). Каждый регулярный язык представим некоторым источником.

Доказательство. Индукция по глубине k построения языка.

Базис. $k = 1$. Одноэлементные языки автоматно представимы; слеовательно, они представимы и источниками.

Предположение индукции. Предположим, что все регулярные языки с глубиной построения меньше k представимы источниками.

Шаг индукции. Покажем, что все регулярные языки с глубиной построения k представимы источниками. Пусть регулярный язык L имеет глубину построения k. Тогда L имеет один из видов: $L_1 \cup L_2$, $L_1 \cdot L_2$, L_1^*. Так как регулярные языки L_1, L_2 имеют глубину построения меньше k, то по предположению индукции они представимы некоторыми источниками. Тогда язык L представим источником по теореме о замкнутости класса языков, представимых источниками, относительно операций Клини.

Следствие. Каждый регулярный язык автоматно представим.

Теорема (анализа). Всякий язык, представимый источником, регулярен.

Доказательство. Так как классы языков, представимых источниками и автоматами, совпадают, то теорему достаточно доказать для автоматов. Пусть автомат $A = (X, Q, q_0, T, F)$, где $Q = \{q_0, q_1, \ldots, q_n\}$, $F = \{q_{i_1}, \ldots, q_{i_m}\}$, представляют язык $L = Beh(A)$.

Построим множество входных слов $E_{ij}^{(k)}$ с помощью следующей индуктивной процедуры:

$$E_{ij}^{(0)} = \{a \in X : T(q_i, a) = q_j\};$$

$$E_{ij}^{(k)} = E_{ij}^{(k-1)} \cup E_{ik}^{(k-1)} \cdot (E_{kk}^{(k-1)})^* \cdot E_{kj}^{(k-1)}, \quad k = 0, 1, \ldots, n.$$

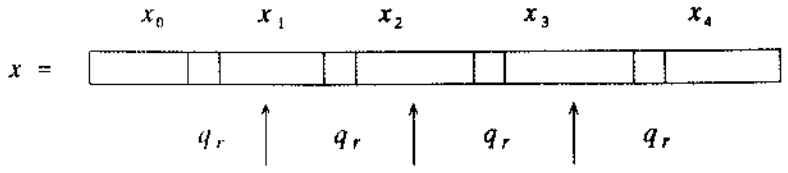
Если автомат A при подаче входного слова $x = x(0)x(1)\ldots x(r-1)$ проходит последовательность состояний $q(0) = q_i, q(1), q(2), \ldots, q(r-1), q(r) = q_j$, то скажем тогда, что x переводит A из q_i в q_j, и автомат A проходит при этом через состояния $q(1), q(2), \ldots, q(r-1)$. Индукцией по k можно показать, что множество входных слов $E_{ij}^{(k)}$ переводят A из q_i в q_j, и не проводят A через состояния q_ν с номерами $\nu > k$. Пусть $M_{ij}^{(k)}$ есть множество всех входных слов, которые переводят A из q_i в q_j, и при этом не проводят A через состояния q_ν с номерами $\nu > k$.

Утверждение. $E_{ij}^{(k)} = M_{ij}^{(k)}$.

Доказательство. Индукция по k. Для $k = 0$ утверждение очевидно. Пусть оно верно для $k = r-1$, т.е. $E_{ij}^{(r-1)} = M_{ij}^{(r-1)}$. Покажем, что оно верно для $k = r$. Пусть $\hat{M}_{ij}^{(r)}$ (или $\hat{E}_{ij}^{(r)}$ есть множество всех входных слов из $M_{ij}^{(r)}$ (или из $E_{ij}^{(r)}$), которые проводят A через состояние q_r. Ясно, что $\hat{E}_{ij}^{(r)} = E_{ir}^{(r-1)} \cdot (E_{rr}^{(r-1)}) \cdot E_{rj}^{(r-1)}$. Множество

$$M_{ij}^{(r)} = M_{ij}^{(r-1)} \cup \hat{M}_{ij}^{(r)} = E_{ij}^{(r-1)} \cup \hat{M}_{ij}^{(r)}.$$

Покажем, что $\hat{M}_{ij}^{(r)} = \hat{E}_{ij}^{(r)}$. Если входное слово $x \in \hat{M}_{ij}^{(r)}$ проводит A через q_r s раз, то x представимо в виде $x = x_0 x_1 \ldots x_{s-1} x_s$, (рис.27.9). Тогда



Здесь состояний q_r нет

Рис.27.9

$x_0 \in M_{ir}^{(r-1)} = E_{ir}^{(r-1)}$; $x_s \in M_{rj}^{(r-1)} = E_{rj}^{(r-1)}$;

$x_i \in M_{rr}^{(r-1)} = E_{rr}^{(r-1)}$, $i = 1,2,\ldots,s-1$;

$x_1 \cdot x_2 \cdot \ldots \cdot x_{s-1} \in (E_{rr}^{(r-1)})^*$. Поэтому

$x \in E_{ir} \cdot (E_{rr}^{(r-1)})^* \cdot E_{rj}^{(r-1)} = \hat{E}_{ij}^{(r)}$. Отсюда $\hat{M}_{ij}^{(r)} = \hat{E}_{ij}^{(r)}$.

Если $x \in \hat{E}_{ij}^{(r)}$, то слово x переводит A из q_i в q_j, проводит A через q_r, не проводит A через состояния q_v с номерами $v > r$, т.е. $x \in \hat{M}_{ij}^{(r)}$. Поэтому $\hat{E}_{ij}^{(r)} = \hat{M}_{ij}^{(r)}$. Равен-ство $\hat{M}_{ij}^{(r)} = \hat{E}_{ij}^{(r)}$ доказано. Тогда

$$M_{ij}^{(r)} = E_{ij}^{(r-1)} \cup \hat{M}_{ij}^{(r)} = E_{ij}^{(r-1)} \cup E_{ij}^{(r)} = E_{ij}^{(r)}.$$

Утверждение доказано. Продолжим доказательство теоремы. Исходный автомат A имеет $n+1$ состояний $\{q_0,q_1,\ldots,q_n\}$.

Язык $L = E_{0,i_1}^{(n)} \cup \ldots \cup E_{0,i_m}^{(n)}$ регулярен согласно утверждению о регулярности всех дизъюнктивных слагаемых.

Теорема доказана.

Замечание. Если пустое слово включается в состав рассматриваемых языков, то итерация $L^* = L^0 \cup L^1 \cup \ldots$, где $L^0 = \{*\}$. Доказанные выше теоремы останутся справедливыми. Следует только слегка изменить их доказательства, сделав поправку на пустое слово.

27.4. Теоремы замкнутости для класса автоматно представимых языков

Пусть X есть конечный алфавит. *Обращение* слова $x=x(0)x(1)\ldots x(k-1)x(k)$ из X^* есть слово $x^{-1} = x(k)x(k-1)\ldots x(1)x(0)$. Если множество $M \subseteq X^*$, то $M^{-1} = \{x^{-1} : x \in M\}$.

Теорема. Класс языков, представимых источниками, замкнут относительно операции обращения.

Доказательство. Пусть язык M представим источником $S = (X,Q,Q_0,D,F)$. Тогда язык M^{-1} представим источником $S' = (X, Q,F,D',Q_0)$, где $D' = \{(q',a,q) : (q,a,q') \in D\}$, т.е. в граф-схеме источника S все стрелки меняют свое направление на противоположное.

Следствие. Класс автоматно представимых языков замкнут относительно операции обращения.

Пусть $X = \{a_0,a_1,\ldots a_k\}$, $Y = \{b_0,b_1,\ldots,b_l\}$ есть два

конечных алфавита, и пусть функция $f: X \to Y$ осуществляет *проекцию* с одного алфавита на другой (т.е. с алфавита X на алфавит Y). Пусть $x = x(0)x(1)\ldots x(r)$ есть слово в алфавите X. Тогда слово $f(x) = f(x(0))f(x(1))\ldots f(x(r))$ есть проекция слова x при отображении f. Если $M \subseteq X^*$, то $f(M) = \{f(x) : x \in M\}$ есть проекция множества M при отображении f.

Теорема. Класс языков, представимых источниками, замкнут относительно проекции.

Доказательство. Пусть язык M представим источником $S = (X,Q,Q_0,D,F)$, и функция $f: X \to Y$ осуществляет проекцию с алфавита X на алфавит Y. Тогда язык $f(M)$ представим источником $S' = (f(X), Q,Q_0,D',F)$, где $D' = \{(q,b,q') : \exists a \in X\ (f(a) = b\ \&\ (q,a,q') \in D\}$, т.е. в граф-схеме источника S всякая пометка a из X заменяется на пометку $f(a)$ из Y.

Следствие. Класс автоматно представимых языков замкнут относительно проекции.

Пусть $x = x(0)x(1)\ldots x(k)$ есть некоторое слово в алфавите X, причем слово x содержит букву a. *Аннулирование* буквы a в слове x есть удаление в слове x буквы a всюду, где она встречается. Обозначим эту операцию через $An(x,a)$. Пусть $M \subseteq X^*$. Пусть тогда $An(M,a) = \{An(x,a) : x \in M\}$.

Теорема. Класс языков, представимых источниками, замкнут относительно операции аннулирования.

Доказательство. Пусть язык M представим источником $S = (X,Q, Q_0,D,F)$. Тогда язык $An(M,a)$ представим источником $S' = (X,Q,Q_0,D',F)$, где $D' = \{(q,b,q') \in D : b \neq a\} \cup \{(q,*,q') : (q,a,q') \in D\}$, т.е. в граф-схеме источника S удаляются буквы a всюду, где они встречаются, но сами стрелки, которые были помечены буквой a, остаются.

Следствие. Класс автоматно представимых языков замкнут относительно операции аннулирования.

Пусть $x = x(0)x(1)\ldots x(k)aa\ldots a$ при $x(k) \neq a$ есть слово в алфавите X, содержащем букву a. Тогда операция *усечения* слова x по букве a (обозначение: $Tranc(x,a)$) определяется как $Tranc(x,a) = x(0)x(1)\ldots x(k)$. Если $M \subseteq X^*$, то множество $Tranc(M,a) = \{Tranc(x, a) : x \in M\}$.

Теорема. Класс языков, представимых источниками, замкнут относительно операции усечения.

Доказательство. Пусть язык M представим источником $S = (X,Q,Q_0,D,F)$. Пусть $a \in X$ и $a^k = aaa\ldots a$, k раз. Построим источник $S' = (X,Q,Q_0,D,G)$, где $G = \{q \in Q : \exists k\, \exists q' \in F\ ((q,$

$a^k,q') \in D\} \cup F$, т.е. G есть множество всех тех состояний $q \in Q$, для которых существует слово a^k при некотором натуральном k, переводящее источник S из состояния q в состояние q', при этом $q' \in F$. Источник S' с поведением $M' = Beh(S')$ допускает все те слова, которые можно продолжить буквами a до слова, допустимого источником S, а также все слова, допустимые источником S.

Пусть S'' есть источник, допускающий множество M'' всех слов в алфавите X, не заканчивающихся на букву a. Источник, допускающий множество $M' \cap M''$ искомый.

Следствие. Класс автоматно представимых языков замкнут относительно операции усечения.

Пусть X и Y есть два алфавита и множество $M \subseteq X^*$. *Цилиндр, восставленный* из множества M по второй оси, есть множество $Cyl_2M = \{(x,y) \in (X \times Y)^* : x \in M\}$.

Теорема. Класс языков, представимых источниками, замкнут относительно операции восстановления цилиндра.

Доказательство. Пусть язык $M \subseteq X^*$ представим источником $S = (X,Q,Q_0,D,F)$ и Y есть другой алфавит. Тогда множество Cyl_2M представимо источником $S' = (X \times Y,Q,Q_0,D',F)$, где $D' = \{(q,(a,b),q') : (q,a,q') \in D$ и $b \in Y\}$, т.е. в граф-схеме источника S все ребра, помеченные символом a из X, помечаем символами (a,b), где $b \in Y$.

Следствие. Класс автоматно представимых языков замкнут относительно операции восстановления цилиндра.

Операцию восстановления цилиндра можно задать и по первой оси. Эту операцию можно распространить на случай восстановления цилиндра по любому количеству осей.

Теорема. Алгоритмически разрешимы следующие свойства конечных автоматов.

1. Представимый автоматом язык пуст (проблема пустоты).
2. Представимый автоматом язык бесконечен.

Доказательство. Пусть $A = (X,Q,q_0,T,F)$ есть детерминированный конечный автомат.

1. Представимый автоматом язык пуст тогда и только тогда, когда A не допускает ни одного слова длины не более мощности алфавита X.

2. Представимый автоматом A язык бесконечен тогда и только тогда, когда автомат A допускает слово $x = x_0x_1x_2$, где $x_i \in X^*$, $i = 0,1,2$, причем для одного из $q \in Q$:

x_0 переводит q_0 в q;

x_1 переводит q в q;
x_2 переводит q в выделенное состояние.

Последние три условия проверяются по п.1.

27.5. Минимизация числа состояний автомата с выходом

Будем рассматривать автоматы с общим входным и общим выходным алфавитами. Два автомата *эквивалентны*, если они реализуют один и тот же оператор.

Пусть $A = (X,Y,Q,q_0,T,B)$ есть автомат с выходом, реализующий оператор $\mathbb{O} : X^* \to Y^*$.

Два состояния q', q'' автомата A *r-отличимы*, если существует входное слово $x = x(0)x(1)\ldots x(r-1)$ длины r, для которого $B(q',x) = y' = y'' = B(q'',x)$. Состояния q,q' *отличимы*, если они r-отличимы при некотором r. Состояния q',q'' *r-неотличимы*, если для любого входного слова x длины не более r $B(q',x) = B(q'',x)$. Состояния q',q'' *строго r-отличимы*, если они r-отличимы, но неотличимы никаким входным словом длины меньше r (т.е. $(r-1)$-неотличимы). Состояния q' и q'' *неотличимы*, если они r-неотличимы при любом r.

Аналогичные определения отличимости состояний q' и q'' можно ввести для двух автоматов A и B, причем q' есть состояние автомата A, а q'' автомата B.

Автоматы A и B, реализующие операторы $\mathbb{O}_A$ и $\mathbb{O}_B$, *r-отличимы*, если существует входное слово x длины r, для которого $\mathbb{O}_A(x) = \mathbb{O}_B(x)$. Автоматы A и B *отличимы*, если они отличимы при некотором r. Автоматы A и B *r-неотличимы*, если $\mathbb{O}_A(x) = \mathbb{O}_B(x)$ для всякого входного слова x длины не более r. Автоматы A и B *строго r-отличимы*, если они r-отличимы и $(r-1)$-неотличимы. Автоматы A и B *неотличимы*, если они r-неотличимы при любом r.

Ясно, что если автоматы A и B неотличимы, то они реализуют одинаковые операторы и, следовательно, эквивалентны.

Состояние q'' *достижимо* из состояния q' автомата A, если $T(q',x) = q''$ для некоторого входного слова x.

Автомат A называется *приведенным*, если все его состояния достижимы из начального состояния и попарно отличимы.

Автомат A называется *минимальным*, если он имеет наименьшее число состояний среди всех автоматов, эквивалентных автомату A.

Теорема. Приведенный автомат является минимальным.

Доказательство. Допустим противное: приведенный автомат A

$= (X,Y,Q,q_0,T,B)$ не является минимальным, а минимальным является автомат $A' = (X,Y,Q',q'_0,T',B')$, эквивалентный автомату A, но с числом состояний меньше, чем в A. Пусть $Q' = \{q'_0, q'_1,\ldots,q'_m\}$. Пусть входные слова $x_0,x_1,\ldots,x_m$ переводят автомат A' из начального состояния q'_0 в состояния $q'_0,q'_1,\ldots,q'_m$, а автомат A в состояния $q_{i0},q_{i1},\ldots,q_{im}$ соответственно. Так как автоматы A и B эквивалентны, то неразличимы пары состояний

$$q'_0,\ q_{i0};\quad q'_1,\ q_{i1};\quad \ldots\quad q'_m,\ q_{im}. \qquad (27.3)$$

Пусть состояние $q \notin \{q_{i0},q_{i1},\ldots,q_{im}\} = S$, но $q \in Q$. Так как A есть приведенный автомат, то состояние q отличимо от каждого состояния из Q. Так как q достижимо из q_0, то для некоторого входного слова x $T(q_0,x) = q$. Пусть $T'(q'_0,x) = q'$ при некотором j. Так как автоматы A и B эквивалентны, то состояния q и q' неотличимы. Пусть состояние q' соответствует согласно (27.3) состоянию $q_{ij} \in S$. По (27.3) состояние q'_j неотличимо от q_{ij}, но q'_j неотличимо от q и потому q неотличимо от q_{ij}. Противоречие. Следовательно, автомат A минимален.

Теорема. Если два состояния q, q' автомата A с k состояниями отличимы, то они отличимы входным словом длины не более k^2 (т.е. k^2-отличимы).

Доказательство. Пусть состояния q, q' отличимы. Тогда они строго r-отличимы при некотором r, т.е. найдется входное слово $x = x(0)x(1)\ldots x(r-1)$, для которого автомат A, начав работу над словом x в состояниях q и q', перерабатывает его в различные слова $y = y(0)y(1)\ldots y(r-1)$ и $y' = y'(0)y'(1)\ldots y'(r-1)$ (т.е. $B(q,x) = y$, $B(q',x) = y'$) и проходит последовательность состояний

$$q(0),\ q(1)\ ,\ldots,q(r); \qquad q(0) = q;$$
$$q'(0),q'(1),\ldots,q'(r); \qquad q'(0) = q'$$

соответственно. При этом две последние буквы в словах y и y' различны (т.е. $y(r) \neq y'(r)$). Для любых (дискретных) моментов времени t_1,t_2 $(0 \leqslant t_1 < t_2 \leqslant r-1)$ пары (столбцы)

$$q(t_1)\ ,\ q(t_1+1)\ ,\ldots\ ,\ q(t_2)$$
$$q'(t_1),\ q'(t_1+1),\ldots\ ,\ q'(t_2)$$

попарно различны, ибо при $q(t_1) = q(t_2)$, $q'(t_1) = q'(t_2)$, выбросив из слова $x = x(0)x(1)\ldots x(t_1)\ldots x(t_2)\ldots x(r-1)$ часть $x(t_1)\ x(t_1+1)\ldots x(t_2-1)$, получим более короткое слово

$x(0)x(1)\ldots x(t_1-1)x(t_2)\ldots x(r-1)$, перерабатываемое автоматом A в различные слова

$$y(0)\ y(1)\ \ldots\ y(t_1-1)\ y(t_2)\ \ldots\ \ y(r-1),$$
$$y'(0)y'(1)\ldots\ y'(t_1-1)y'(t_2)\ \ldots\ y'(r-1),$$

в противоречие со строгой r-отличимостью состояний q и q'. Так как число различных пар состояний равно k^2, то длина входного слова x не превосходит k^2.

Следствие. Если два состояния автомата A с k состояниями k^2-неотличимы, то они вообще неотличимы.

Доказанная теорема позволяет разбить множество состояний автомата A с k состояниями на попарно непересекающиеся классы неотличимых состояний. Использование этой теоремы при больших k затруднительно, ибо необходимо проверить все состояния на отличимость для всех входных слов длины k^2. Поэтому сначала упрощают автомат, исходя из особенностей задания его функций переходов и выходов, а затем уже применяют доказанную теорему.

27.5.1. Склеивание неразличимых состояний

Пусть задан автомат $A = (X,Y,Q,q_0,T,B)$ с выходом. Разобьем множество его состояний на попарно непересекающиеся классы неотличимых состояний: $Q = Q_0 \cup Q_1 \cup \ldots \cup Q_p$, $Q_i \cap Q_j = \varnothing$; $i,j = 0, 1,\ldots,p$; $i \neq j$. Для всякого i множество Q_i содержит попарно неотличимые состояния. Пусть q_i есть какой-либо представитель из Q_i. Построим автомат $A' = (X,Y,Q',Q_0, T',B')$, где $Q' = \{Q_0,Q_1,\ldots,Q_p\}$; Q'_0 есть тот класс неотличимых состояний, который содержит q_0; $T'(Q_i,a) = \{T(q,a) : q \in Q_i\}$; $B'(Q_i,a) = B(q,a)$, $q \in Q_i$. Автоматы A и A' эквивалентны. В самом деле, для всякого входного слова $x = x(0)x(1)\ldots x(r)$, если автомат A перерабатывает его в выходное слово $y = y(0)y(1)\ldots y(r)$, проходя последовательность состояний $q(0),q(1),\ldots,q(r),q(r+1)$ при $q(0) = q_0$, то автомат A', проходя последовательность состояний $Q(0),Q(1),\ldots,Q(r),Q(r+1)$ при $Q(0) = Q_0$, выдает выходное слово $y' = y'(0)y'(1)\ldots y'(r)$, причем $y'(t) = B(Q(t),x(t)) = B(q(t),x(t)) = y(t)$, где $q(t) \in Q(t)$, т.е. $y' = y$.

Состояния автомата A' попарно отличимы, ибо если бы некоторые Q_i, Q_j были неотличимы, то их объединили бы на этапе построения автомата A в один класс.

27.5.2. Алгоритм минимизации автомата

1. Если автомат $A = (X,Y,Q,q_0,T,B)$ имеет состояния, недостижимые из начального состояния q_0, то построим автомат $A' = (X,Y,Q',q_0,T',B')$, эквивалентный A, где Q' есть множество состояний автомата A, достижимых из начального состояния q_0; функции T' и B' есть ограничения функций T и B на множество Q'. Если все состояния в Q достижимы из q_0, то A' есть A.

2. Построим классы неотличимости состояний автомата A' и проведем операции их склеивания. Получим автомат A'', эквивалентный автомату A'.

3. Автомат A'' приведенный и, следовательно, минимальный.

Пример 1. Пусть автомат A задан в табл.27.5.

Множество столбцов, помеченных состояниями 1,6 и 3,4, одинаковы; эти пары состояний заведомо склеиваются. В результате склеивания указанных пар состояний получаем автомат с шестью состояниями (табл.27.6). Одинаковых столбцов в табл. 27.6 больше нет.

Строим теперь разбиение множества состояний автомата, приведенного в табл.27.6, на попарно непересекающиеся классы неотличимых состояний. Классы эти строятся последовательно: сначала классы 1-неотличимых состояний, затем 2-неотличимых, 3-неотличимых, и так далее.

Исходим из множества состояний {0,1,2,3,5,7} автомата из табл.27.6. Запишем это множество матрицей-строкой M_0 = 0123 57. Подаем на вход автомата слово x = 0. Автомат переходит в состояния 023173 и выдает на выходе 001110. Состояния 012357 разбиваются на два класса {0,1,7} и {2,3,5} состояний, неотличимых словом x = 0. Запишем эти два множества строкой 017.235, разделив точкой два полученных класса. Слову x = 0 сопоставим матрицу M_1, приведенную на рис.27.10.

Исходим теперь из последней строки состояний 017.235 в M_1. На вход автомата подаем слово x = 1. Автомат переходит в состояния 175231 и выдает на выходе 111000. Классы состояний 017 и 235 словом x = 1 неотличимы. Слову x = 1 ставим в соответствие матрицу M_2 (рис.27.10).

Таблица 27.5

	0	1	2	3	4	5	6	7
0	0,0	2,0	3,1	1,1	1,1	7,1	2,0	4,0
1	1,1	7,1	2,0	4,0	4,0	6,0	7,1	5,1

Таблица 27.6

	0	1	2	3	5	7
0	0,0	2,0	3,1	1,1	7,1	3,0
1	1,1	7,1	2,0	3,0	1,0	5,1

0 и 1 исчерпывают все слова длины 1, поэтому состояния классов 017 и 235 1-неотличимы.

Исходим из последней строки 017.235 в M_2. На вход автомата подаем слово x = 00. Автомат перейдет в состояния 031123 и выдаст на выходе 011100. Класс 017 разбивается на два класса 0.17, а класс 235 на два класса 2.35 состояний. Запишем эти множества строкой 0.17.2.35, разделив точками полученные классы. Сопоставим слову x = 00 матрицу M_3 (рис.16.10).

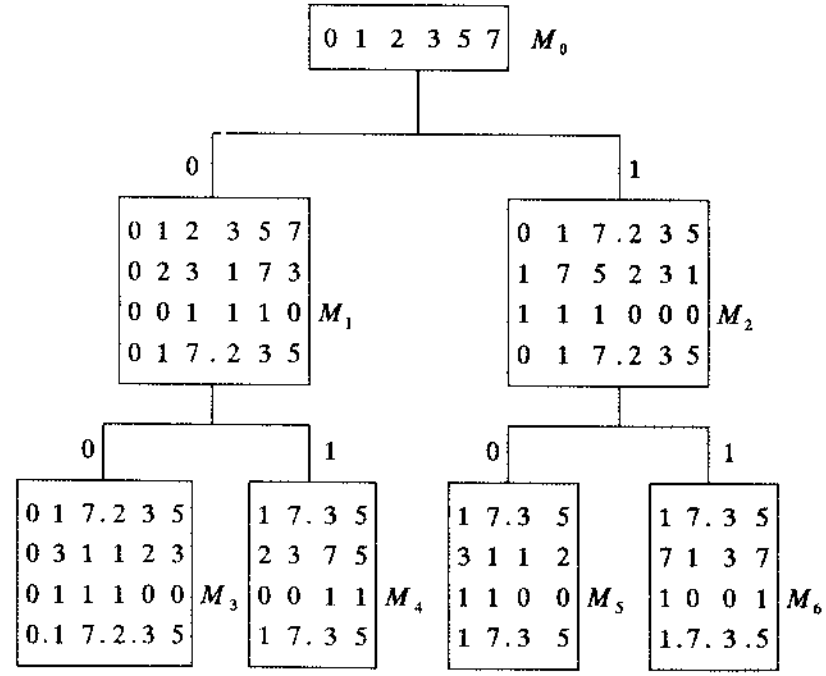


Рис.27.10

Исходим из последней строки 0.17.2.35 в M_3. Так как одноэлементные множества {0} и {2} дальнейшего подразбиения не допускают, то достаточно взять строку состояний 17.35. На вход автомата подаем слово x = 01; сопоставим ему матрицу M_4, (рис.27.10). Разбиения классов 17 и 35 не произошло. Не произойдет нового разбиения и при подаче на вход слова x = 10 (матрица M_5).

Исходим из последней строки 17.35 матрицы M_5. На вход автомата подаем слово x = 11; ему соответствует матрица M_6, приведенная на рис.27.10. Классы 17 и 35 разделились на одноэлементные множества и дальнейшего разбиения не допускают.

На этом построение классов неотличимости заканчивается. Все шесть состояний оказались попарно отличимыми словами (табл.27.7). На пересечении строки i и столбца j ($i < j$) стоит слово, которое отличает состояния i и j.

Автомат (см.табл.27.6) является минимальным. Построение классов неотличимости состояний автомата из табл.27.6, оформленное в виде дерева, приведено на рис.27.10. Дерево строится от корня к листьям, сверху вниз, а в каждом ярусе узлов слева направо.

Пример 2. Минимизировать автомат с функцией переходов и выходов, приведенных в табл.27.8.

Одинаковых столбцов в табл.27.8 нет. Построим разбиение множества состояний автомата на попарно непересекающиеся классы неотличимых состояний. Построение оформим в виде дерева (рис.27.11).

Исходим из множества состояний автомата, заданного строкой M_0 = 012345. Слову x = 0, поданному на вход автомата, сопоставим матрицу M_1 так же, как это делали в предыдущем

Таблица 27.7

	1	2	3	5	7
0	00	0	0	0	00
1		0	0	0	11
2			00	00	0
3				11	0
5					0

Таблица 27.8

	0	1	2	3	4	5
0	2,0	3,0	2,0	3,0	4,1	5,1
1	4,0	5,0	4,1	5,1	0,1	1,1

примере. Разделение множества состояний на одноэлементные множества не произошло. Затем так же, как это делали в пре-

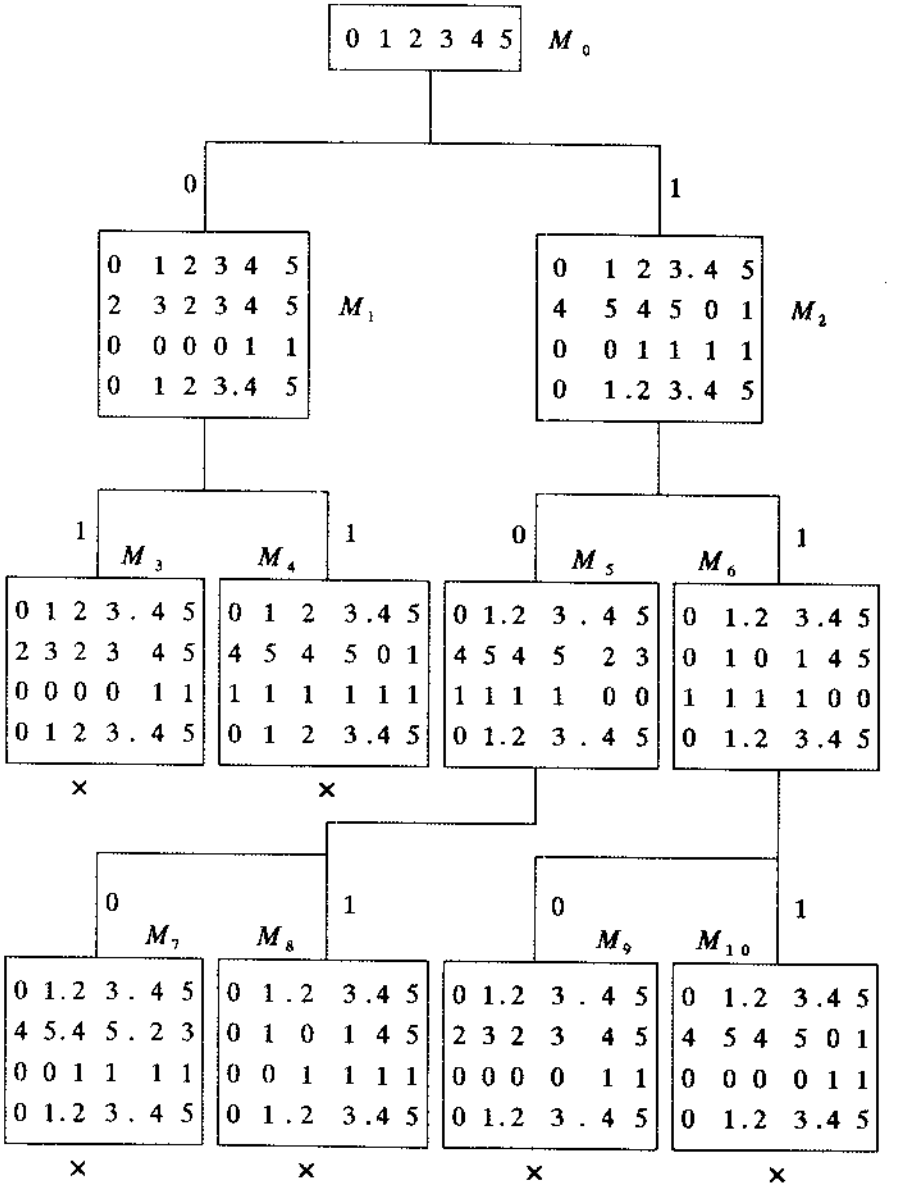


Рис.27.11

дыдущем примере, последовательно строим матрицы M_0, M_1, ..., M_{10}. Первые две строки матрицы M_i задают фунцию $f_i : Q \to Q$ с графиком G_i. Если при построении матрицы M_i окажется, что $G_i \subseteq G_j$ для некоторого $j < i$, то эту матрицу помечаем крестом; на такой матрице дальнейший рост дерева прекращается: ответвлений от этого узла больше не будет. На рис.16.11 таковы матрицы M_3, M_4, M_7, M_8, M_9, M_{10}, у которых $G_3 \subseteq G_1$, $G_4 \subseteq G_2$, $G_7 \subseteq G_5$, $G_8 \subseteq G_6$, $G_9 \subseteq G_3$, $G_{10} \subseteq G_2$. Строка 4 в последней построенной матрице M_{10} указывает классы неотличимости состояний $\{0,1\}$, $\{2,3\}$, $\{4,5\}$. Склеиваем их и получаем эквивалентный исходному минимальный автомат; он приведен в табл.27.9. Входные слова, отличающие его состояния 0, 2 и 4, указаны в табл.27.10.

27.5.3. Алгоритм разбиения множества состояний на классы неотличимых состояний

Пусть $A = (X,Y,Q,q_0,T,B)$ есть автомат, в котором $X = \{1, 2,\ldots,|X|\}$, $Y = \{1,2,\ldots,|Y|\}$, $Q = \{1,2,\ldots,|Q|\}$, $q_0 = 1$. Разбиение множества состояний Q на классы неотличимых состояний оформим в виде дерева. Будем строить его от корня к листьям, сверху вниз, а в каждом ярусе узлов слева направо.

0. В корне дерева помещаем матрицу-строку $M_0 = 12\ldots|Q|$ состояний автомата.

1. Допустим, что матрицы $M_0, M_1, \ldots, M_{i-1}$ построены. Переходим к построению матрицы M_i. Возможны следующие случаи:

1) из матрицы M_{i-1} исходит ребро, помеченное входным символом $a < |X|$. Тогда M_i строим следующим образом. Из последней строки матрицы M_{i-1} удаляем одноэлементные классы. Результат

$$q_{11}\ q_{12}\ \cdots\ q_{1,n_1} \cdot q_{21}\ q_{22}\ \cdots q_{2,n_2} \cdot\ \cdots \cdot q_{r1}\ q_{r2}\ \cdots\ q_{r,n_r}$$

делаем строкой 1 в M_i. Строки

$$T(q_{11},a+1) \cdot T(q_{12},a+1) \cdot \ldots \cdot T(q_{r,n_r},a+1),$$
$$B(q_{11},a+1)\ B(q_{12},a+1)\ \ldots\ B(q_{r,n_r},a+1)$$

Таблица 27.9

	0	2	4
0	2,0	2,0	4,1
1	4,0	4,1	0,1

Таблица 27.10

	2	4
0	1	1
2		0

делаем в M_i строками 2 и 3 соответственно. Каждый класс $q_{m1}\ \cdots\ q_{m,n_m}$ $(m=1,2,\ldots,r)$ строки 1 в M_1 разбиваем на подклассы так, что два состояния q и q' из $\{q_{m1},\ldots,q_{m,n_m}\}$ попадут в один класс тогда и только тогда, когда $B(q,a+1) = B(q',a+1)$. Вновь полученные классы $q_{m1}\ \cdots\ q_{m,l_1} \cdot q_{m,l_1+1}\ \cdots\ q_{m,l_2} \cdot \ldots \cdot q_{m,l_{s-1}} \ldots\ q_{m,l_s}$ $(m = 1,2,\ldots,r)$, отделенные друг от друга точками, запишем в одну строку и сделаем ее строкой 4 в M_i. Матрица M_i построена.

Первые две строки в M_i задают функцию $f_i : Q \to Q$ с графиком G_i. Если $G_i \subseteq G_v$ для некоторого v, $0 < v < i$, то матрицу M_i помечаем крестом.

Если строка 4 в M_i представляет собой разбиение на одноэлементные множества, то алгоритм работу заканчивает, и строка 4 в M_i дает искомое разбиение;

2) из матрицы M_{i-1} исходит ребро (с матрицей на конце), помеченное входным символом $|X|$. Возможны следующие случаи:

а) все построенные матрицы помечены крестом. Тогда алгоритм заканчивает работу, и строка 4 в матрице с наибольшим номером дает искомое разбиение;

б) не все построенные матрицы помечены крестом. Среди таких матриц выбираем матрицу с наименьшим номером, испускаем из нее ребро, помечаем его входным сигналом 1 и в концевой вершине этого ребра строим матрицу M_i так же, как это делали в случае 1.

3. Переходим к п.2.

Замечание. При функционировании алгоритма не нужно хранить все дерево матриц M_i. Из всего дерева при его построении следует сохранять только множества ML листьевых матриц этого дерева, а также получающиеся в процессе построения дерева матриц M_i множество SF функций f_i и текущее разбиение D множества состояний на классы неотличимости.

Приведенный алгоритм минимизации числа состояний автомата имеет достаточно большую вычислительную сложность. Поэтому исходный автомат предпочтительнее упростить сначала каким-либо другим более простым способом, например, склеив состояния, столбцы для которых в таблице переходов (и выходов) автомата одинаковы.

Замечание. Порождение классов неотличимых состояний можно провести согласно следующему алгоритму. Если корень дерева (рис.27.11) пометим начальным состоянием, а в каждый другой узел дерева поместим состояние, в которое перейдет автомат,

если на его вход подать входной сигнал, помечающий ребро, входящее в этот узел, и соответствующий выходной сигнал, то в результате получим бесконечное дерево переходов состояний автомата и его выходов. Это дерево имеет конечное число попарно различных бесконечных поддеревьев. Можно ограничиться конечными поддеревьями, с длиной ветвей, равной числу состояний автомата. Класс неотличимых состояний образуют все состояния, стоящие в корнях тех конечных деревьев, которые совпадают, если с их узлов убрать (стереть) пометки состояний, оставив пометки выходных символов.

Детерминизацию источника тоже можно организовать в виде дерева перехода состояний. В корень дерева помещается замыкание начального состояния источника. Если затем в некотором узле дерева помещено подмножество состояний S и если от этого узла отходит ребро, помеченное входным символом a, то концевой узел этого ребра помечается соответствующим макросостоянием. Построение дерева в узле прекращаем, если этот узел помечается макросостоянием, уже построенным ранее.

Аналогично в виде дерева перехода состояний можно реализовать алгоритм построения автомата для объединения, пересечения, конкатенации, итерации автоматных языков. Если соответствующие языки заданы источниками, то результирующий автомат можно получить детерминированным и минимальным.

28. АВТОМАТЫ И СВЕРХЪЯЗЫКИ

28.1. Макроавтоматы

Пусть X есть произвольный алфавит. Бесконечное слово $x = x(0)x(1)x(2)\ldots$ в алфавите X называется *сверхсловом*. *Сверхъязык* есть некоторое множество сверхслов в алфавите X. Пусть X^∞ есть множество всех сверхслов в алфавите X. Пусть $\lim x(t)$ есть предельное множество сверхслова x при $t \to \infty$, т.е. $\lim x(t)$ есть множество всех тех букв в алфавите X, которые в сверхслове x встречаются бесконечное число раз. Если Q есть множество внутренних состояний некоторого автомата, то *макросостояние* есть подмножество $Q' \subseteq Q$.

(Детерминированный) *макроавтомат* есть система объектов $MA = (X,Q,q_0,T,F)$, где X,Q,q_0,T задаются как в обычном (детерминированном) автомате, а $F \subseteq P(Q)$ есть некоторое множество макросостояний. Вместо слова макроавтомат иногда будем говорить просто автомат, если из контекста ясно, о каком автомате идет речь.

Пусть на вход макроавтомата MA подается сверхслово $x = x(0)x(1)\ldots$ из X^∞. Автомат MA, начиная работу в начальном состоянии, выдает бесконечную последовательность состояний $q(0),q(1),\ldots$ (или сверхслово $q = q(0)q(1)\ldots$ состояний), называемую *ходом* автомата MA на сверхслове x. Пусть $\lim q(t)$ есть (предельное) множество всех состояний, которые в последовательности $q(t)$ встречаются бесконечно много раз. Макроавтомат MA *допускает* сверхслово x, если $\lim q(t) \in F$.

Поведение $Beh(MA)$ макроавтомата MA есть множество всех входных сверхслов x, допустимых автоматом MA. Макроавтомат MA *представляет* (допускает, определяет) сверхъязык L, если $L = Beh(MA)$. Сверхъязык L *автоматно представим* (допустим, определим), если существует макроавтомат MA, для которого $L = Beh(MA)$. Пусть множество $L(MA,Q,F)$ есть сверхъязык, допустимый макроавтоматом MA с множеством состояний Q и с множеством макросостояний F.

Макроисточник есть система $MS = (X,Q,Q_0,D,F)$, где X,Q,Q_0, D задаются как в источнике, а множество $F \subseteq P(Q)$ есть некоторое подмножество макросостояний.

Иногда вместо слова макроисточник будем говорить просто источник, если из контекста ясно, о каком источнике идет речь.

Сверхслово $q(0)q(1)q(2)\ldots$ в алфавите состояний есть ход макроисточника MS на входном сверхслове $x = x(0)x(1)x(2)\ldots$ (множество всех таких ходов обозначим через $Rn(MS,x)$), если при последовательной подаче входных символов $x(0),x(1),\ldots$ (возможна подача и пустых символов) источник, исходя из начального состояния, проходит последовательность состояний $q(0),q(1),\ldots$ согласно своей таблице переходов.

Макроисточник MS *допускает* сверхслово x, если существует ход $q(t)$ MS на x, для которого $\lim q(t) \in F$. В противном случае макроисточник MS сверхслово x отвергает (не допускает). *Поведение макроисточника* MS есть множество $Beh(MS)$ всех входных сверхслов, допустимых источником MS. Макроисточник MS *представляет* (допускает, определяет) сверхъязык L, если $L = Beh(MS)$. Два макроисточника эквивалентны, если их поведения совпадают.

Введем некоторые операции над языками и сврхъязыками. Пусть M есть язык, а L,L_1,L_2 есть сверхъязыки в алфавите X.

Объединение $L_1 \cup L_2 = \{x \in X^\infty : x \in L_1 \vee x \in L_2\}$.

Пересечение $L_1 \cap L_2 = \{x \in X^\infty : x \in L_1,\ x \in L_2\}$.

Дополнение $CL = X^\infty - L$.

Сверхитерация (сильная итерация) языка M есть сверхъязык $M^\infty = M \cdot M \cdot M \cdot \ldots$.

Конкатенация (умножение) языка M и сверхъязыка L есть сверхъязык $M \cdot L = \{x = y \cdot z \in X^\infty : y \in M,\ z \in L\}$.

Пусть множество $L(MS,Q_0,F)$ означает сверхъязык, допустимый макроисточником MS с множеством начальных состояний Q_0 и с множеством выделенных макросостояний F.

Теорема. Алгоритмически разрешимы (распознаваемы) следующие свойства макроавтоматов.

1. Представляемый макроавтоматом сверхъязык пуст (проблема пустоты).

2. Представляемый макроавтоматом сверхъязык бесконечен.

Доказательство. Достаточно рассмотреть случай, когда макроавтомат $MA = (X,Q,q_0,T,F)$ имеет одноэлементное множество F макросостояний, именно, $Q = \{q_0,q_1,\ldots,q_k\}$, $Q' = \{q_1,q_2,\ldots,q_s\}$, $s \leqslant k$, $F = \{Q'\}$. Докажем для этого автомата два утверждения теоремы.

1. Допустим, что существует сверхслово x, переводящее начальное состояние в макросостояние Q', т.е. для некоторого хода $q(t)$ предельное макросостояние $\lim q(t) = Q'$. Тогда, во-первых, существует слово $y \in X^*$, переводящее начальное состояние q_0 в q_1, и во-вторых, для всякого $i = 1,2,\ldots,s-1$ существует слово $z_i \in X^*$, переводящее автомат из состояния q_i в состояние q_{i+1}, и слово $z_s \in X^*$, переводящее MA из q_s в q_1, причем все промежуточные состояния лежат в Q'. Наличие или отсутствие таких слов проверяется для каждого автомата, в том числе и для MA.

2. Представимый макроавтоматом сверхъязык либо пуст, либо бесконечен.

Теорема. Класс автоматно представимых сверхъязыков замкнут относительно булевых операций.

Доказательство. Пусть сверхъязыки L' и L'' представимы макроавтоматами $A'=(X,Q',q_0',T',F')$ и $A''=(X,Q'',q_0'',T'',F'')$ соответственно.

Пересечение $L' \cap L''$ представимо макроавтоматом $A = (X,Q,q_0,T,F)$, где $Q = Q' \times Q''$; $q_0 = (q_0',q_0'')$; $T((q',q''),a) = (T'(q',a),T''(q'',a))$; множество выделенных макросостояний F включает в себя любое подмножество из $F' \times F''$, в котором первая компонента лежит в F', а вторая в F''.

Дополнение $X^\infty - L'$ представимо макроавтоматом $A = (X,Q',q_0',T,P(Q)-F)$.

Объединение сверхъязыков выражается через их пересечение и дополнение. Макроавтомат для объединения можно построить как и в случае пересечения, с той лишь разницей, что множество выделенных макросостояний F включает в себя любое подмножество из $F' \times F''$, в котором первая компонента лежит в F', или вторая лежит в F''. На практике множество всех достижимых выделенных макросостояний для объединения можно получить, построив из начального состояния все пути до первого повтора состояния на каждом пути, выделив в каждом полученном пути цикл и отнеся все состояния этого цикла в множество выделенных состояний, если множество первых компонент этих состояний лежит в F' или множество вторых лежит в F''. Аналогично поступаем и для пересечения.

28.2. Конкатенация языка и сверхъязыка

Теорема. Конкатенация конечно автоматного языка и конечно автоматного сверхъязыка есть конечно автоматный сверхъязык.

Доказательство. Пусть $L = L(A,q_0,Q')$ есть автоматный язык, представимый автоматом $A = (X,Q,q_0,T,Q')$ и $SL = SL(MA, s_0,G)$ есть сверхъязык, представимый макроавтоматом $MA = (X, S,s_0,P,G)$. Тогда искомая конкатенация

$$L \cdot SL = \bigcup_{q' \in Q',\, F \in G} L(A,\ q_0,q') \cdot SL(MA,s_0,F).$$

Теорему достаточно доказать для случая, когда $L \cdot SL = L(A, q_0,\ q') \cdot SL(MA,s_0,F)$.

Пусть имеем один автомат A и неограниченный запас копий макроавтоматов MA. Принадлежность сверхслова $x = x(0)x(1)\ldots$ сверхъязыку $L \cdot SL$ распознаем следующим образом. Запустим на сверхслове x автомат A. Если далее в некоторый момент t_1 автомат A впервые придет в состояние q', то в этот момент запускаем копию 1 макроавтомата MA из неограниченного запаса копий макроавтоматов MA. Копия 1 макроавтомата MA запускается на сверхслове $x\{t_1,\infty) = x(t_1)x(t_1+1)\ldots$. Если в момент t_2 автомат A во второй раз придет в состояние q', то в этот момент запускаем копию 2 для MA. И так далее. Пусть $c_0(t)$ есть состояние автомата A в момент t, а $c_m(t)$ есть состояние копии m в момент t, если она в момент t активна (т.е. работает), и $c_m(t) = \Lambda$, если копия m в момент t не активна (т.е. пребывает в запасе). Тогда сверхслово x принадлежит $L \cdot SL$ тогда и только тогда, когда 1) активна хотя бы одна копия мак-

роавтомата MA; 2) хотя бы одна из активных копий макроавтомата MA вырабатывает предельное макросостояние F. Поэтому

$$x \in L \cdot SL \longleftrightarrow (\exists m)(\lim c_m(t) = F).$$

Заметим, что если для некоторого момента t' имеет место $c_m(t') = c_n(t')$, то $\forall t > t'$ $c_m(t) = c_n(t)$; поэтому дальнейшая работа одной из этих копий становится излишней. Следовательно, копию m или копию n можно вернуть в запас, например, копию с большим номером. Так как макроавтомат MA имеет $k+1$ состояний $s_0, s_1, \ldots, s_k$, то понадобится $k+1$ копий макроавтомата MA.

Построим макроавтомат MB, допускающий сверхъязык $L \cdot SL$. Состояния MB есть векторы $c = (c_0, c_1, \ldots, c_{k+1})$ из множества $Q \times F' \times F' \times \ldots \times F'$, где $F' = F \cup \{\Lambda\}$, c_0 есть состояние автомата A; $c_1, c_2, \ldots, c_{k+1}$ есть состояния макроавтомата MA или знаки Λ, при этом среди $c_1, \ldots, c_{k+1}$ все состояния из MA попарно различны; $(q_0, \Lambda, \ldots, \Lambda)$ есть начальное состояние. При воздействии буквы a из X на состояние $c(t) = (c_0(t), c_1(t), \ldots, c_{k+1}(t))$ вектор $c(t+1) = (c_0(t+1), c_1(t+1), \ldots, c_{k+1}(t+1))$ получим следующим образом. Компонетна $c_0(t+1) = T(c_0(t), a)$. Каждое не равное Λ состояние $c_i(t)$ для $i \geqslant 1$ перейдет в состояние $c_i(t+1) = P(c_i(t), a)$ макроавтомата MA. Далее осуществим "чистку": в векторе $c(t+1)$ из нескольких равных компонент (если таковые имеются) оставляем компоненту с меньшим номером, а остальные заменяем на Λ. Если теперь $c_0(t+1) = q'$, то первая (по порядку) компонента Λ в $c(t+1)$ замещается на начальное состояние s_0 макроавтомата MA. Если же $c_0(t+1) \neq q'$, то все "пустые" компоненты Λ в $c(t+1)$ остаются пустыми. Функция переходов макроавтомата MB построена.

Множество выделенных макросостояний $\{F_1, \ldots, F_h, \ldots, F_l\}$ для MB содержит всякое макросостояние $F_h = \{(c_0^j, c_1^j, \ldots, c_i^j, \ldots, c_{k+1}^j)$: $j = 1,2,\ldots,r$ при некотором $r\}$, для которого некоторая проекция на ось i $(1 \leqslant i \leqslant k+1)$: $\mathrm{Pr}_i F_h = F$. При этом компоненты вне i в векторах из F_h безразличны.

Покажем, что сверхъязык $L \cdot SL$ представим построенным макроавтоматом MB. Пусть ход $c(0)c(1)\ldots c(t)\ldots$ макроавтомата MB на сверхслове x из X^∞ таков, что $\lim c(t) = F_h$ при некотором $h \in \{1,2,\ldots,l\}$. Тогда $(\exists m)(1 \leqslant m \leqslant k+1$ & $\lim c_m(t) = F)$. Это означает, что включившись в некоторый момент, копия m для MA уже никогда не отключается (т.е. не переводится в запас). Далее копия m вырабатывает предельное макросостояние

F. Поэтому входное сверхслово x допускает представление $x = x[0,t) \cdot x[t,\infty)$, причем $x[0,t) \in L$, $x[t,\infty) \in SL$ и потому $x \in L \cdot SL$.

Пусть $x \in L \cdot SL$. Покажем, что ход $c(0)c(1)\ldots$ макроавтомата MB на x имеет предельное макросостояние из числа построенных, т.е. $\lim c(t) \in \{F_1, \ldots, F_l\}$. Так как $x \in L \cdot SL$, то сверхслово x допускает представление $x = x[0,t) \cdot x[t,\infty)$, причем $x[0,t) \in L$, $x[t,\infty) \in SL$. Так как $x[0,t) \in L$, то для вектора $c(t)$ имеем $(\exists p)(1 \leqslant p \leqslant k+1$ & $c_p(t) = s_0)$. В ходе дальнейшей работы копия MA может перейти в запас, "передав" свою работу копии с меньшим номером. И так далее. Так как индекс p не может убывать неограниченно, то $(\exists m \leqslant p)(\lim c_m(t) = F)$. Тогда макроавтомат MB в компоненте m вектора состояний $c(t)$ даст предельное макросостояние F, и потому $\lim c(t) \in \{F_1, \ldots, F_l\}$, т.е. сверхслово x допустимо построенным макроавтоматом MB. Теорема доказана.

28.3. Сверхитерация автоматных языков

Цель этого параграфа состоит в доказательстве того факта, что сверхитерация автоматного языка является автоматным сверхъязыком.

Определение. Сверхслово $x = x(0)x(1)\ldots$ в алфавите X *устойчиво* относительно автомата $A = (X, Q, q_0, T, Q')$, если существует бесконечно много хвостов $x[t_i,\infty)$, $i = 1,2,\ldots$, для которых:

1) все слова $x[0,t_i) \in L(A, q_0, Q')$, $i = 1,2,\ldots$;

2) автомат A склеивает каждый из хвостов $x[t_i,\infty)$, $i = 1,2,\ldots$, с исходным словом x.

Пусть L_{st} есть множество всех сверхслов, устойчивых относительно автомата A, допускающего язык L.

Лемма 1. Если язык L представим автоматом A, то $L_{st} \subseteq L^\infty$.

Доказательство. Пусть сверхслово $x = x(0)x(1)\ldots \in L_{st}$. Тогда существует бесконечно много хвостов $x[t_i,\infty)$, $i = 1,2,\ldots$, для которых:

1) все слова $x[0,t_i) \in L$; 2) автомат A склеивает каждый из хвостов $x[t_i,\infty)$ с исходным сверхсловом x.

Запустим на x автомат A (рис.28.1). В момент t_1' запустим копию A_1 автомата A. Слово $x[0,t_1') \in L$. Копия A_1 в некоторый момент $\nu_1 \geqslant t_1'$ склеивается с A. Так как хвостов $x[t_i,\infty)$ бесконечно много, то существует хвост $x[t_2',\infty)$ с $t_2' > \nu_1$. В момент t_2' запустим копию A_2 автомата A. Слово $x[0,t_2') \in L$, по-

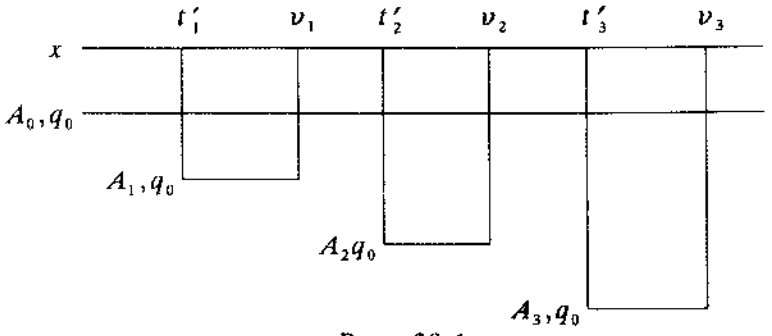


Рис. 28.1

этому автомат A вместе со склеившейся с ним копией A_1 придет в выделенное состояние. Тогда слово $x[t'_1,t'_2) \in L$. Хвост $x[t'_2,\infty)$ склеится автоматом A с x в неоторый момент ν_2. Так как хвостов $x[t_i,\infty)$ бесконечно много, то найдется хвост $x[t'_3,\infty)$ с $t'_3 > \nu_2$. Слово $x[0,t'_3) \in L$, поэтому автомат A вместе со склеившейся с ним копией A_2 в момент t'_3 придет в выделенное состояние. Тогда слово $x[t'_2,t'_3) \in L$. Продолжая эти рассуждения, мы покажем, что сверхслово $x=x[0,t'_1)\cdot x[t'_1,t'_2)\cdot x[t'_2,t'_3)\cdot\ldots$ составлено из слов $x[t'_i,t'_{i+1})$ из L, т.е. $x\in L^\infty$.

Лемма 2. Если представимый автоматом A язык L замкнут относительно итерации (т.е. $L = L^*$), то $L^\infty = L\cdot L_{st}$.

Доказательство. Пусть язык L представим автоматом $A = (X, Q,q_0,T,Q')$, где $|Q| = k$ и $L^* = L$. По лемме 1 множество $L_{st} \subseteq L^\infty$, поэтому любое сверхслово z из L_{st} разбивается на бесконечное число слов из L.

Добавив к z спереди любое слово из L, снова получим сверхслово из L^∞. Следовательно, $L\cdot L_{st} \subseteq L^\infty$.

Пусть теперь сверхслово $x \in L^\infty$. Для распознавания принадлежности сверхслова $x = x(0)x(1)\ldots$ сверхъязыку L^∞ поступим следующим образом. Запасем бесконечно много копий автомата A. Запустим на сверхслове x первую копию автомата A.

В каждый момент времени $t_0,t_1,t_2,\ldots$, в который копия для A приходит в выделенное состояние, запускаем следующую (находящуюся в запасе) копию автомата A. Все начальные отрезки $x[0,t_i)$, $i = 1,2,\ldots$, входного сверхслова лежат в L. Если автомат A имеет k состояний, то достаточно иметь работающими (активными) не более k копий автомата A, ибо при большем их числе найдутся такие, которые пребывают в одном и том же состоянии. Из всех таких копий достаточно оставить активной

только одну копию A, например, имеющую меньший номер, т.е. копию, ранее всех включившуюся. Заметим, что первая копия не отключается никогда.

Так как $x \in L^\infty$, то сверхслово x составлено из бесконечного числа слов, принадлежащих L. Так как все начальные отрезки для x до мест соединений этих слов лежат в L, то существует бесконечно много моментов, в которые включаются все новые копии автомата A. Так как в любой момент активны не более k копий автомата A, то происходит бесконечное число склеиваний активных копий автомата A. Пусть несколько копий A склеивается с копией с номером l_1, затем происходит склеивание с копией l_2, и так далее. В последовательности склеиваний копий A

$$(l_1,\ldots);\ (l_2,\ldots);\ (l_3,\ldots);\ldots;\ (l_i,\ldots);\ldots$$

Скобка $(l_i,\ldots)$ означает, что с копией l_i автомата A склеились все копии этой скобки (находящиеся в том же состоянии, что и копия l_i). Номера копий в скобке $(l_i,\ldots)$ возрастают по величине. В ряду чисел $l_1,l_2,\ldots$ некоторые числа встречаются бесконечно много раз. Выберем среди них наименьшее число l. С копией l автомата A происходит бесконечно много склеиваний других копий; при этом копия l, включившись в некоторый момент t, уже никогда не отключается. Сверхслово x разбивается на два куска: $x[0,t)$ и $x[t,\infty)$, причем слово $x[0,t) \in L$, а сверхслово $x[t,\infty)$ таково, что существует бесконечно много хвостов $x[t'_1,\infty)$, $x[t'_2,\infty),\ldots$ этого сверхслова, для которых:

1) все слова $x[t'_1,t'_i) \in L$, $i = 1,2,\ldots$;

2) автомат A склеивает каждый из хвостов $x[t'_i,\infty)$ со сверхсловом $x[t'_1,\infty)$.

Следовательно, $x[0,t'_1) \in L$, $x[t'_1,\infty) \in L_{st}$, т.е. $x \in L\cdot L_{st}$. Показано, что $L^\infty \subseteq L\cdot L_{st}$, что вместе с ранее доказанным обратным включением дает искомое равенство.

Лемма 3 (об устойчивости). Существует алгоритм, который по любому конечному автомату A строит макроавтомат MA, поведение которого есть множество всех сверхслов, устойчивых относительно A.

Доказательство. Пусть автомат $A = (X,Q,q_0,T,Q')$ с k состояниями допускает язык L. Пусть L_{st} есть множество всех сверхслов, устойчивых относительно автомата A. Состояниями автомата MA являются векторы $c = (c_1,c_2,\ldots,c_{k+1})$, где c_1 есть состояние автомата A; $c_{k+1} \in \{\Lambda,*\}$; $c_i \in Q \cup \{\Lambda\}$, $i =$

$2,3,\ldots,k$; при этом те из $c_2,c_3,\ldots,c_{k+1}$, которые являются состояниями автомата A, попарно различны. Вектор c всегда имеет в запасе пустую компоненту Λ. Множество S состояний для MA построено. Вектор $s_0 = (q_0,\Lambda,\ldots,\Lambda)$ есть начальное состояние макроавтомата MA. Функция M переходов для MA строится следующим образом. Пусть при состоянии $c(t) = (c_1(t), \ldots,c_{k+1}(t))$ на вход в MA поступает входной символ a. Построим вектор $c(t+1) = (c_1(t+1),\ldots,c_{k+1}(t+1))$, положив $c_1(t+1) = T(c_1(t),a)$. Для $i = 2,3,\ldots,k+1$ при $c_i(t) \neq \Lambda$ полагаем $c_i(t+1) = T(c_i(t),a)$, а при $c_i(t) \in \{\Lambda,*\}$ полагаем $c_i(t+1) = \Lambda$. Вектор $c'(t+1)$ (и тогда $M(c(t),a) = c'(t+1)$) сделаем результатом следующих операций.

1. Если хотя бы одна координата из $c_2(t+1),\ldots, c_{k+1}(t+1)$ совпадает с $c_1(t+1)$, то $c'(t+1) = (c_1(t+1),\Lambda,\ldots,\Lambda,*)$.

2. Если ни одна из компонент $c_2(t+1),\ldots,c_{k+1}(t+1)$ не есть $c_1(t+1)$, то проводим следующую "чистку" вектора $c(t+1)$: из всех одинаковых компонент в $c_2(t+1),\ldots,c_{k+1}(t+1)$ оставляем одну с меньшим номером. Такую "чистку" проводим до тех пор, пока все состояния в получившемся векторе не станут попарно различными.

Выделенным для MA является множество макросостояний $H = \{H_1,H_2,\ldots,H_i,\ldots,H_l\}$, где H_i есть любое множество векторов c, содержащих звездчатое состояние $(q,\Lambda,\ldots,\Lambda,*)$.

Макроавтомат $MA = (X,S,s_0,M,H)$ построен.

На практике вектор состояний экономнее строить без пустых символов, изменяя длину вектора, когда это необходимо.

Утверждение 1. Сверхслово $x \in L_{st}$ тогда и только тогда, когда звездчатые состояния в ходе макроавтомата MA встречаются бесконечно много раз.

Доказательство. Пусть в ходе MA на x первое звездчатое состояние есть $c(t_1)$. Это значит, что в момент t_1 одна из копий автомата A, включившись в момент между 0 и t_1, склеилась с первой копией. Если $c(t_2)$ есть второе звездчатое состояние, то это значит, что некоторая копия для A, включившись в момент между t_1 и t_2, склеилась в момент t_2 с первой копией автомата A. И так далее. Следовательно, существует бесконечно много хвостов $x\{t_i,\infty)$, $i = 1,2,\ldots$, для которых:

1) все слова $x\{0,t_i) \in L$;

2) автомат A склеивает каждый из хвостов $x\{t_i,\infty)$ с исходным сверхсловом x.

Каждая склейка означает появление звездчатого состояния в ходе MA на x.

Утверждение 1 доказано.

Утверждение 2. $L_{st} = Beh(MA)$.

Доказательство. Справедливы следующие высказывания.

1. Сверхслово $x \in L_{st}$.

2. Звездчатое состояние в ходе MA на x встречается бесконечно много раз.

3. Предельное макросостояние в ходе MA на x содержит звездчатое состояние.

4. Для $c(t) \in Rn(MA,x)$ предельное множество $\lim c(t) \in H$.

5. $x \in Beh(MA)$.

Следовательно, $L_{st} = Beh(MA)$.

Утверждение 2 установлено. Лемма 3 доказана.

Теорема (о сверхитерации). Если L есть конечно автоматный язык, то L^∞ есть конечно автоматный сверхъязык.

Доказательство. Пусть L есть автоматный язык, представимый автоматом A с начальным состоянием q_0. Так как $L^\infty = (L^*)^\infty$, то можно считать, что множество L замкнуто относительно итерации. Если этого нет, то вместо L следует взять язык L^* и представляющий его автомат. По лемме 2 сверхъязык $L^\infty = L\cdot L_{st}$. По лемме 3 сверхъязык L_{st} автоматно представим. По теореме из предыдущего параграфа автоматно представим сверхъязык $L\cdot L_{st}$. Следовательно, сверхъязык L^∞ автоматно представим.

Замечание. По автомату A макроавтомат MA, представляющий сверхъязык L^∞, строится эффективно.

28.4. Детерминизация макроисточника

Теорема. Существует алгоритм, который по любому заданному макроисточнику строит эквивалентный ему детерминированный макроавтомат.

Доказательство. Пусть $MA = (X,Q,Q_0,D,F)$ есть макроисточник; $F \subseteq P(Q)$ и множество $F = \{F_1,\ldots,F_k\}$. Тогда представимый макроисточником MA сверхъязык

$$SL = L(MA,Q_0,F) = \bigcup_{q_0\in Q_0,H\in F} L(MA,q_0,H).$$

Пусть $H = \{q_1,q_2,\ldots,q_s\}$ и $MA_1 = (X,Q,q_0,D,H)$ есть макроисточник, полученный из MA с сохранением в MA только состояний из H и ребер, им инцидентных (т.е. этим состояниям принадлежащих). Макроисточник MA_1 можно рассматривать как источник A_1, допускающий или не допускающий конечный язык обыч-

ным образом. Макроисточник MA_1 представляет сверхъязык SL_1, содержащий все сверхслова x из X^∞, на которых любой ход MA_1 содержит состояния только лишь из H. Если множество всех сверхслов $SL_1 = L(MA_1,q_1,H)$ пусто, то исходный сверхъязык пуст. Если нет, то $SL_1 =$

$$(L(A_1,q_1,q_2)\cdot L(A_1,q_2,q_3)\cdot\ldots\cdot L(A_1,q_{s-1},q_s)\cdot(L(A_1,q_s,q_1))^\infty,$$

и тогда представимое в макроисточнике MA множество сверхслов

$$SL = L(MA,q_0,H) = \begin{cases} L(A,q_0,q_1)\cdot SL_1, & \text{если } L(A,q_0,q_1) \neq \varnothing, \\ SL_1, & \text{если } L(A,q_0,q_1) = \varnothing \text{ и } q_1 = q_0, \\ \varnothing & \text{в остальных случаях.} \end{cases}$$

Далее, применяя теорему о сверхитерации языка, представимого конечным автоматом, и теорему о конкатенации языка, представимого автоматом, с сверхъязыком, представимым макроавтоматом, строим искомый детерминированный макроавтомат, представляющий сверхъязык SL.

Следствие. Класс сверхъязыков, представимых макроисточниками, замкнут относительно дополнения.

28.4.1. Общерегулярные сверхъязыки

Утверждение. 1. Если язык A (множество конечных слов) регулярен, то сверхъязык A^∞ общерегулярен.

2. Если язык A регулярен, а сверхъязык B общерегулярен, то сверхъязык $A\cdot B$ общерегулярен.

3. Если сверхъязыки A и B общерегулярны, то сверхъязык $A \cup B$ общерегулярен.

Теорема. Всякий сверхъязык общерегулярен тогда и только тогда, когда он автоматно представим.

Доказательство. *Синтез*. Автоматность общерегулярного сверхъязыка доказывается индукцией по его построению.

Анализ. Пусть макроисточник $MA = (X,Q,q_0,T,F)$ допускает сверхъязык L. Достаточно показать общерегулярность L для одноэлементного множества $F = \{q_1,\ldots,q_r\}$. Пусть источник $A = (X,Q,q_0,T,\{q_1\})$ допускает множество слов, переводящих A_1 из q_0 в q_1. Построим источники $A_{1\to2}$, $A_{2\to3}$, ..., $A_{r-1\to r}$, A_{r-1} следующим образом. Источник $A_{i\to j} = (X,Q,q_i,D,\{q_j\})$, где $(q,a,q') \in D \longleftrightarrow (q,a,q')\in T$ & $q\in F$ & $q'\in F$. Тогда $A_{i\to j}$ допускает такие и только такие слова, которые переводят A из q_i в q_j и при этом A не проходит через состояния вне F. Тогда $Beh(A) =$

$$Beh(A_1)\cdot(Beh(A_{1\to2})\cdot Beh(A_{2\to3})\ldots Beh(A_{r-1\to r})\cdot Beh(A_{r-1}))^\infty.$$

Поэтому множество $Beh(MA)$ общерегулярно. Теорема доказана.

Построение общерегулярного множества сверхслов, представляющего дополнение сверхъязыка $Beh(MA)$, представимого макроисточником $MA = (X,Q,Q_0,D,F)$, можно осуществить следующим образом. Сверхслово $w \in X^\infty$ принадлежит $Beh(MA) \longleftrightarrow$

$$(\forall r)_{r\in Rn(MA,w)\,\&\,r(0)\in Q_0}\ \lim(r) \in F.$$

Пусть дополнение $Pow(Q)$ есть F для простоты состоит из единственного макросостояния $G = \{q_1,\ldots,q_r\}$. Построим источники A_1, $A_{1\to2}$, $A_{2\to3}$, ..., $A_{r-1\to r}$, A_{r-1} следующим образом. A_1 строится по $A_1' = (X,Q,Q_0,D,\{q_1\})$ так, что для всякого хода A_1' на входном конечном слове заключительным оказалось бы состояние q_1. Для источника $A_{i\to j} = (X,Q,q_i,T,\{q_j\})$ положим $(q,a,q')\in T \longleftrightarrow (q,a,q')\in D$ & $q\in G$ & $q'\in G$. Тогда $A_{i\to j}$ допускает такие и только такие слова, которые переводят A из q_i в q_j и при этом A не проходит через состояния вне G. Тогда дополнение

$$X^\infty - Beh(MA) =$$
$$Beh(A_1)\cdot(Beh(A_{1\to2})\cdot Beh(A_{2\to3})\ldots Beh(A_{r-1\to r})\cdot Beh(A_{r-1}))^\infty,$$

т.е. дополнение $X^\infty - Beh(MA)$ общерегулярно.

Проекция сверхслова с одного алфавита на другой осуществляется побуквенно, как и в случае проекции конечного слова. Проекция сверхъязыка осуществляется проекцией каждого его сверхслова.

Теорема. Класс автоматно представимых сверхъязыков замкнут относительно проекции.

Доказательство. Пусть макроавтомат $MA = (X,Q,q_0,T,F)$ представляет сверхъязык $SL \subseteq X^\infty$ и функция $f\colon X \to Y$ осуществляет проекцию сверхъязыка SL с алфавита X на алфавит Y. Тогда макроисточник $MA_f = (X,Q,Q_0,D,F)$, где $(q,f(a),q') \in D \longleftrightarrow (q,a,q')\in T$, представляет сверхъязык $f(SL)$. По макроисточнику MA_f можно построить детерминированный макроавтомат, представляющий сверхъязык $f(SL)$.

29. ПРОБЛЕМА УНИФОРМИЗАЦИИ

29.1. Языки и операторы

Пусть зафиксированы входной X и выходной Y алфавиты. Всякий конечный автомат с выходом реализует некоторый словарный

(сверхсловарный) оператор $\mathbb{O}(x) = y$, определенный на множестве слов (сверхслов) в алфавите X и принимающий значения в множестве слов (сверхслов) в алфавите Y. График оператора $\mathbb{O}$ есть множество $G = \{(x,y) : \mathbb{O}(y) = y\}$. В словарном операторе пара $(x,y) \in X^* \times Y^*$, а в сверхсловарном операторе пара $(x,y) \in X^\infty \times Y^\infty$.

Поведение автоматов без выхода с выделенными состояниями определяли в терминах представимости языков (сверхъязыков). Поведение автоматов с выходом определяется в терминах реализации (т.е. представимости) операторов. Между обоими понятиями представимости имеется естественная связь, основанная на сводимости друг к другу свойств множеств и свойств функций. Именно, произвольное множество, рассматриваемое как подмножество некоторого универсального множества (универсума), однозначно определяется своей характеристической функцией. Произвольная функция в свою очередь может рассматриваться как множество точек ее графика. В нашем случае в роли множеств выступают языки (сверхъязыки), а в роли функций – операторы.

С каждым подалфавитом $Y' \subseteq Y$ оператор $\mathbb{O}$ ассоциирует язык $\mathbb{O}^{-1}(Y') = \{x = x(0)x(1)\ldots x(k) \in X^* : \mathbb{O}(x) = y = y(0)y(1)\ldots y(k)\} \in (Y')^*$. Скажем тогда, что оператор $\mathbb{O}$ *представляет* язык $\mathbb{O}^{-1}(Y')$ множеством выходов (т.е. выходными буквами из Y'). Оператор $\mathbb{O}$ допускает все слова из $\mathbb{O}^{-1}(Y')$ и отвергает все другие слова.

Аналогичные построения можно выполнить относительно сверхъязыков. Именно, с каждым подмножеством $Y' \subseteq Y$ оператор $\mathbb{O}$ ассоциирует сверхъязык $\mathbb{O}^{-1}(Y') = \{x = x(0)x(1)\ldots \in X^\infty : \mathbb{O}(x) = y = y(0)y(1)\ldots \in Y^\infty, \lim y(t) = Y'\}$. Скажем тогда, что оператор $\mathbb{O}$ представляет сверхъязык $\mathbb{O}^{-1}(Y')$, т.е. допускает все сверхслова из $\mathbb{O}^{-1}(Y')$ и отвергает все слова вне $\mathbb{O}^{-1}(Y')$.

Автомат $A = (X,Y,Q,q_0,T,B)$ с выходом представляет язык (сврхъязык) L выходами из $Y' \in Y$, если оператор $\mathbb{O}$, соответствующий этому автомату, представляет этот язык (сверхъязык) выходами из алфавита Y'.

Теорема. Класс языков (сверхъязыков), представимых в конечных автоматах с выходом, совпадает с классом языков (сверхъязыков), представимых в конечных автоматах без выхода.

Доказательство. С произвольным автоматом без выхода $A = (X,Q,q_0,T,F)$, где $Q = \{q_0,q_1,\ldots,q_k\}$, $F = \{q_{i1},\ldots,q_{ir}\}$, можно ассоциировать автомат с выходом $A' = (X,Y,Q,q_0,T,B)$, где

$Y = \{y_0,y_1,\ldots,y_k\}$, а функция выхода B строится следующим образом. Если $T(q_i,x) = q_j$, то $B(q_i,x) = y_j$. Язык $Beh(A)$ совпадает с языком, представимым автоматом A' множеством выходов $Y' = \{y_{i1},\ldots,y_{ir}\}$.

Пусть макроавтомат без выхода $A = (X,Q,q_0,T,F)$, где $Q = \{q_0,q_1,\ldots,q_k\}$, $F = \{\{q_{i1},\ldots,q_{ir}\},\ldots,\{q_{j1},\ldots,q_{jl}\}\} = P(Q)$, представляет сверхъязык $Beh(A)$. Построим автомат с выходом $A' = (X,Q,q_0,T,B)$, где $Y = \{y_0,y_1,\ldots,y_k\}$, а функция выходов B строится так же, как и в предыдущем случае. Сверхъязык $Beh(A)$ представим автоматом с выходом A' макромножеством выходов $Y' = \{\{y_{i1},\ldots,y_{ir}\},\ldots,\{y_{j1},\ldots,y_{jl}\}\}$.

Пусть автомат Мура с выходом $A = (X,Y,Q,q_0,T,B)$ представляет язык $Beh(A)$ множеством $Y' \subseteq Y$. Положим $F = \{q \in Q : \exists y \in Y'\ (B(q) = y)\}$. Тогда автомат без выхода $A' = (X,Q,q_0,T,F)$ представляет тот же язык $Beh(A)$.

Если теперь автомат Мура с выходом $A = (X,Y,Q,q_0,T,B)$ представляет сверхъязык $\mathbb{O}^{-1}(Y')$ некоторым макромножеством $Y' = \{\{y_{i1},\ldots,y_{ir}\},\ldots,\{y_{j1},\ldots,y_{jl}\}\} = \{Y_i,\ldots,Y_j\} \in P(Y)$, то при

$F = \{F_i,\ldots,F_j\}$, где

$F_i = \{q \in Q : \exists y \in Y_i\ B(q) = y\}$,

$\ldots$

$F_j = \{q \in Q : \exists y \in Y_j\ B(q) = y\}$,

автомат $A = (X,Q,q_0,T,F)$ представляет тот же самый язык.

Замечание. Алгоритм минимизации автоматов без выхода получается сочетанием алгоритма минимизации автоматов с выходом с доказанной теоремой.

Операторы можно рассматривать и вне понятия конечного автомата. Словарный оператор можно рассматривать как функцию $\mathbb{O} : X^* \to Y^*$, а сверхсловарный оператор – как функцию $\mathbb{O} : X^\infty \to Y^\infty$.

Сверхсловарный оператор $\mathbb{O} : X^\infty \to Y^\infty$ есть оператор без предвосхищения (его еще называют ограниченнодетерминированным оператором), если для всяких входных сверхслов

$$x' = x'(0)x'(1)\ \ldots\ x'(t)\ \ldots$$
$$x'' = x''(0)x''(1)\ \ldots\ x''(t)\ \ldots$$

соответствующие им выходные слова

$$\mathbb{O}(x') = y' = y'(0)y'(1)\ \ldots\ y'(t)\ \ldots,$$

$$\mathbb{O}(x'') = y'' = y''(0)y''(1) \ldots y''(t) \ldots$$

таковы, что для всякого t равенство

$$x'(0)x'(1) \ldots x'(t) = x''(0)x''(1) \ldots x''(t)$$

влечет равенство

$$y'(0)y'(1) \ldots y'(t) = y''(0)y''(1) \ldots y''(t).$$

В противном случае оператор $\mathbb{O}$ является оператором с предвосхищением.

Словарный оператор $\mathbb{O}: X^* \to Y^*$ есть оператор без предвосхищения, если:

1) равенство $\mathbb{O}(x) = y$ влечет равенство длин слов x и y;

2) для всяких входных слов x', x'' и для всякого t равенство $x'(0)x'(1) \ldots x'(t) = x''(0)x''(1) \ldots x''(t)$ влечет равенство $\mathbb{O}(x') = y' = y'(0)y'(1)\ldots y'(t) = y''(0)y''(1)y''(2)\ldots y''(t) = y'' = x''$;

3) оператор $\mathbb{O}$, применимый к слову x, применим и ко всякому его начальному отрезку.

В противном случае словарный оператор $\mathbb{O}$ есть оператор с предвосхищением.

В операторе без предвосхищения выходная буква в момент t не зависит от входных букв в момент $t' > t$. В операторах с предвосхищением это условие может не выполняться. В конечных автоматах реализуются только операторы без предвосхищения.

Утверждение. Если оператор $\mathbb{O}$ реализуем в некотором конечном автомате, то его график представим конечным автоматом.

Доказательство. Пусть для определенности $\mathbb{O}$ есть сверхсловарный оператор, реализуемый автоматом $A = (X,Y,Q,q_0,T,B)$. Если граф-схему для A (где каждое ребро снабжено и входом, и выходом) рассматривать как макроисточник в алфавите $X \times Y$, то график оператора A совпадает с поведением этого макроисточника, имеющего в качестве множества выделенных макросостояний множество всех подмножеств множества состояний исходного автомата. Утверждение доказано.

Поставим теперь следующий вопрос. Если график некоторого оператора $\mathbb{O}: X^\infty \to Y^\infty$ представим в некотором макроавтомате, то можно ли реализовать оператор $\mathbb{O}$ в подходящем конечном автомате? Ответ отрицателен. Это показывает следующий пример.

Пусть сверхсловарный оператор $\mathbb{O}: X^\infty \to Y^\infty$ с $X = Y = \{0, 1\}$, сопоставляя сверхслову 00...0... (сплошь из нулей) то же самое сверхслово 00...0..., а всякому другому сверхслову – сверхслово 11...1... (сплошь из единиц). Оператор $\mathbb{O}$ является

оператором с предвосхищением, и потому он не реализуем с помощью конечного автомата. Между тем, график оператора $\mathbb{O}$ представим в макроисточнике

$MS = (X \times Y, Q; Q_0, D, F)$, где $Q = \{q_0,q_1,q_2\}$; $Q_0 = \{q_0,q_1\}$; $D = \{(q_0,00,q_0),(q_1,01,q_1),(q_1,11,q_2),(q_2,11,q_2),(q_2,01,q_2)\}$; $F = \{\{q_0\},\{q_2\}\}$.

Ранее показано, что классы сверхслов, представимых в макроисточниках и макроавтоматах, совпадают.

29.1.1. Униформизация

Пусть дан сверхъязык SL в алфавите $X \times Y$. Сверхсловарный оператор $\mathbb{O}: X^\infty \to Y^\infty$ униформизует сверхъязык SL, если для всякого $x \in X^\infty$ равенство $\mathbb{O}(x) = y$ влечет $(x,y) \in SL$.

Можно говорить об униформизации сверхъязыка SL оператором $Y^\infty \to X^\infty$. Понятие униформизации можно сформулировать и для сверхъязыков в алфавите $X_1 \times X_2 \times \ldots \times X_n$ по любой переменной.

Понятие униформизации есть одно из важнейших понятий теории множеств. Очень важно оно и в теории конечных автоматов. Проблема униформизации конечно автоматного сверхъязыка состоит в следующем. Пусть задан макроавтомат A, представляющий некоторый сверхъязык SL в алфавите $X \times Y$. Требуется выяснить, существует ли сверхсловарный конечно автоматный оператор $\mathbb{O}$, униформизующий сверхъязык SL; если да, то построить конечный автомат с выходом, реализующий такой оператор.

Покажем, что проблема униформизации конечно автоматного сверхъязыка решается положительно. Решение этой задачи дадим в терминах предложенной Макнотоном игровой интерпретации, в той редакции, которую предложил для этого Б.А.Трахтенброт.

29.2. Игры

Будем рассматривать игры с полной информацией. Назовем их просто играми. Игре присущи следующие особенности.

1. В игре участвуют два игрока, называемых белыми и черными. Первый ход делают белые; черные отвечают своим ходом. Затем ходы чередуются: белые, черные, белые, черные, и так далее, образуя партию игры.

2. Ход белых есть выбор одной буквы из алфавита $X = \{x_1, \ldots, x_m\}$. Ход черных есть выбор одной буквы из алфавита $Y = \{y_1, \ldots, y_n\}$.

3. При каждом ходе очередной игрок знает все предыдущие

ходы данной партии (полная информация).

4. Для каждой партии имеет место либо выигрыш белых, либо выигрыш черных.

Партия может быть конечной и бесконечной. В конечной партии партия прекращается после конечного числа шагов. Конечную партию можно рассматривать как слово в алфавите $X \times Y$:

$$x(0)x(1)\ \ldots\ x(t),$$
$$y(0)y(1)\ \ldots\ y(t),$$

или как пару слов $x[0,t]$, $y[0,t]$ отдельно ходов белых и ходов черных. Длина партии есть t.

Бесконечная партия продолжается неограниченно. Бесконечную партию можно рассматривать как сверхслово

$$x(0)x(1)\ \ldots\ x(t)\ \ldots$$
$$y(0)y(1)\ \ldots\ y(t)\ \ldots$$

в алфавите $X \times Y$, или как пару сверхслов $x[0,\infty)$, $y[0,\infty)$ отдельно ходов белых и ходов черных.

Игра конечна, если для некоторого числа p любая партия не длинее p. В противном случае игра бесконечна.

Всякую конечную игру можно задать в виде дерева, ребра которого помечены буквами из $X \times Y$, причем ребра, выходящие из одной вершины, по разному помечены. Концевые вершины (листья) разбиты на два класса: класс вершин, помеченных знаком "+" (это выигрыш белых), и класс вершин, помеченных знаком "–" (это выигрыш черных). В корне дерева помещена пешка. Белые выбирают любую букву $x(0)$ из X; черные в ответ выбирают любую букву $y(0)$ из Y. Пешка перемещается в вершину по ребру, помеченному парой $(x(0),y(0))$. Далее белые выбирают одну X-букву (т.е. букву из X) $x(1)$, на что черные отвечают выбором одной Y-буквы $y(1)$. Пешка перемещается в вершину по ребру, помеченному парой $(x(1),y(1))$. И так далее. Партия заканчивается, если пешка окажется в концевой вершине; ее пометка указывает, кто выиграл. Множество партий, выигрышных для белых, есть язык в алфавите $X \times Y$.

На рис.29.1 приведена конечная игра с $X = \{0,1\}$, $Y = \{a,b\}$.

Для всякой конечной игры можно определить, какая из сторон имеет выигрыш, и указать последовательность ходов, приводящую к выигрышу. Покажем, как это делается для игры, дерево которой изображено на рис.29.1. Просматриваем все поддеревья, имеющие корень в предпоследнем ярусе 2; это под-

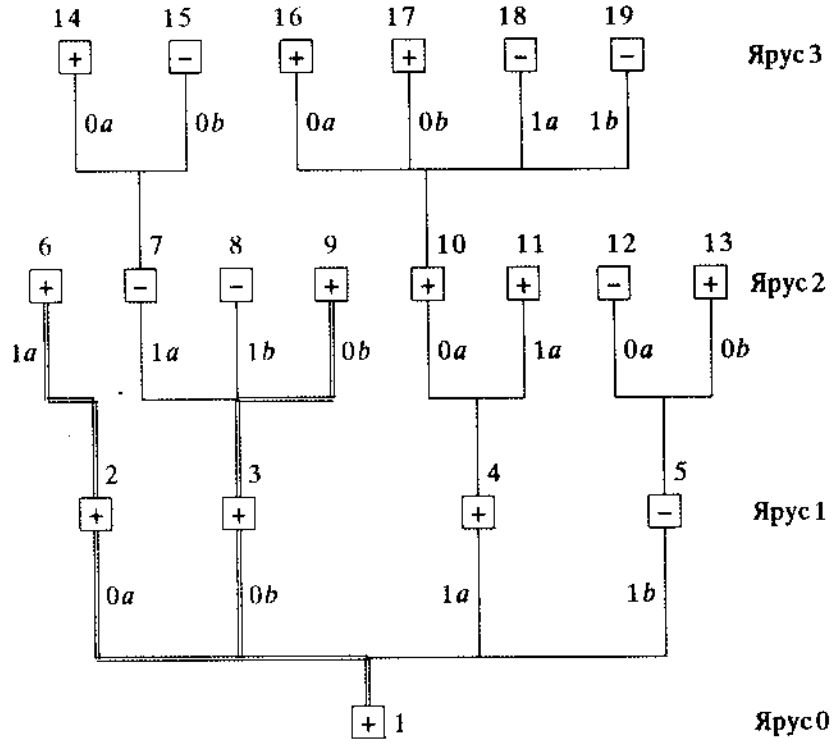


Рис. 29.1

деревья с корнем в вершинах 7 и 10. Из вершины 7 выигрывают черные, ибо при единственно возможном ходе белых 0 черные отвечают выбором b и выигрывают. Вершину 7 в знак выигрыша черных помечаем знаком "–". Из вершины 10 выигрывают белые, выбирая ход 0; любой ответ черных приводит их к проигрышу. Вершину 10 в знак выигрыша белых помечаем знаком "+". Далее просматриваем все поддеревья с корнем в ярусе 1. Из вершины 2 выигрывают белые; из вершины 3 ходом 0 выигрывают белые; из вершины 4 произвольным ходом выигрывают белые; из вершины 5 при единственно возможном ходе белых 0 ответом a выигрывают черные. Вершины 2,3,4 помечаем знаком "+", а вершину 5 – знаком "–". Из вершины 1 ходом 0 выигрывают белые; помечаем ее знаком "+". В этой игре белые начинают и выигрывают (в любой партии). Соответствующие ходы отмечены двойными линиями. В этой игре белые имеют выигрывающую стратегию (т.е. белые начинают и выигрывают).

В бесконечных играх все партии бесконечны. Такую игру можно изобразить бесконечным деревом (рис.29.2) из каждой вершины которого исходят mn ребер, помеченных парами (x,y) из $X \times Y$, где $X = \{x_1,\ldots,x_m\}$, $Y = \{y_1,\ldots,y_n\}$.

Игра есть дерево. Партия есть путь в дереве. В этом бесконечном дереве отмечается множество путей (партий), выигрышных для белых. Остальные пути (партии) выигрышны для черных. Таким образом различные игры при фиксированных X и Y отличаются только множествами партий (путей), выигрышных для белых. Это множество есть некоторый сверхъязык в алфавите $X \times Y$. С другой стороны, всякий сверхъязык L из $(X\times Y)^{\infty}$ задает бесконечную игру, в которой множество партий (путей) из L выигрышно для белых.

29.2.1. Игры с конечным числом состояний

Пусть $A = (X\times Y,Q,q_0,T,F)$ есть макроавтомат, где $F \subseteq P(Q)$. Игра с конечным числом состояний (автоматная игра) осуществляется следующим образом. Пешка помещается в вершину $q(0) = q_0$ граф-схемы автомата A. Белые называют свой первый ход

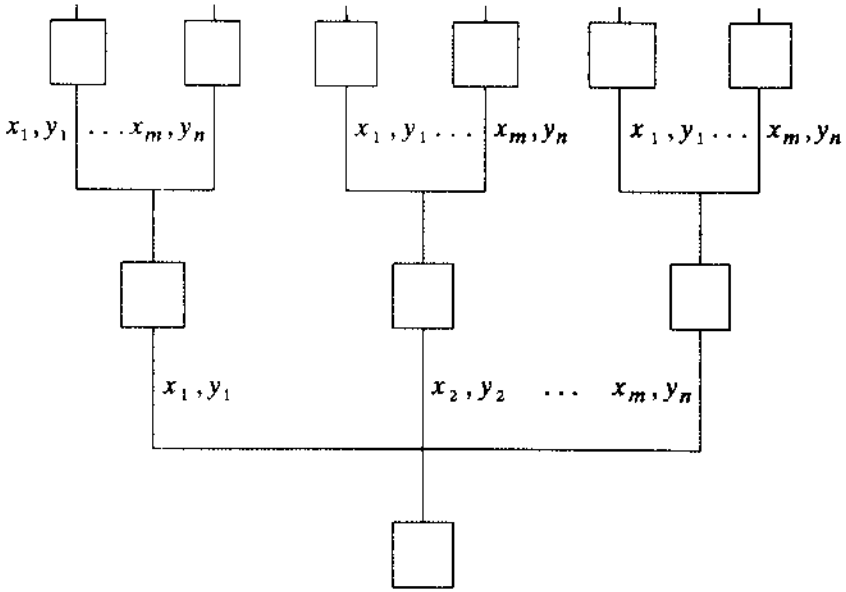


Рис.29.2

$x(0)$ из X; черные отвечают выбором $y(0)$ из Y. Пешка перемещается в вершину

$$q(1) = T(q(0),(x(0),y(0))).$$

Далее белые выбирают $x(1)$ из X, а черные отвечают выбором $y(1)$ из Y. Пешка перемещается в вершину

$$q(2) = T(q(1),(x(1),y(1))).$$

Такой обмен ходами продолжается неограниченно:

$$x(0)\quad x(1)\quad \ldots\quad x(t)\ \ldots$$
$$y(0)\quad y(1)\quad \ldots\quad y(t)\ \ldots;$$

в результате этого пешка последовательно перемещается в вершины

$$q(0)\quad q(1)\quad \ldots\quad q(t)\ \ldots,$$

образуя сверхслово состояний q из Q^{∞}. Белые выигрывают, если $\lim q(t)$ лежит в F, и проигрывают в противном случае.

29.3. Стратегии

Пусть бесконечная игра с конечным числом состояний осуществляется с помощью макроавтомата $A = (X\times Y,Q,q_0,T,F)$.

Стратегия черных есть оператор без предвосхищения, перерабатывающий входные сверхслова $x = x(0)x(1)\ldots$, являющиеся последовательными ходами белых, в выходные сверхслова $y = y(0)y(1)\ldots$, являющиеся ответными ходами черных. При этом ход черных $y(t)$ зависит от предшествующих ходов белых $x(0)x(1)\ldots x(t)$. *Стратегия белых* есть оператор без предвосхищения, перерабатывающий сверхслова $y = y(0)y(1)\ldots$, являющиеся ходами черных, в сверхслова $x = x(0)x(1)\ldots$, являющиеся ходами белых. При этом ход белых $x(t)$ в момент t зависит от $y(0)y(1)\ldots y(t-1)$, а ход $x(0)$ от ходов черных не зависит.

Стратегия белых есть оператор с задержкой. Всякий оператор $y = \mathbb{O}''(x)$ без предвосхищения есть некоторая стратегия черных. Всякий оператор $x = \mathbb{O}'(y)$ без предвосхищения с задержкой есть некоторая стратегия белых.

Если белые и черные выбрали стратегии (операторы) $\mathbb{O}'$ и $\mathbb{O}''$ соответственно, то партия, которая будет сыграна, полностью определена; обозначим ее через $(\mathbb{O}',\mathbb{O}'')$. Если в партии $(\mathbb{O}',\mathbb{O}'')$, белые начинают и выигрывают, то будем говорить, что стратегия белых $\mathbb{O}'$ *бьет* стратегию черных $\mathbb{O}''$. Если стратегия белых $\mathbb{O}'$ бьет любую стратегию черных $\mathbb{O}''$, то $\mathbb{O}'$ есть выигрывающая стратегия белых. Аналогично для черных.

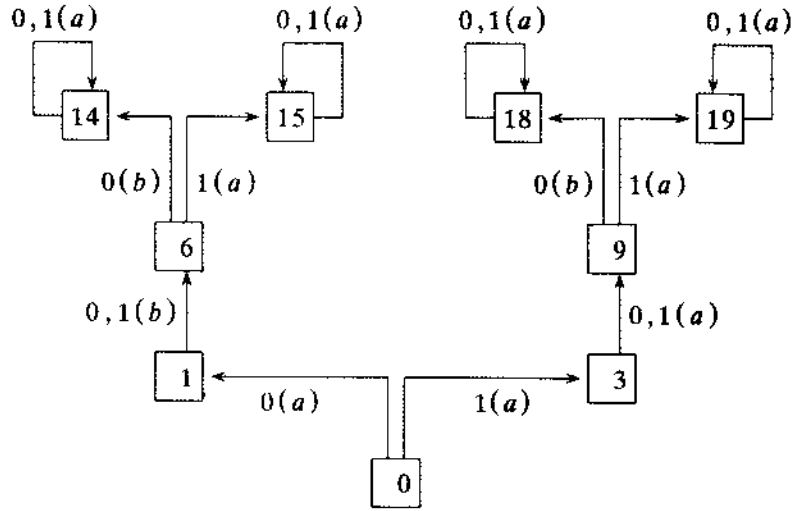


Рис.29.5

В классе всех бесконечных игр существуют такие игры, в которых никакая из сторон не имеет выигрывающей стратегии. В случае конечно автоматных игр положение иное: здесь всегда одна из сторон имеет выигрывающую стратегию и даже конечно автоматную. К ее построению и переходим.

29.4. Униформизация конечно автоматных языков

Проблема униформизации конечно автоматного сверхъязыка $L = (X \times Y)^{\infty}$, представимого макроавтоматом A, может быть интерпретирована как задача о существовании и построении выигрывающей конечно автоматной стратегии, которая является оператором без предвосхищения $y = \mathbb{O}''(x)$ для черных, или оператором без предвосхищения $x = \mathbb{O}'(y)$ для белых, ибо в конечно автоматной игре всегда одна из сторон имеет выигрывающую конечно автоматную стратегию.

29.4.1. Порядковые векторы и порядковые стратегии

Пусть $A = (X \times Y, Q, q_0, T, F)$ есть макроавтомат, представляющий сверхъязык $L = L(A, q_0, F)$; $Q = \{q_0, q_1, \ldots, q_n\}$ есть множество состояний в конечно автоматной игре $G(A)$, соответствующей сверхъязыку L. Возможны $(n+1)!$ перестановок состояний из

Q; назовем каждую такую перестановку $(q_{i_0}, q_{i_1}, \ldots, q_{i_n})$ *порядковым* вектором и будем обозначать его через $(i_0, i_1, \ldots, i_n)$. Рассмотрим некоторую партию

$$\begin{array}{l} x(0)\ x(1)\ \ldots\ x(t)\ \ldots \\ y(1)\ y(2)\ \ldots\ y(t)\ \ldots \end{array} \qquad (29.1)$$

в конечно автоматной игре $G(A)$. Партии (29.1) соответствует последовательность состояний

$$q(0)\ q(1)\ \ldots\ q(t)\ \ldots \qquad (29.2)$$

Наряду с последовательностью (29.2) сопоставим партии (29.1) последовательность порядковых векторов

$$\mathbf{q}(0)\ \mathbf{q}(1)\ \ldots\ \mathbf{q}(t)\ \ldots, \qquad (29.3)$$

индуктивно определяемую следующим образом:

1) $\mathbf{q}(0) = (0, 1, \ldots, n)$;

2) если $\mathbf{q}(t) = (i_0, i_1, \ldots, i_k, i_{k+1}, \ldots, i_n)$ и $q(t+1) = q_k$, то $\mathbf{q}(t+1) = (i_k, i_1, i_2, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n)$.

Ниже следует пример последовательности состояний и соответствующей последовательности порядковых векторов.

q_0	q_2	q_2	q_2	q_1	q_0	q_1	q_1	q_3	q_4	q_1	q_2	q_4
0	2	2	2	1	0	1	1	3	4	1	2	4
1	0	0	0	2	1	0	0	1	3	4	1	2
2	1	1	1	0	2	2	2	0	1	3	4	1
3	3	3	3	3	3	3	3	3	0	0	3	3
4	4	4	4	4	4	4	4	4	2	2	0	0

Справедлива следующая лемма.

Лемма 1. Если $\mathbf{q}(t) = (\ldots, i, \ldots, j, \ldots)$, $\mathbf{q}(t+n) = (\ldots, j, \ldots, i, \ldots)$, то в промежутках между t и $t+n$ вершина j (состояние q_j) посещалась, т.е. занимала в векторе первое место, по крайней мере один раз.

Пусть макросостояние $Q' = \{q_0, q_1, \ldots, q_s\} \subseteq Q$. Пусть $n(t)$ есть номер позиции в векторе $\mathbf{q}(t)$, которая в следующем векторе $\mathbf{q}(t+1)$ стала первой. Например,

$\mathbf{q}(t) = (q_2, q_1, q_4, q_0, q_3, q_5)$, $n(t) = 3$,
$\mathbf{q}(t+1) = (q_4, q_2, q_1, q_0, q_3, q_5)$.

Пусть в последовательности (29.3), которую запишем в виде

$$q(0)\ q(1)\ \ldots\ q(t')\ q(t'+1)\ \ldots\ q(t'')\ q(t''+1)\ \ldots,$$

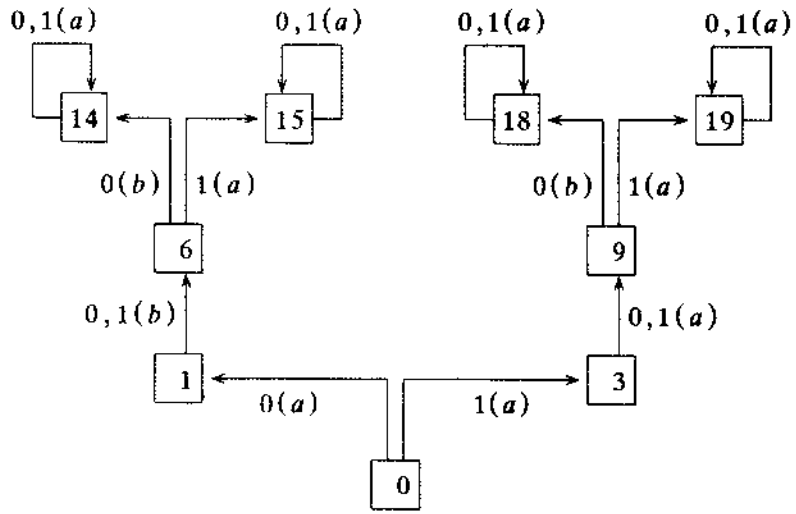


Рис.29.5

В классе всех бесконечных игр существуют такие игры, в которых никакая из сторон не имеет выигрывающей стратегии. В случае конечно автоматных игр положение иное: здесь всегда одна из сторон имеет выигрывающую стратегию и даже конечно автоматную. К ее построению и переходим.

29.4. Униформизация конечно автоматных языков

Проблема униформизации конечно автоматного сверхъязыка $L = (X \times Y)^{\infty}$, представимого макроавтоматом A, может быть интерпретирована как задача о существовании и построении выигрывающей конечно автоматной стратегии, которая является оператором без предвосхищения $y = \mathbb{O}''(x)$ для черных, или оператором без предвосхищения $x = \mathbb{O}'(y)$ для белых, ибо в конечно автоматной игре всегда одна из сторон имеет выигрывающую конечно автоматную стратегию.

29.4.1. Порядковые векторы и порядковые стратегии

Пусть $A = (X \times Y, Q, q_0, T, F)$ есть макроавтомат, представляющий сверхъязык $L = L(A, q_0, F)$; $Q = \{q_0, q_1, \ldots, q_n\}$ есть множество состояний в конечно автоматной игре $G(A)$, соответствующей сверхъязыку L. Возможны $(n+1)!$ перестановок состояний из

Q; назовем каждую такую перестановку $(q_{i_0}, q_{i_1}, \ldots, q_{i_n})$ *порядковым* вектором и будем обозначать его через $(i_0, i_1, \ldots, i_n)$. Рассмотрим некоторую партию

$$\begin{array}{l} x(0)\ x(1)\ \ldots\ x(t)\ \ldots \\ y(1)\ y(2)\ \ldots\ y(t)\ \ldots \end{array} \qquad (29.1)$$

в конечно автоматной игре $G(A)$. Партии (29.1) соответствует последовательность состояний

$$q(0)\ q(1)\ \ldots\ q(t)\ \ldots \qquad (29.2)$$

Наряду с последовательностью (29.2) сопоставим партии (29.1) последовательность порядковых векторов

$$\mathbf{q}(0)\ \mathbf{q}(1)\ \ldots\ \mathbf{q}(t)\ \ldots, \qquad (29.3)$$

индуктивно определяемую следующим образом:

1) $\mathbf{q}(0) = (0, 1, \ldots, n)$;

2) если $\mathbf{q}(t) = (i_0, i_1, \ldots, i_k, i_{k+1}, \ldots, i_n)$ и $q(t+1) = q_k$, то $\mathbf{q}(t+1) = (i_k, i_1, i_2, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n)$.

Ниже следует пример последовательности состояний и соответствующей последовательности порядковых векторов.

q_0	q_2	q_2	q_2	q_1	q_0	q_1	q_1	q_3	q_4	q_1	q_2	q_4
0	2	2	2	1	0	1	1	3	4	1	2	4
1	0	0	0	2	1	0	0	1	3	4	1	2
2	1	1	1	0	2	2	2	0	1	3	4	1
3	3	3	3	3	3	3	3	3	0	0	3	3
4	4	4	4	4	4	4	4	4	2	2	0	0

Справедлива следующая лемма.

Лемма 1. Если $\mathbf{q}(t) = (\ldots, i, \ldots, j, \ldots)$, $\mathbf{q}(t+n) = (\ldots, j, \ldots, i, \ldots)$, то в промежутках между t и $t+n$ вершина j (состояние q_j) посещалась, т.е. занимала в векторе первое место, по крайней мере один раз.

Пусть макросостояние $Q' = \{q_0, q_1, \ldots, q_s\} \subseteq Q$. Пусть $n(t)$ есть номер позиции в векторе $\mathbf{q}(t)$, которая в следующем векторе $\mathbf{q}(t+1)$ стала первой. Например,

$\mathbf{q}(t) = (q_2, q_1, q_4, q_0, q_3, q_5)$, $n(t) = 3$,
$\mathbf{q}(t+1) = (q_4, q_2, q_1, q_0, q_3, q_5)$.

Пусть в последовательности (29.3), которую запишем в виде

$$q(0)\ q(1)\ \ldots\ q(t')\ q(t'+1)\ \ldots\ q(t'')\ q(t''+1)\ \ldots,$$

предельное макросостояние Q' формируется, начиная с момента t'. Пусть t'' выбрано так, что множество состояний, встречающихся в слове $q(t')q(t'+1)\ldots q(t'')$, совпадает с множеством Q'. Тогда

$\forall t \geqslant t'$:

1) в порядковом векторе $\mathbf{q}(t)$ первые $s+1$ позиций заняты состояниями предельного макросостояния Q';

2) $n(t) = s$; кроме того, $(\exists^{\infty}t)(n(t) = s)$, где $\exists^{\infty}t$ означает: существует бесконечно много t.

В последовательности $t''+1$, $t''+2,\ldots$ существует подпоследовательность натуральных чисел $m_0 < m_1 < \ldots$, для которой $n(m_0) = n(m_1) = \ldots = s$, и между m_i и m_{i+1} в первой позиции вектора $\mathbf{q}(t)$ побывали все состояния из Q' (и только они). Так как в последовательности $m_0 < m_1 < \ldots$ встречается не более $(n+1)!$ векторов, то один из них, например, $\mathbf{q} = (i_0,i_1,\ldots,i_s,i_{s+1},\ldots,i_n)$ в последовательности порядковых векторов $\mathbf{q}(m_0),\mathbf{q}(m_1),\ldots$ встречается бесконечно много раз, допустим, что на местах $t_0 < t_1 < \ldots$, составляющих подпоследовательность в последовательности $m_0 < m_1 < \ldots$. Показана справедливость следующей леммы.

Лемма 2. Для макросостояния $Q' = \{q_0,q_1,\ldots,q_s\}$ существует последовательность натуральных чисел $t_0 < t_1 < \ldots$, для которой

1) $\mathbf{q}(t_0) = \mathbf{q}(t_1) = \ldots = (i_0,i_1,\ldots,i_s,\ldots,i_n)$;

2) $n(t_0) = n(t_1) = \ldots = s$ (а потому $\mathbf{q}(t_0+1) = \mathbf{q}(t_1+1) = \ldots = (i_s,i_0,i_1,\ldots,i_{s-1},i_{s+1},\ldots,i_n)$);

3) $\forall t \geqslant t_0 \quad n(t) = s$;

4) $\forall t \geqslant t_0$ в векторе $\mathbf{q}(t)$ первые s позиций заняты элементами из Q'; причем между всякими t_i и t_{i+1} в первой позиции порядковых векторов $\mathbf{q}(t)$ побывали все состояния из Q'.

Стратегия белых, т.е. оператор без предвосхищения $x = \mathbb{O}(y)$, называется *порядковой*, если ход белых $x(t)$ зависит только от порядкового вектора $\mathbf{q}(t)$, именно, $x(t) = F_x(\mathbf{q}(t))$ для некоторой функции F_x. Стратегия черных называется *порядковой*, если ход черных $y(t)$ зависит только от $\mathbf{q}(t)$ и $x(t)$, именно, $y(t) = F_y(\mathbf{q}(t),x(t))$ для некоторой функции F_y.

Порядковые стратегии являются конечно автоматными стратегиями с числом состояний не более $(n+1)!$. В самом деле, пусть в конечно автоматной игре отображение $G(\mathbf{q}(t),x(t),y(t))$ есть функция, указывающая порядковый вектор, в которой пара ходов $(x(t),y(t))$ переводит порядковый вектор $\mathbf{q}(t)$. По заданной конечно автоматной игре функция G может быть построена эффективно. Тогда для белых

$$\mathbf{q}(0) = (0,1,\ldots,n) = \mathbf{q}_0;$$
$$\mathbf{q}(t+1) = G(\mathbf{q}(t),F_x(\mathbf{q}(t)),y(t)) = G_x(\mathbf{q}(t),y(t));$$
$$x(t) = F_x(\mathbf{q}(t));$$

а для черных

$$\mathbf{q}(0) = (0,1,\ldots,n) = \mathbf{q}_0;$$
$$\mathbf{q}(t+1) = G(\mathbf{q}(t),x(t),F_y(\mathbf{q}(t),x(t))) = G_y(\mathbf{q}(t),x(t));$$
$$y(t) = G_y(\mathbf{q}(t),x(t)).$$

Пусть $\mathbf{Q}$ есть множество всех порядковых векторов. Тогда автоматы $A_x = (X,Y,\mathbf{Q},\mathbf{q}_0,G_x,F_x)$, $A_y = (X,Y,\mathbf{Q},\mathbf{q}_0,G_y,F_y)$ вычисляют (реализуют) стратегии $\mathbb{O}'$ и $\mathbb{O}''$ белых и черных соответственно.

Заметим, что в рассмотренном случае выходные функции $x(t) = F_x(\mathbf{q}(t)),y(t) = F_y(\mathbf{q}(t),x(t))$, т.е. в порядковых стратегиях функция $x(t)$ не зависит от входа, поэтому построенные автоматы A_x и A_y являются автоматами Мура и Мили соответственно.

29.4.2. Теоремы о порядковых стратегиях

Теорема 1. Если в конечно автоматной игре, осуществляемой макроавтоматом $A = (X\times Y,Q,q_0,T,F)$, стратегия черных $y = \mathbb{O}''(x)$ бьет порядковую стратегию белых $x = \hat{\mathbb{O}}'(y)$, то черные имеют порядковую стратегию $\hat{\mathbb{O}}''$, которая бьет $\hat{\mathbb{O}}'$.

Доказательство. Пусть в данной конечно автоматной игре пара стратегий $(\hat{\mathbb{O}}',\mathbb{O}'')$ определяет партию

$$\begin{array}{l} x(0)\ x(1)\ \ldots\ x(t)\ \ldots \\ y(0)\ y(1)\ \ldots\ y(t)\ \ldots \\ q(0)\ q(1)\ \ldots\ q(t)\ \ldots \\ \mathbf{q}(0)\ \mathbf{q}(1)\ \ldots\ \mathbf{q}(t)\ \ldots, \end{array} \qquad (29.4)$$

выигрышную для черных. Тогда в этой партии предельное макросостояние $\lim\ q(t) = Q' = \{q_0,q_1,\ldots\} \in F$. Ходы противников в партии (29.4) в момент t

$$x(t) = F_x(\mathbf{q}(t)) \ , \quad y(t) = \mathbb{O}''(x(t)) \qquad (29.5)$$

таковы, что белые действуют порядково, т.е. ход белых зависит только от порядкового состояния, а ход черных выбирается, вообще говоря, не порядково. По (непорядковой)

выигрышной стратегии черных $\mathbb{O}''$ построим порядковую стратегию $\hat{\mathbb{O}}''$, тоже выигрышную для черных. Пусть $t_0 < t_1 < \ldots < t_s < \ldots$ есть существующая по лемме 2 последовательность, для которой

1) предельное макросостояние Q' $\forall t \geqslant t_0$ уже определилось;

2) $\mathbf{q}(t_0) = \mathbf{q}(t_1) = \ldots = (i_0, i_1, \ldots, i_s, i_{s+1}, \ldots, i_n)$;

3) $n(t_1+1) = n(t_2+1) = \ldots = s$, т.е. $\mathbf{q}(t_0+1) = \mathbf{q}(t_2+1) = \ldots = (i_s, i_0, i_1, \ldots, i_{s-1}, i_{s+1}, \ldots, i_n)$;

4) в порядковых векторах на первом месте между t_i и t_{i+1} побывали все состояния из Q' (и только они), т.е. первые компоненты (позиции) порядковых векторов между t_i и t_{i+1} дают в точности множество Q'.

Рассмотрим в партии $(\hat{\mathbb{O}}', \mathbb{O}'')$ следующее начало:

$$\begin{array}{l} x(0)\ x(1)\ \ldots\ x(t0)\ x(t0+1)\ \ldots\ x(t1)\ x(t1+1) \\ y(0)\ y(1)\ \ldots\ y(t0)\ y(t0+1)\ \ldots\ y(t1)\ y(t1+1) \\ q(0)\ q(1)\ \ldots\ q(t0)\ q(t0+1)\ \ldots\ q(t1)\ q(t1+1) \\ \mathbf{q}(0)\ \mathbf{q}(1)\ \ldots\ \mathbf{q}(t_0)\ \mathbf{q}(t_0+1)\ \ldots\ \mathbf{q}(t_1)\ \mathbf{q}(t_1+1) \end{array} \qquad (29.6)$$

Пусть $\mathbf{q}$ есть некоторый порядковый вектор в (29.6). Так как стратегия белых $\hat{\mathbb{O}}'$ порядковая, то при появлении $\mathbf{q}$ белые действуют всегда одинаково: $x' = F_x(\mathbf{q})$. Ход черных при повторных появлениях вектора $\mathbf{q}$ не обязан быть одинаковым, ибо стратегия черных $\mathbb{O}''$ не является порядковой. Построим порядковую стратегию черных $\hat{\mathbb{O}}''$ следующим образом. Полагаем вектор $\mathbf{r}(0) = \mathbf{q}(0)$. Проверяем, есть ли в (29.6) другая пара $(\mathbf{q}, x)$, равная $(\mathbf{r}(0), x(0))$. Если нет, то ход черных $y(0)' = y(0)$. Если да, то пусть n_0 есть ближайший слева к t_1 момент, для которого пара $(\mathbf{r}(0), x(0)) = (\mathbf{q}(n_0), x(n_0))$. Полагаем тогда $y'(0) = y(n_0) = F_y(\mathbf{r}(0), x(0))$.

Далее, пусть $\mathbf{r}(1) = F(\mathbf{r}(0), x(0), y'(0))$. Проверяем, есть ли в (29.6) другая пара $(\mathbf{q}, x)$, равная $(\mathbf{r}(1), x(1))$. Если нет, то $y'(1) = y(1)$. Если да, то пусть n_1 есть ближайший слева к t_1 момент, для которого $(\mathbf{r}(1), x(1)) = (\mathbf{q}(n_1), x(n_1))$. Тогда полагаем $y'(1) = y(n_1) = F_y(\mathbf{r}(1), x(1))$. И так далее до t_1. Во всех остальных случаях значение $F_y(\mathbf{q}, x)$ выбираем произвольно. Порядковая стратегия черных $\hat{\mathbb{O}}''$ построена. Партия $(\hat{\mathbb{O}}', \hat{\mathbb{O}}'')$ имеет вид

$$\begin{array}{l} x(0)\ x(1)\ \ldots\ x(t_0)\ \ldots\ x(t_1)\ \ldots \\ y(0)\ y(1)\ \ldots\ y(t_0)\ \ldots\ y(t_1)\ \ldots \\ r(0)\ r(1)\ \ldots\ r(t_0)\ \ldots\ r(t_1)\ \ldots \\ \mathbf{r}(0)\ \mathbf{r}(1)\ \ldots\ \mathbf{r}(t_0)\ \ldots\ \mathbf{r}(t_1)\ \ldots \end{array} \qquad (29.7)$$

Для построенной партии (29.7) последовательность состояний, проходимых автоматом A, и соответствующая последовательность порядковых векторов $\mathbf{r}(t)$, в которых первая компонента есть $r(t)$, вообще говоря, отличается от таковой в партии (29.5); при этом $\mathbf{r}(t) = \mathbf{q}(t)$ до тех пор, пока в (29.6) не встретилась ситуация $(\mathbf{q}, x)$, которая встречалась в (29.6) позже.

Покажем, что последовательность (29.7) порядковых векторов $\mathbf{r}(t)$ имеет начальный отрезок векторов

$$\mathbf{r}(0)\ \mathbf{r}(1)\ \ldots\ \mathbf{r}(m)\ \ldots\ \mathbf{r}(m+n), \qquad (29.8)$$

первые позиции в которых на отрезке $[m, m+n]$ дадут в точности предельное макросостояние Q'. Другими словами, мы покажем, что в последовательности (29.7) на отрезке $[m, m+n]$ в первой позиции побывали все состояния предельного макросостояния Q' и только они. Пусть $\forall p \in [0, t_1]$

$$a_p = (\max\ t)_{t<t_1}(\mathbf{r}(p) = \mathbf{q}(t)).$$

Так как $\mathbf{r}(0) = \mathbf{q}(0)$, то величина

$$a_0 = (\max\ t)_{t<t_1}(\mathbf{r}(0) = \mathbf{q}(t))$$

существует, причем $0 \leqslant a_0 < t_1$ и

$$\mathbf{r}(0) = \mathbf{q}(a_0)\ \&\ (\forall t)\ (a_0 < t \leqslant t_1 \to \mathbf{r}(0) \neq \mathbf{q}(t)).$$

Если $a_0 = t_1$, то полагаем $m = 1$. Если $a_0 < t_1$, то по построению порядковой стратегии $\hat{\mathbb{O}}''$ ввиду

$$\mathbf{r}(0) = \mathbf{q}(a_0),\ x(0) = x(a_0),\ y'(0) = y(a_0)$$

вектор $\mathbf{r}(1) = \mathbf{q}(a_0+1)$, и потому a_1 существует; при этом

$$0 \leqslant a_0 < a_1 \leqslant t_1 \text{ и}$$

$$\mathbf{r}(1) = \mathbf{q}(a_1)\ \&\ (\forall t)\ (a_1 < t < t_1 \to \mathbf{r}(1) \neq \mathbf{q}(t)).$$

Если $a_1 = t_1$, то полагаем $m = t_1$. Если $a_1 < t_1$, то указанное выше построение продолжается. Процесс выделения возрастающей цепи $0 < a_0 < a_1 < \ldots$ необходимо оборвется на некотором $a_m = t_1$ и тогда $\mathbf{r}(m) = \mathbf{q}(t_1) = (i_0, i_1, \ldots, i_s, i_{s+1}, \ldots, i_n)$.

Так как $\mathbf{q}(t_1) = \mathbf{q}(t_0)$, то $\mathbf{r}(m) = \mathbf{q}(t_0) = \mathbf{q}(t_1)$.

Так как $x(m) = x(t_1)$, $y'(m) = y(t_1)$, то

$$\mathbf{r}(m+1) = \mathbf{q}(t_1+1) = \mathbf{q}(t_0+1) = (i_s, i_0, \ldots, i_{s-1}, i_{s+1}, \ldots, i_n).$$

Поэтому a_{m+1} существует, $t_0 < a_{m+1} \leqslant t_1$ и

$$\mathbf{r}(m+1) = \mathbf{q}(a_{m+1}) \;\&\; (\forall t)\ (a_{m+1} < t < t_1 \to \mathbf{r}(m+1) \neq \mathbf{q}(t)).$$

Если $a_{m+1} = t_1$, то $n = 1$. Если $a_{m+1} < t_1$, то ввиду

$$\mathbf{r}(m+1) = \mathbf{q}(a_{m+1}),\ x(m+1) = x(a_{m+1}),\ y'(m+1) = y(a_{m+1})$$

вектор $\mathbf{r}(m+2) = \mathbf{q}(a_{m+1}+1)$; поэтому a_{m+2} существует и для него $\mathbf{r}(m+2) = \mathbf{q}(a_{m+2}) \;\&\; (\forall t)(a_{m+1} < a_{m+2} < t_1 \to \mathbf{r}(m+2) \neq \mathbf{q}(t))$. Продолжая таким же образом, получим последовательность

$$t_0 < a_{m+1} < a_{m+2} < \ldots < a_{m+n} = t,$$

причем $\mathbf{r}(m+n) = \mathbf{q}(t_1) = (i_0, i_1, \ldots, i_s, i_{s+1}, \ldots, i_n)$.

Так как $\mathbf{r}(m+1) = (i_s, i_0, i_1, \ldots, i_{s-1}, i_{s+1}, \ldots, i_n)$, то по лемме 1 между $m+1$ и $m+n$ первые позиции векторов $\mathbf{r}(m+1), \ldots, \mathbf{r}(m+n)$ дадут в точности предельное макросостояние Q'.

Теорема 1 доказана.

Аналогично доказывается следующая теорема.

Теорема 2. Если стратегия белых $\mathbb{O}'$ бьет порядковую стратегию черных $\hat{\mathbb{O}}''$, то белые имеют порядковую стратегию $\hat{\mathbb{O}}'$, которая бьет стратегию черных $\hat{\mathbb{O}}''$.

Теорема 3. Пусть в конечно автоматной бесконечной игре G белые и черные прибегают только к порядковым стратегиям. Тогда одна из сторон имеет выигрышную порядковую стратегию, (которая бьет любую порядковую стратегию противника); эта выигрышная стратегия по игре G может быть эффективно построена.

Доказательство. Пусть оба игрока прибегают только к порядковым стратегиям. Тогда последовательности порядковых векторов, соответствующих сыгранным партиям, будут периодическими с длиной предпериода и периода в сумме не более $(n+1)!$. И предпериод, и период можно эффективно найти и по периоду найти предельное макросостояние (первые позиции в порядковых векторах периода). Если поменяем порядковую стратегию, то сыгранная партия поменяется, но выигрыш той или иной стороны определяется партией длины не более $(n+1)!$. То есть бесконечная игра G эквивалентна конечной игре G' с длительностью каждой партии не более $(n+1)!$. Правила игры G' таковы. Пусть уже сделаны t ходов

$$\begin{array}{l} x(0)\ x(1)\ \ldots\ x(t-1) \\ y(0)\ y(1)\ \ldots\ y(t-1) \\ \mathbf{q}(0)\ \mathbf{q}(1)\ \ldots\ \mathbf{q}(t-1)\ \mathbf{q}(t), \end{array}$$

после которых возник порядковый вектор $\mathbf{q}(t)$. Если этот вектор ранее не встречался, то игроки могут выбирать любой ход из своих алфавитов X и Y. Если вектор $\mathbf{q}(t)$ ранее возникал, т.е. $\mathbf{q}(t) = \mathbf{q}(t')$ при некотором $t' < t$, то белые и черные обязаны повторять ходы: $x(t) = x(t')$, $y(t) = y(t')$. Партия прекращается при первом повторении порядкового вектора. При этом в концевой вершине дерева игры ставим знак +, если предельное макросостояние между этими повторениями благоприятно для белых, и знак – в противном случае.

Далее построенную конечную игру G' анализируем на предмет определения выигрышной стороны и построения для нее конечного автомата, вычисляющего для нее выигрышную стратегию, как это делали для конечных игр.

Теорема 4. (Основная о конечно автоматных играх.) Во всякой игре с конечным числом состояний одна из сторон имеет выигрышную конечно автоматную стратегию. Существует алгоритм, который по любой заданной автоматной игре:

1) выясняет, какая из сторон имеет выигрышную стратегию;

2) строит для выигрышной стороны одну из выигрывающих конечно автоматных стратегий.

Доказательство. Алгоритм теоремы 3 строит порядковую стратегию $\hat{\mathbb{O}}$, которая бьет любую порядковую стратегию $\hat{\mathbb{O}}'$ противника. Но эта же стратегия $\hat{\mathbb{O}}$ бьет и любую другую стратегию $\mathbb{O}'$ противника (не обязательно порядковую). В самом деле, если бы это было не так, т.е. если бы некоторая стратегия $\mathbb{O}'$ била $\hat{\mathbb{O}}$, то по теореме 1 и некоторая порядковая стратегия $\hat{\mathbb{O}}'$ била бы $\hat{\mathbb{O}}$, чего нет.

Следствие. Если в конечно автоматной игре одна из сторон имеет какую-нибудь выигрывающую стратегию, то она имеет и конечно автоматную выигрывающую стратегию.

Теорема 5. (Перефразировка теоремы 4). Пусть дан произвольный сверхъязык L в алфавите $X \times Y$, который представим в некотором макроавтомате. Тогда либо сверхъязык L униформизуется конечно автоматным оператором $y = \mathbb{O}'(x)$, либо его дополнение $(X \times Y)^{\infty} - L$ униформизуется конечно автоматным операто-

ром с задержкой $x = \mathbb{O}(y)$. Существует алгоритм, который по макроавтомату, представляющему сверхъязык L, выясняет, какая ситуация в указанной альтернативе имеет место, и строит конечный автомат, реализующий соответствующий оператор.

Следствие. Если существует оператор без предвосхищения, униформизующий конечно автоматный сверхъязык L, то существует и конечно автоматный оператор, униформизующий сверхъязык L. В частности, если график некоторого оператора без предвосхищения представим в конечном автомате, то этот оператор реализуем в конечном автомате.

Замечание. Пусть бесконечная игра осуществляется с помощью макроавтомата $A = (X \times Y, Q, q_0, T, F)$, где $F \subseteq P(Q)$ есть множество макросостояний, благоприятных для белых. Если выигрывают белые, то белые имеют выигрывающую стратегию $x = \mathbb{O}'(y)$, униформизующую сверхъязык $L(A, q_0, F)$, ибо стратегия $\mathbb{O}'$ "загоняет" фишку в предельное макросостояние из F. Если выигрывают черные, то черные имеют выигрывающую стратегию $y = \mathbb{O}''(x)$, униформизующую дополнительный сверхъязык, ибо стратегия $\mathbb{O}''$ "загоняет" фишку в предельное макросостояние из $P(Q)$, лежащее вне F.

29.4.3. Пример построения выигрывающего автомата

Пусть конечный автомат для бесконечной игры задается граф-схемой, приведенной на рис.29.6. $X = \{a,b\}$; $Y = \{c,d\}$. Макросостояния $\{0\}$, $\{1\}$; $\{0,1,2\}$ благоприятны для белых; остальные возможные для этого автомата макросостояния $\{2\}$, $\{0,1\}$, $\{1,2\}$ благоприятны для черных. Начальное состояние есть 0. Строим конечную игру (рис.18.7), вершины которой помечены порядковыми векторами (i_0, i_1, i_2). Построение конечного дерева на некотором пути прекращается при появлении на этом пути пары одинаковых порядковых векторов. Множество состояний между одинаковыми порядковыми векторами на одном и том же пути составляет предельное макросостояние.

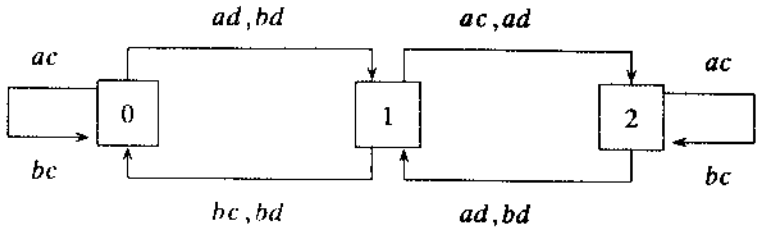


Рис.29.6

Просматривая дерево игры на рис.29.7, видим, что черные имеют выигрывающую стратегию (рис.29.8). Начальное состояние 012 расположено в корне дерева. Граф-схема автомата (Мили), ее вычисляющую, приведена на рис.29.9. В скобках указаны выходы автомата. Входные символы автомата есть ходы белых; выходные (символы в скобках) есть ответы черных.

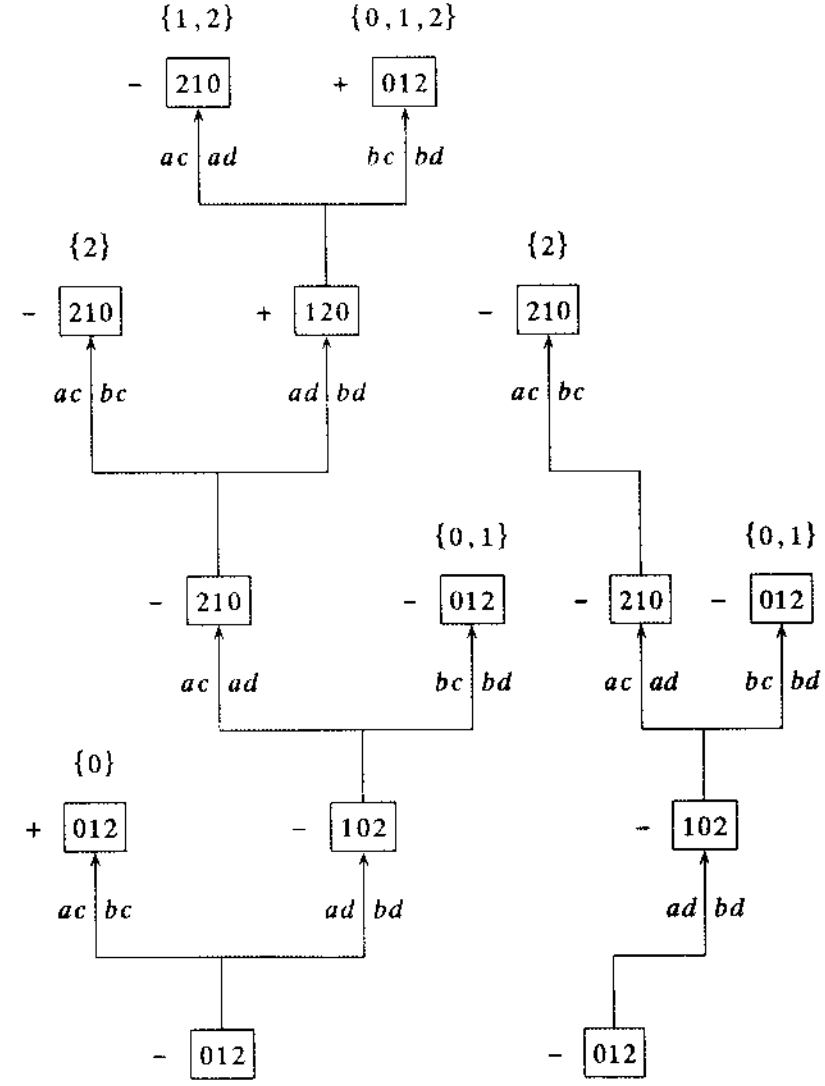


Рис.29.7 Рис.29.8

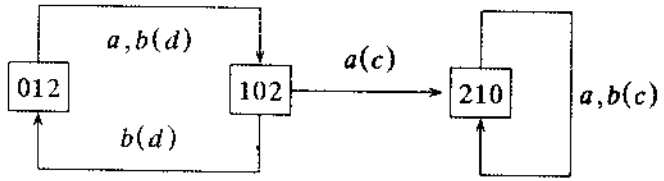


Рис.29.9

Если для того же самого автомата, граф-схема которого изображена на рис.29.6, к числу макросостояний, благоприятных для белых, добавить макросостояние {2}, то получим игру (рис.29.10), в которой белые имеют выигрывающую стратегию (рис.29.11). Граф-схема автомата (Мура), вычисляющего выиг-

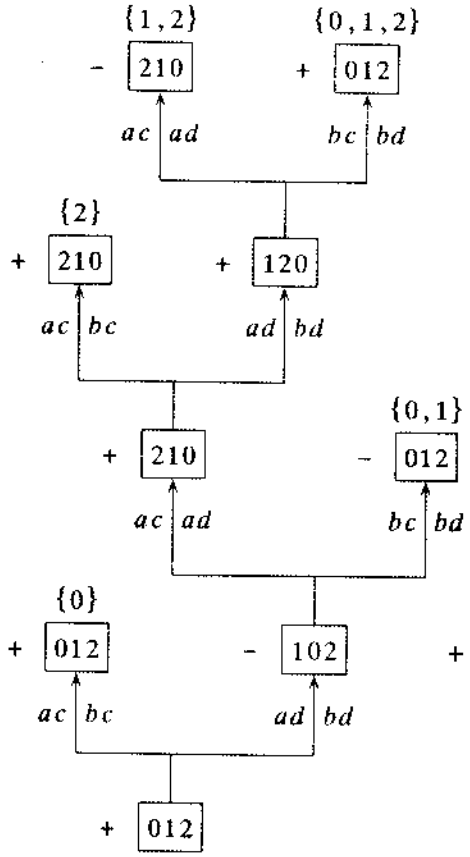


Рис.29.10

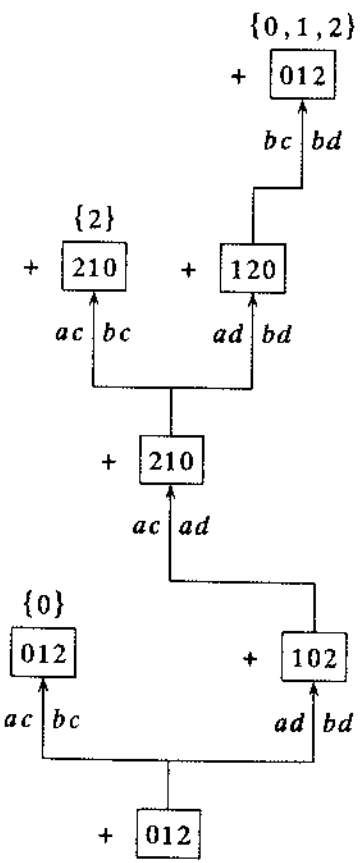


Рис.29.11

рывающую стратегию белых, приведена на рис.29.12; начальное состояние 012; состояния помечены ходами белых; стрелки переходов – ходами черных.

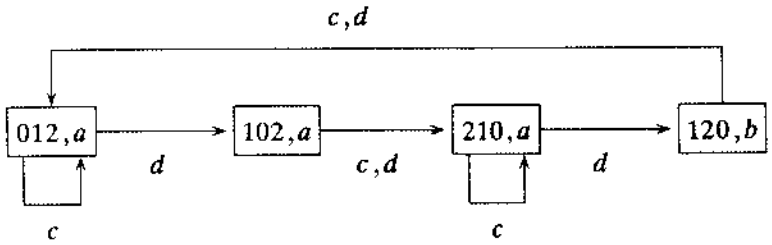


Рис.29.12

30. МОНАДИЧЕСКАЯ ЛОГИКА

30.1. Логика одноместных предикатов

Опишем монадическую логику – интерпретируемый формализм, связанный с логикой одноместных предикатов (ЛОП) второго порядка на натуральных числах с функцией следования (т.е. прибавления единицы) $y = x + 1$ на натуральном ряду.

Формальные символы.

1. 0,1,2,... есть символы натуральных чисел.

2. $x,y,z,\ldots$ есть символы предметных переменных, пробегающих натуральный ряд, возможно с индексами.

3. $'$ есть штрих, символ функции следования на натуральном ряду, $x' = x+1$.

4. $a_1,a_2,\ldots,a_n$ есть символы конечного алфавита $\Sigma = \{a_1, \ldots,a_n\}$.

5. $X,Y,Z,\ldots$ есть символы предикатных переменных (возможно с индексами), аргументы которых пробегают натуральный ряд, а значения – множество символов алфавита Σ. Правильнее было бы сказать, что $X,Y,Z,\ldots$ есть символы для конечнозначных функций, с аргументами, пробегающими натуральный ряд, и со значениями в множестве Σ. Сохраним все же традиционную предикатную терминологию.

6. &, $\vee$, $\neg$, $\rightarrow$, $\exists$, $\forall$ есть логические символы: конъюнкция, дизъюнкция, отрицание, импликация, квантор существования, квантор общности.

7. = есть знак равенства.

8. (,) есть скобка левая, запятая, скобка правая.

Термы.

1. Символ натурального числа есть терм.
2. Символ предметной переменной есть терм.
3. Если t есть терм, то t' есть терм.

Замечание. Вместо t', t'', t''',... будем писать также t+1, t+2, t+3,....

Формулы.

1. Если X есть предикатный символ, а t есть терм, содержащий предметную переменную или без нее, и a есть буква из Σ, то $X(t) = a$ есть атомарная формула с глубиной построения 1, в которой предикатная переменная X и предметная переменная x, если она есть, являются свободными переменными.

2. Если A и B есть формулы с глубиной построения r и s соответственно, то $(A \,\&\, B)$, $(A \vee B)$, $(A \to B)$ есть формулы с глубиной построения $\max(r,s)+1$, свободные переменные которых есть вместе взятые свободные переменные формул A и B. Выражение $(\neg A)$ тоже формула с глубиной построения r+1. Свободные переменные формулы $(\neg A)$ есть свободные переменные формулы A.

3. Если $A(x)$ есть формула со свободной предметной переменной x и с глубиной построения r, то $(\exists x)A(x)$, $(\forall x)A(x)$ есть формулы с глубиной построения r+1, переменная x в которой связана (квантором). Свободные переменные формул $(\exists x)A(x)$, $(\forall x)A(x)$ есть все свободные переменные формулы $A(x)$, кроме x.

4. Если $A(X)$ есть формула со свободной предикатной переменной X и с глубиной построения r, то $(\exists X)A(X)$, $(\forall X)A(X)$ есть формулы с глубиной построения r+1, переменная X в которых связана (квантором). Свободные переменные формул $(\exists X)A(X)$, $(\forall X)A(X)$ есть все свободные переменные формулы $A(X)$, кроме X.

Формула называется *замкнутой*, если она не имеет свободных переменных. Всякая замкнутая формула содержательно представляет собой высказывание о натуральном ряде и о его подмножествах, истинное или ложное.

Построенный интерпретируемый формализм составляет монадическую логику (МЛ) натуральных чисел, или логику одноместных предикатов (ЛОП) второго порядка натуральных чисел с функцией следования на натуральном ряду.

Вместо $X(t) = a$ для краткости будем писать также $X^a(t)$. Предикат $X(x)$, определенный на натуральном ряду, можно рассматривать как сверхслово $X(0)X(1)X(2)\ldots$ в алфавите Σ.

Пусть $A(X_1,X_2,\ldots,X_k)$ есть формула ЛОП без свободных

предметных переменных и со свободными предикатными переменными $X_1,X_2,\ldots,X_k$. Набор фиксированных предикатов $X_1,X_2,\ldots,X_k$ (назовем этот набор k-предикатом)

$$\begin{array}{l} X_1(0)X_1(1)\ \ldots\ X_1(t)\ \ldots \\ X_2(0)X_2(1)\ \ldots\ X_2(t)\ \ldots \\ \qquad\ldots \\ X_k(0)X_k(1)\ \ldots\ X_k(t)\ \ldots, \end{array}$$

написанных один под другим, можно рассматривать как "толстое" слово "толщины" k в алфавите $\Sigma\times\Sigma\times\ldots\times\Sigma$ (k раз). Тогда множество истинности формулы $A(X_1,\ldots,X_k)$ есть некоторый сверхъязык в алфавите $\Sigma\times\Sigma\ldots\times\Sigma$ (k раз).

Пусть $A(X,X_1,\ldots,X_k)$ есть формула в ЛОП без свободных предметных переменных, и $X,X_1,\ldots,X_k$ есть полный список ее свободных предикатных переменных. Пусть $\lceil A\rceil = \{(X,X_1,\ldots,X_k)$: высказывание $A(X,X_1,\ldots,X_k)$ истинно$\}$ есть множество истинности формулы A. Пусть формула $B(X_1,\ldots,X_k) = (\exists X)A(X,X_1,\ldots,X_k)$. Тогда множество истинности формулы B может быть получено из множества истинности формулы A вычеркиванием (удалением) во всех наборах $(X,X_1,\ldots,X_k)$ из $\lceil A\rceil$ первой компоненты X. Если, например, в алфавите $\Sigma = \{0,1,2\}$ 4-предикат

$$\begin{array}{l} X = 1\,0\,1\,1\,2\,2\,0\,1\,2\,3\ \ldots \\ X_1 = 2\,1\,0\,1\,1\,2\,0\,0\,1\,1\ \ldots \\ X_2 = 0\,1\,2\,2\,1\,2\,0\,1\,2\,0\ \ldots \\ X_3 = 1\,2\,2\,1\,1\,0\,1\,2\,0\,1\ \ldots \end{array}$$

принадлежит множеству истинности некоторой формулы $A(X,X_1,X_2,X_3)$, то 3-предикат

$$\begin{array}{l} X_1 = 2\,1\,0\,1\,1\,2\,0\,0\,1\,1\ \ldots \\ X_2 = 0\,1\,2\,2\,1\,2\,0\,1\,2\,0\ \ldots \\ X_3 = 1\,2\,2\,1\,1\,0\,1\,2\,0\,1\ \ldots \end{array}$$

принадлежит множеству истинности формулы $B(X_1,X_2,X_3) = (\exists X)A(X,X_1,X_2,X_3)$. Можно считать, что этот 3-предикат получается из исходного 4-предиката поразрядной заменой, т.е. проекцией $\{0,1,2\}^4 \to \{0,1,2\}^3$, при которой $(b_0,b_1,b_2,b_3)^T \to (b_1,b_2,b_3)^T$, где каждое b_i лежит в Σ, а буква T означает транспонирование.

Если $B(X_1,\ldots,X_k)$ есть $(\exists x)A(X,X_1,\ldots,X_k)$, то можно считать, что множество истинности формулы B получается из множества истинности формулы A проектированием последнего на оси $1,2,\ldots,k$ с помощью проекции $\Sigma^{k+1} \to \Sigma^k$, при которой $(b_0,$

$b_1,\ldots,b_k)^T \to (b_1,\ldots,b_k)^T$, где все b_i лежат в Σ. Таким образом, $\lceil B \rceil = \mathrm{Pr}_{1,2,\ldots,k}(\lceil A \rceil)$. Переменная X в формуле A может стоять на любом аргументном месте формулы A.

30.2. Выразимость в ЛОП

Покажем, что монадическая логика второго порядка есть довольно гибкий инструмент для записи работы автоматов, источников, макроисточников, регулярных и общерегулярных языков. Анализ указанных объектов с помощью ЛОП довольно прост и естественен. Синтез осуществляется сложнее. Рассмотрим его ниже.

В ЛОП выразимы формулами следующие предикаты.

$x = y \longleftrightarrow (\forall X)(X^a(x) \to X^a(y))$, $a \in \Sigma$.

$x = 0 \longleftrightarrow (\forall y)(x \neq y+1)$.

$x = 1 \longleftrightarrow (\forall y)(x \neq y+2) \;\&\; x \neq 0$.

$x = 2 \longleftrightarrow (\forall y)(x \neq y+3) \;\&\; x \neq 0 \;\&\; x \neq 1$.

...

$x = k \longleftrightarrow (\forall y)(x \neq y+k+1) \;\&\; x \neq 0 \;\&\; \ldots \;\&\; x \neq k-1$.

Заметим, что

$$X(t) \neq a \longleftrightarrow \Big(\bigvee_{b \in \Sigma - (a)} X(t) = b\Big); \qquad (30.1)$$

$$X^a(x+1) \longleftrightarrow (\exists y)(x+1=y \;\&\; X^a(y)).$$

Здесь переменная y в $X(y)$ не имеет штриха.

$x \leqslant y \longleftrightarrow (\forall X)(X^a(x) \;\&\; (\forall x)(X^a(x) \to X^a(x')) \to X^a(y))$, $a \in \Sigma$.

$x < y \longleftrightarrow x < y \;\&\; x \neq y$.

$x \geqslant y \longleftrightarrow y \leqslant x$.

$x > y \longleftrightarrow y < x$.

Введем ограниченные кванторы:

$(\exists y)_{\leqslant x}\, A(y) \longleftrightarrow (\exists y)(y \leqslant x \;\&\; A(y))$.

$(\forall y)_{\leqslant x}\, A(y) \longleftrightarrow (\forall y)(y \leqslant x \to A(y))$.

$(\exists t)_x^y\, A(t) \longleftrightarrow (\exists t)(x \leqslant t < y \;\&\; A(t))$.

$(\forall t)_x^y\, A(t) \longleftrightarrow (\forall t)(x \leqslant t < y \to A(t))$.

Введем также неограниченные кванторы:

$(\exists^\infty x)\, A(x) \longleftrightarrow (\forall y)(\exists x)_{>y}\, A(x)$.

$(\forall^\infty x)\, A(x) \longleftrightarrow (\exists y)(\forall x)_{>y}\, A(x)$.

Определим формулами ЛОП следующие предикаты:

$$a \in \lim X(x) \longleftrightarrow (\exists^\infty x)(X(x) = a);$$

$$\lim X(x) = \{a_{i_1},\ldots,a_{i_s}\} \longleftrightarrow \Big(\&_{k=1}^{s} (a_{ik} \in \lim X(x)) \;\&\; \&_{a \in \Sigma} \Big(a \in \lim X(x) \to \bigvee_{k=1}^{s} (a = a_{ik})\Big)\Big).$$

Теорема (анализа). По всякому источнику S можно построить формулу ЛОП $A(x,y)$, для которой слово

$$X(x)X(x+1)\ldots X(y-1) \in Beh(S) \longleftrightarrow A(X,x,y).$$

Доказательство. Источнику без пустых ребер $B = (X,Q_0,D,F)$, где $F \subseteq Q$, поставим в соответствие формулу ЛОП

$$A(X,x,z) \text{ есть } (\exists Q)\Big(\Big(\bigvee_{q \in Q} Q(x) = q \;\&\; (\forall y)_x^z \Big(\bigvee_{(q,a,q') \in D} (Q(y) = q \;\&\; X(y) = a \;\&\; Q(y+1) = q')\Big) \to \Big(\bigvee_{q \in F} Q(z) = q\Big)\Big)\Big).$$

Тогда слово $X(x)X(x+1)\ldots X(z-1) \in Beh(S) \longleftrightarrow A(x,z)$.

Заметим, что формула A фиктивна вне $[x,z)$ в том смысле, что значения предиката X вне $[x,z)$ не влияют на характер истинности формулы A. Если хотим иметь в источнике и пустые ребра, то алфавит Σ следует расширить пустым символом.

30.2.1. Макроисточники и ЛОП

Теорема (анализа). По всякому макроисточнику MS можно построить формулу ЛОП $A(X,x)$, для которой сверхслово $X(x)X(x+1)\ldots \in Beh(MS) \longleftrightarrow A(X,x)$. Значение предиката X вне $[0,x)$ не существенно.

Доказательство. Макроисточнику $MS = (X,Q,Q_0,D,F')$, где $F' \in P(Q)$, сопоставим формулу

$$A(X,y) \text{ есть } (\exists Q)\Big(\Big(\bigvee_{q \in Q_0} Q(y) = q\Big) \;\&\; (\forall x)_{\geqslant y} \Big(\bigvee_{(q,a,q') \in D} (Q(x) = q \;\&\; X(x) = a \;\&\; Q(x+1) = q')\Big) \;\&\; \Big(\bigvee_{F \in F'} \lim Q(x) = F\Big)\Big).$$

Сверхслово $X(0)X(1)\ldots X(y)X(y+1) \in Beh(MS) \longleftrightarrow A(X,x)$.

Формула $A(X,x)$ фиктивна вне $[0,y)$.

30.2.2. *Регулярные языки и ЛОП*

Теорема (анализа). Для всякого регулярного языка R можно построить формулу ЛОП $A(X,x,y)$, для которой слово

$$X(x)X(x+1)\ldots X(y-1) \in R \leftrightarrow A(X,x,y).$$

Значение предиката X вне $[x,y)$ не существенно.

Доказательство. Индукция по глубине k построения R.

Базис. $k = 1$. R есть a, где $a \in \Sigma$. Тогда $A(X,x,y)$ есть $X(x) = a \;\&\; y = x+1$.

Предположение индукции. Допустим, что для всякого регулярного языка с глубиной построения меньше k требуемая формула ЛОП может быть построена.

Шаг индукции. Покажем, что для всякого регулярного языка с глубиной построения k требуемая формула может быть построена. Пусть регулярный язык R имеет глубину построения k. Возможны следующие случаи.

1. $R = R_1 * R_2$, где $* \in \{\vee, \cdot\}$. Так как глубина построения регулярных языков R_1 и R_2 меньше k, то по предположению индукции существуют формулы ЛОП $A_1(X,x,y)$ и $A_2(X,x,y)$, для которых слово

$$X(x)X(x+1)\ldots X(y-1) \in R_i \leftrightarrow A_i(X,x,y),\ i = 1,2.$$

Тогда формула A, равная $A_1 \vee A_2$, соответствует языку $R_1 \vee R_2$, а формула $A(X,x,y)$, равная $(\exists t)_x^y\ (A_1(X,x,t) \;\&\; A_2(X,t,y))$, соответствует языку $R_1 \cdot R_2$.

2. $R = R_1^*$. Так как R_1 имеет глубину построения меньше k, то по предположению индукции для R_1 существует формула ЛОП $A_1(X,x,y)$, для которой $X(x)X(x+1)\ldots X(y-1) \in R_1 \leftrightarrow A_1(X,x,y)$. Пусть формула ЛОП $B(X,x,y)$ есть

$$Y(x) = a \;\&\; Y(y) = a \;\&\; (\forall t)(x < t < y \to Y(t) \neq a)).$$

Положим формулу $A(X,x,y)$ равной

$$x < y \;\&\; (\exists Y)\ (Y(x) = a \;\&\; Y(y) = a \;\&$$

$$(\forall u)(\forall v)(x \leqslant u < v \leqslant y \;\&\; B(Y,u,v) \to A_1(X,u,v))).$$

Тогда слово $X(x)X(x+1)\ldots X(x,y-1) \in R \leftrightarrow A(X,x,y)$. Значения предиката X вне $[x,y)$ не существенно.

Замечание. Сверхслово Y имеет вхождение буквы a на тех позициях отрезка $[x,y)$, которые разбивают слово $X(x)X(x+1)\ldots X(y-1)$ на куски, принадлежащие итерируемому языку.

30.2.3. *Общерегулярные языки и ЛОП*

Теорема (анализа). Для всякого общерегулярного языка Ω можно построить формулу ЛОП $A(X,x)$, для которой сверхслово

$$X(x)X(x+1)\ \ldots \in \Omega \leftrightarrow A(X,x).$$

Значения предиката X вне $[0,x)$ не существенны.

Доказательство. Индукция по глубине r построения Ω.

Базис. $k = 1$. Пусть R есть регулярный язык и $\Omega \subseteq R^\infty$ есть общерегулярный язык. Для R существует формула ЛОП $A_1(X,x,y)$, для которой слово $X(x)X(x+1)\ldots X(y-1) \in R \leftrightarrow A_1(X,x,y)$.

Пусть формула $A(X,x)$ есть

$$(\exists Y)\ (Y(x) = a \;\&\; (\exists^\infty y)(Y(y) = a) \;\&$$

$$(\forall u)(\forall v)(x \leqslant u < v \;\&\; B(Y,u,v) \to A_1(X,u,v))),$$

где B есть построенная в предыдущей теореме формула. Тогда сверхслово $X(x)X(x+1)\ldots \in R^\infty \leftrightarrow A(X,x)$.

Значения предиката X вне $[0,x)$ не существенны.

Предположение индукции. Пусть для всякого общерегулярного сверхъязыка с глубиной построения меньше k требуемая формула существует.

Шаг индукции. Покажем, что для всякого общерегулярного сверхъязыка с глубиной построения k требуемая формула может быть построена. Пусть общерегулярный сверхъязык Ω имеет глубину построения k. Тогда $\Omega = R \cdot \Omega_1$ для некоторого регулярного языка R и общерегулярного сверхъязыка Ω_1. Для R существует формула ЛОП $C(X,x,y)$, для которой слово

$$X(x)X(x+1)\ldots X(y-1) \in R \leftrightarrow C(X,x,y).$$

Так как глубина построения Ω_1 меньше k, то по предположению индукции для Ω_1 существует формула $D(X,x)$, для которой сверхслово

$$X(x)X(x+1)\ldots \in \Omega_1 \leftrightarrow D(X,x).$$

Пусть формула $A(X,x)$ есть

$$(\exists y)(x < y \;\&\; C(X,x,y) \;\&\; D(X,y)).$$

Тогда сверхслово $X(x)X(x+1)\ldots \in \Omega \leftrightarrow A(X,x)$. Значения предиката X вне $[0,x)$ не существенны.

30.3. Специальная префиксная форма

Формула A в ЛОП имеет специальную префиксную форму, если A находится в префиксной форме и при этом в кванторной приставке все предикатные кванторы предшествуют всем предметным

кванторам. Покажем, что всякая формула ЛОП эквивалентна некоторой формуле ЛОП в специальной префиксной форме. Пусть A есть формула ЛОП. В формуле A устраним все вхождения нуля, именно, если формула A имеет вид $A(0)$, то заменим ее на формулу $(\exists x)(x = 0 \ \& \ A(x))$, где $x = 0$ есть $(\forall y)(x \neq y + 1)$. В результате получим формулу A_1, не имеющую нулей и эквивалентную формуле A.

В формуле A_1 устраним итерации штриха, заменяя вхождения атомарной формулы $X(x'\ldots')$, где штрих встречается k раз, на формулу

$$(\exists y_1)(\exists y_2)\ldots(\exists y_k)\ (y_1 = x + 1 \ \& \ y_2 = y_1 + 1 \ \& \ \ldots \ \& \ y_k = y_{k-1} + 1 \ \& \ X(y_k)).$$

В результате получим формулу A_2, эквивалентную исходной формуле, при этом каждая предметная переменная в A_2 имеет не более одного штриха, а предметные переменные, являющиеся аргументами свободных предикатных переменных формулы A вовсе не имеют штрихов.

Приведем формулу A_2 к префиксной форме. В результате получим формулу A_3 в префиксной форме, эквивалентную исходной формуле A.

Справедливы следующие эквивалентности:

$$(\exists x)(\forall X)\ C(X,x) \longleftrightarrow \tag{30.2}$$

$$(\exists Y)(\forall X)(\forall x)\ ((\exists y)\ Y^a(y) \ \& \ (Y^a(x) \to C(X,x)));$$

$$(\forall x)(\exists X)\ C(X,x) \longleftrightarrow \tag{30.3}$$

$$(\forall Y)(\exists X)(\exists x)\ ((\forall y)\ Y^a(y) \to Y^a(x) \ \& \ C(X,x))); \ a \in \Sigma.$$

Докажем первую из них. Пусть в формуле (30.2) истинна правая часть. Тогда имеем следующий ряд эквивалентных формул:

$$(\exists Y)(\forall X)(\forall x)\ ((\exists y)\ Y_0^a(y) \ \& \ (Y_0^a(x) \to C(X,x)));$$

$$(\forall X)(\forall x)\ ((\exists y)\ Y_0^a(y) \ \& \ (Y_0^a(x) \to C(X,x)));$$

$$(\forall x)\ (Y_0^a(y_0) \ \& \ (Y_0^a(x) \to C(X,x)));$$

$$Y_0^a(y_0) \ \& \ (Y_0^a(y_0) \to C(X,y_0));$$

$$Y_0^a(y_0); \quad Y_0^a(y_0) \to C(X,y_0);$$

$$C(X,y_0);$$

$$(\forall X)\ C(X,y_0);$$

$$(\exists x)(\forall X)\ C(X,x).$$

Пусть теперь в (30.2) истинна левая часть $(\exists x)(\forall X)C(X,x)$. Тогда $(\forall X)C(X,y_0)$ истинно при некотором y_0. Зададим предикат $Y_0(y)$ так, чтобы $Y_0^a(y_0) \ \& \ (\forall x)(x \neq y_0 \to Y_0(x) \neq a)$.

Следующие далее формулы последовательно эквивалентны друг другу (во всех этих формулах $a \in \Sigma$):

$$Y_0^a(y_0); \quad (\forall X)C(X,y_0);$$

$$Y_0^a(y_0); \quad Y^a(y_0) \to (\forall X)C(X,y_0);$$

$$Y_0^a(y_0); \quad (\forall x)(Y_0^a(x\ ) \to (\forall X)C(X,x);$$

$$(\exists y)\ Y_0^a(y); \quad (\forall x)(\forall X)(Y_0^a(x) \to C(X,x));$$

$$(\exists y)\ Y_0^a(y) \ \& \ (\forall X)(\forall x)(Y_0^a(x) \to C(X,x));$$

$$(\forall X)(\forall x)\ ((\exists y)\ Y_0^a(y) \ \& \ (Y_0^a(x) \to C(X,x)));$$

$$(\exists Y)(\forall X)(\forall x)\ \ ((\exists y)\ Y^a(y) \ \& \ (Y^a(x) \to C(X,x))).$$

Справедливость соотношения (30.2) установлена. Справедливость формулы (30.3) показывается аналогично. Левая и правая часть в (30.3) двойственны соответствующим частям в (30.2).

Используем эквивалентности (30.2) и (30.3) для разделения предикатных и предметных кванторов в формуле A_3. В результате получим формулу A_4 в специальной префиксной форме (говорят еще в специальной предваренной форме), эквивалентную исходной формуле A.

30.4. Синтез автомата по формуле ЛОП

Пусть $A(X_1,\ldots,X_n)$ есть формула ЛОП; $X_1,\ldots,X_n$ есть полный список ее свободных предикатных переменных; свободных предметных переменных формула A не имеет. Покажем, что по формуле A можно построить макроавтомат, поведение которого совпадает с множеством истинности формулы A.

Так как для всякой формулы ЛОП существует ей эквивалентная формула в специальной префиксной форме, то можно сразу считать, что формула A находится в специальной префиксной форме, при этом согласно предыдущему пункту свободные предикатные переменные $X_1,\ldots,X_n$ формулы A имеют предметные переменные без штрихов. Формула A имеет следующий вид:

$$(Q_1Y_1)\ldots(Q_rY_r)\ (Q_1'x_1)\ \ldots\ (Q_k'x_k)\ B,$$

где B есть бескванторная формула ЛОП. Формулу B представим в виде СДНФ; далее согласно формуле (19.1) заменим отрицания

вида $X(t) \neq a$ на эквивалентную формулу $\bigvee_{b \in \Sigma - \{a\}} X(t) = b$, не содержащую отрицаний; результат приведем к СДНФ. Тогда получим формулу A_1, эквивалентную исходной формуле и имеющую вид

$$(Q_1 Y_1)\ldots(Q_r y_r)\ (Q'_1 x_1)\ldots(Q'_k x_k)\ \Big(\bigvee_i$$
$$A_{i11}(x_1)\ldots A_{i1r}(x_1)B_{i11}(x'_1)\ldots B_{i1r}(x'_1)\ C_{i11}(x_1)\ldots C_{i1n}(x_1)$$
$$\ldots$$
$$A_{ik1}(x_k)\ldots A_{ikr}(x_k)B_{ik1}(x'_k)\ldots B_{ikr}(x'_k)C_{ik1}(x_k)\ldots C_{ikn}(x_k)),$$

где каждые $A_{ijl}(x_l)$, $B_{ijl}(x'_l)$, $C_{ijl}(x'_l)$ есть некоторые равенства

$$Y_j(x_l) = a_{ijl}, \quad Y_j(x'_l) = b_{ijl},$$
$$X_j(x_l) = c_{ijl}; \quad a_{ijl}, b_{ijl}, c_{ijl} \in \Sigma.$$

Так как в бескванторную часть формулы A_1 входят только одноместные предикатные переменные, то можно использовать процедуру Бемана продвижения предметных кванторов в глубину бескванторной СДНФ в формуле A_1. Продвигаем сначала квантор $(Q'_k x_k)$. Если Q'_k есть квантор $\exists$, то в A_1 продвигаем квантор $(\exists x_k)$ сначала по дизъюнкции, а потом по конъюнкции. В результате получим формулу

$$(Q_1 Y_1)\ldots(Q_r Y_r)\ (Q'_k x_1)\ldots(Q'_{k-1} x_{k-1})\ \Big(\bigvee_i$$
$$A_{i_{11}}(x_1)\ldots A_{i1r}(x_1)\ B_{i11}(x'_1)\ldots B_{i1r}(x'_1)\ C_{i11}(x_1)\ldots C_{i1n}(x_1)$$
$$\ldots$$
$$A_{i,k-1,1}(x_{k-1})\ \ldots\ B_{i,k-1,1}(x'_{k-1})\ \ldots\ C_{i,k-1,n}(x_{k-1})$$
$$(\exists x_k)(A_{ik1}(x_k)\ldots A_{ikr}(x_k)B_{ik1}(x'_k)\ldots B_{ikr}(x'_k)\ldots C_{ikn}(x_k)).$$

$$E_{ik} = (\exists x)\ D_{ik}(x)$$

Формула E_{ik} выступает в дальнейших преобразованиях единым неизменным блоком. Квантор $(\exists x)$ в формулу A_1 продвинут.

Если Q'_k есть $\forall$, то справедлива следующая последовательность эквивалентных формул:

$$(\forall x_k)(\text{СДНФ}) \longleftrightarrow \neg\neg(\forall x_k)(\text{СДНФ}) \longleftrightarrow \neg(\exists x_k)\neg(\text{СДНФ}) \longleftrightarrow \neg(\exists x_k)(\text{СДНФ1}).$$

Далее квантор $(\exists x_k)$ продвигаем в глубину СДНФ1, как это делалось выше, а внешнее отрицание распределяется над получившейся в результате продвижения квантора существования ДНФ. В обоих случаях продвижения квантора в глубину формулы

A_1 и последующего превращения бескванторной части в СДНФ получим формулу A_2 вида

$$(Q_1 Y_1)\ldots(Q_r Y_r)\ (Q'_1 x_1)\ldots(Q'_{k-1} x_{k-1})\ \Big(\bigvee_i$$
$$A_{i11}(x_1)\ldots A_{i1r}(x_i)B_{i11}(x'_1)\ldots B_{i1r}(x'_i)C_{i11}(x_1)\ldots C_{i1r}(x_1)E'_{ikl}$$
$$\ldots$$
$$A_{i,k-1,1}(x_{k-1})\ldots B_{i,k-1,1}(x'_{k-1})\ldots C_{i,k-1,n}(x_{k-1})E'_{i,k,k-1},$$

где каждое E'_{ijl} есть E'_{ik}, взятое с отрицанием или без него.

Последовательно продвигая в глубину формулы A_2 все другие предметные кванторы, получим формулу A_3 вида

$$(Q_1 Y_1)\ldots(Q_r Y_r)\ \Big(\bigvee_i\ E'_{i1}\ E'_{i2}\ \ldots\ E'_{ik}\Big),$$

где каждое E'_{ij} есть E_{ij} с отрицанием или без него.

Пусть одна из E_{ij} есть формула E, равная

$$(\exists x)\ (Y_1(x) = a_1\ \&\ Y_1(x+1) = b_1\ \&$$
$$Y_2(x) = a_2\ \&\ Y_2(x+1) = b_2\ \&$$
$$\ldots$$
$$Y_r(x) = a_r\ \&\ Y_r(x+1) = b_r\ \&$$
$$X_1(x) = c_1\ \&$$
$$X_2(x) = c_2\ \&$$
$$\ldots$$
$$X_n(x) = c_n).$$

Формула E утверждает, что набор предикатов (сверхслов)

$$\begin{array}{l} Y_1(0)\ Y_1(1)\ \ldots\ Y_1(x)\ Y_1(x+1)\ \ldots \\ Y_2(0)\ Y_2(1)\ \ldots\ Y_2(x)\ Y_2(x+1)\ \ldots \\ \ldots \\ Y_r(0)\ Y_r(1)\ \ldots\ Y_r(x)\ Y_r(x+1)\ \ldots \\ X_1(0)\ X_1(1)\ \ldots\ X_1(x)\ X_1(x+1)\ \ldots \\ \ldots \\ X_n(0)\ X_n(1)\ \ldots\ X_n(x)\ X_n(x+1)\ \ldots \end{array} \qquad (30.4)$$

принадлежит множеству истинности формулы E тогда и только тогда, когда два соседних разряда x и $x+1$, высекают в сверхсловах (30.4) двухбуквенное "толстое" слово

$$(\hat{a},\hat{b}) = (a_1\ldots a_r c_1\ldots c_n)^T(b_1\ldots b_r d_1\ldots d_n)^T$$

в алфавите Σ^{r+n} (знак "T" означает транспонирование). Пусть H есть множество пар $(\hat{a},\hat{b})$ всех таких двубуквенных слов в

алфавите Σ^{r+n}. Пусть $H_1 = \{\hat{a}: \exists\hat{b}\ (\hat{a},\hat{b}) \in H\}$; $H_2 = \{\hat{b}: \exists\hat{a}\ (\hat{a},\hat{b}) \in H\}$, т.е. H_1 и H_2 есть проекции множества H на первую и вторую ось соответственно.

Построим макроавтомат $MA = (\Sigma^{r+n},Q,q_0,T,F)$, поведение которого совпадает с множеством истинности формулы E. В качестве состояний макроавтомата MA возьмем множество

$$Q = \{q_{\hat{a},\hat{b}} : (\hat{a},\hat{b}) \in H\} \cup \{q_{\hat{a}} : \hat{a} \in H_1\} \cup \{q_0\}.$$

Функция переходов $T: Q\times\Sigma^{r+n} \to Q$ строится следующим образом:

$$T(q_0,a) = \begin{cases} q_0, & \text{если } a \notin H_1, \\ q_{\hat{a}}, & \text{если } a \in H_1; \end{cases}$$

$$T(q_{\hat{a}},\hat{b}) = \begin{cases} q_{\hat{a},\hat{b}}, & \text{если } \hat{b} \in H_2, \\ q_{\hat{b}}, & \text{если } \hat{b} \notin H_2 \ \&\ \hat{b} \in H_1, \\ q_0, & \text{если } \hat{b} \notin H_2 \ \&\ \hat{b} \notin H_1; \end{cases}$$

$$T(q_{\hat{a},\hat{b}},c) = q_{\hat{a},\hat{b}}, \quad \forall c \in \Sigma^{r+n}.$$

Множество выделенных состояний $F = \{\{q_{\hat{a},\hat{b}}\}: (\hat{a},\hat{b}) \in H\}$.

Макроавтомат MA построен; для него $\hat{E} = Beh(MA)$.

Аналогично строим макроавтоматы для остальных формул E_{ij} из формулы A_3. Далее действуем по построению формулы A_3. Используя теорему о замкнутости класса сверхъязыков, определимых макроавтоматами, относительно булевых операций и проекции, строим макроавтомат, поведение которого совпадает с множеством истинности формулы A_3, а потому и ей эквивалентной исходной формулы A.

Требуемый макроавтомат построен.

Теорема. Монадическая логика натуральных чисел алгоритмически разрешима, т.е. по всякой замкнутой формуле ЛОП можно установить истинна она или ложна.

Доказательство. Пусть A есть замкнутая формула ЛОП. Приведем A к специальной префиксной форме; она имеет вид (QX)B(X), где формула $B(X)$ имеет единственную свободную переменную X. По формуле $B(X)$ построим макроавтомат MB, поведение которого совпадает с множеством истинности формулы B. Если квантор (QX) есть $(\exists X)$, то вопрос об истинности или ложности формулы $(\exists X)B(X)$ эквивалентен вопросу о непустоте или пустоте сверхъязыка, представимого макроавтоматом MB. Проблема пустоты для макроавтомата алгоритмически разрешима. Следовательно, распознаваема и проблема истинности экзистенциальной формулы $(\exists X)B(X)$. Вопрос об истинности формулы $(\forall X)B(X)$ с квантором общности сводится к аналогичному вопросу для квантора существования.

Замечание. Проблему униформизации конечно автоматного языка в алфавите $X\times Y$ можно решить с помощью указанного алгоритма. Именно, по макроавтомату A, представляющему сверхъязык L, строим формулу ЛОП $A(X,Y)$, множество истинности которой совпадает с множеством L. Если истинна формула $(\forall X)(\exists Y)A(X,Y)$, то существует конечно автоматный оператор $Y = \mathbb{O}(X)$, униформизующий сверхъязык L. Если истинно отрицание исходной формулы, то существует конечно автоматный оператор $X = \mathbb{O}(Y)$, униформизующий дополнение сверхъязыка L. Конечный автомат, реализующий оператор $\mathbb{O}$, строится согласно алгоритму, приведенному в главе 29.

СПИСОК СОКРАЩЕНИЙ И ЗНАКОВ

ДНФ дизъюнктивная нормальная форма
И, 1 истина
КНФ конъюнктивная нормальная форма
Л, 0 ложь
ЛП логика предикатов
МДНФ минимальная дизъюнктивная нормальная форма
нод наибольший общий делитель
нок наименьшее общее кратное
од общий делитель
ок общее кратное
ПИ простой импликант
СДНФ совершенная дизъюнктивная нормальная форма
СКНФ совершенная конъюнктивная нормальная форма
ТДНФ тупиковая дизъюнктивная нормальная форма
$\forall$ квантор общности
$\exists$ квантор существования
$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ множества натуральных, целых, рациональных, вещественных, комплексных чисел соответственно
$=$, $\neq$ равенство и неравенство
$\leqslant$, $<$ неравенство нестрогое и строгое
$b \mid a$, b делит a (без остатка)
$b \nmid a$, b не делит a (без остатка)
$a \vdots b$, a делится на b (без остатка)
(a,b) наибольший общий делитель для a и b
$[a,b]$ наименьшее общее кратное для a и b
$\mathrm{mod}(a,b)$, $\mathrm{rest}(a,b)$ остаток от деления a на b
$\lfloor a/b \rfloor$ частное от деления a на b
$\rightleftharpoons$ обозначение, "есть обозначение для формулы"
f^n n-местная функция
C_n^k, $\binom{n}{k}$ число сочетаний из n по k
$\log_b a$ логарифм a по основанию b
&, $\vee$ конъюнкция, дизъюнкция
$-$, $\neg$ отрицание
$\rightarrow$, $\rightarrow$ импликация
$\equiv$ эквивалентность логическая
$\mid$ штрих Шеффера
$\uparrow$ стрелка Пирса
$\longleftrightarrow$ тогда и только тогда когда, если и только если

$\Leftrightarrow$ равносильность
$\models$ выполнимость
$+, -, \cdot$ сложение, вычитание, умножение
$/$, $:$ деление
$\in$, $\notin$ принадлежит, не принадлежит
$\cup$, $\cap$, $-$ объединение, пересечение, вычитание множеств
$\subseteq$, $\subset$ включение множеств, нестрогое и строгое
$\times$ декартово произведение множеств
$\varnothing$ пустое множество
$\sim$ эквивалентность (бинарное отношение)
$\|$ знак конкатенации слов: ab45tu$\|$hj9cv = ab45tu hj9cv
$\sum_{i=1}^{n}$, $\sum_{i=1}^{n}$ сумма по i от 1 до n
$\prod_{i=1}^{n}$, $\prod_{i=1}^{n}$ произведение по i от 1 до n
$\sum_{d \mid n}$ сумма по всем делителям d числа n
$\prod_{d \mid n}$ произведение по всем делителям d числа n

ЛИТЕРАТУРА

Алексеев В.Б. Введение в теорию сложности алгоритмов. М.: Макс ПРЕСС, 2002. 82 с.

Алексеев В.Б. Лекции по дискретной математике. М.: Изд. отд. Фак. ВМиК МГУ им. М.В.Ломоносова, 2004. 74 с.

Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии: Учебное пособие, 2-е изд., испр. и доп. М.: "Гелиос АРВ", 2002. 480 с.

Ахо А, Хопкрофт Дж, Ульман Дж. Построение и анализ вычислительных алгоритмов. М.: Мир, 1989. 536 с.

Бабаш А.В., Шанкин Г П Криптография. М.: СОЛОН-Р, 2002. 512 с.

Басакер Р., Саати Т. Конечные графы и сети. М.: Наука. 1974.

Берж К. Теория графов и ее применения. М.: ИЛ,1962. 319с.

Белоусов А.И., Ткачев С.Б. Дискретная математика: Учебник для вузов /Под ред. В.С.Зарубина и А.П.Крищенко. – 4-е изд. М.: Изд-во МГТУ им.Н.Э.Баумана, 2006. 743 с.

Биркгоф Г., Барти Т. Современная прикладная алгебра. М.: Мир, 1976. 400 с. .

Биркгоф Г. Теория решеток. М.: Наука, 1984. 568 с.

Боднарчук В.Г., Калужнин Л.А., Котов В.Н., Ромов Б.А. Теория Галуа для алгебр Поста. Кибернетика, 1969, N 3, стр. 1-10 (Киев); Кибернетика, 1969, N 5, стр.1-9.

Болотов А.А., Мещанинов Д.Г., Фролов А.Б. Алгебраические структуры. М.: Изд-во МЭИ, 2005. 80 с.

Болотов А.А., Гашков С.Б., Фролов А.Б. Часовских А.А. Элементарное введение в эллиптическую криптографию. Алгебраические и алгоритмические основы. М.: КомКнига, 2006. 328 с.

Болотов А.А., Гашков С.Б., Фролов А.Б. Элементарное введение в эллиптическую криптографию. Протоколы криптографии на эллиптических кривых. М.: КомКнига, 2006. 280 с.

Бухштаб А.А. Теория чисел. М.: Просвещение, 1966. 384 р.

Виноградов И.М. Основы теории чисел. М.: Высшая школа, 1965. 172 с.

Гаврилов Г.П., Сапоженко А.А. Сборник задач по дискретной математике. М.: Наука, 1977. 368 с.

Гельфонд А.О. Исчисление конечных разностей. М.: Наука, 1967. 375 с.

Горбатов В.А., Горбатов А.В., Горбатова М.В. Дискретная

математика: Учеб. для студентов втузов. – Издательство АСТ>: ООО <Издательство Астрель>, 2003. 447 с.

Диффи У., Хелман М. Защищенность и имитостойкость. Введение в криптографию. //ТИИЭР. 1979. Т.67. N 3. С.71-109.)

Дискретная математика и математические вопросы кибернетики. Т.1. / Под ред. С.В.Яблонского и О.Б.Лупанова. М.: Наука, 1974. 312 с.

Емеличев В.А., Мельников О.И., Сарванов В.И., Тышкевич Р.И. Лекции по теории графов. М.: Наука, 1990. 384 с.

Зыков А.А. Основы теории графов. М.: Наука, 1987. 384 с.

Калужнин Л.А., Сущанский В.И. Преобразования и перестановки. М.: Наука, 1985. 160 с.

Клини С.К. Введение в метаматематику. М.: МИР, 1957.526с.

Клини С.К. Математическая логика. М.: МИР, 1973. 480 с.

Кон П. Универсальная алгебра. М.: Мир, 1968. 352 с.

Кофман А. Введение в прикладную комбинаторику. М.: Наука, 1975. 480 с.

Кудрявцев В.Б. Функциональные системы. М.: Изд-во МГУ, 1982. 157 с.

Кудрявцев В.Б., Алешин С.В., Подколзин А.С. Введение в теорию автоматов. М.: Наука, 1985. 319 с.

Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженера. М.: Энергоатомиздат, 1988. 480 с.

Кук Д., Бейз Г. Компьютерная математика. М.: Наука, 1990. 384 с.

Лидл Р., Нидеррайтер Г. Конечные поля. М.: Мир, 1988. 820 с.

Липский В. Комбинаторика для программистов. М.: Мир, 1988. 213 с.

Лупанов О.Б. Асимптотические оценки сложности управляющих систем. М.: Изд-во МГУ, 1984. 137 с.

Мальцев А.И. Алгоритмы и рекурсивные функции. М.: Наука, 1965. 392 с.

Мальцев А.И. Итеративные алгебры и многообразия Поста. Алгебра и логика (Новосибирск), 1966, N 5, стр.5-24.

Мальцев А.И. Алгебраические системы. М.: Наука, 1970.392с

Мальцев А.И. Основы линейной алгебры. М.: Наука, 1975, 400 с.

Марченков С.С. Замкнутые классы булевых функций. М.: Физматгиз, 2000. 128 с.

Матросов В.Л. Стеценко В.А. Лекции по дискретной матема-

тике. М.: Изд-во Моск. педагог. гос. ун-та, 1997. 220 с.

Меньшиков М.В., Ревякин А.М., Копылова А.Н., Макаров Ю.Н., Стечкин Б.С. Комбинаторный анализ. Задачи и упражнения: Учебное пособие /Под ред. К.А.Рыбникова. М.: Наука, 1982. 368с.

Набебин А.А. Логика и Пролог в дискретной математике. М.: Изд-во Моск.энерг.института, 1996. 452 с.

Набебин А.А. Модулярная арифметика в криптографии. М.: МЭИ, 2007. 201 с.

Набебин А.А., Кораблин Ю.П. Математическая логика и теория алгоритмов. М.: Научный мир, 2008. 343 с.

Набебин А.А. Сборник заданий по дискретной математике. М.: Научный мир, 2009. 280 с.

Нечаев В.И. Элементы криптографии (Основы защиты информации). М.: Высшая школа, 1999. 109 с.

Новиков П.С. Элементы математической логики. М.: Физматгиз, 1959. 400 с.

Оре О. Теория графов. М.: Наука, 1968. 352 с.

Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженера. М.: Энергоатомиздат, 1988. 480 с.

Питерсон У. Коды, исправляющие ошибки. М.:Мир, 1964,339с.

Райзер Г.Дж. Комбинаторная математика. М.: Мир,1966.154с.

Риордан Дж. Комбинаторные тождества. М.: Наука,1982.256с.

Риордан Дж. Введение в комбинаторный анализ. М.: ИЛ, 1963, 288 с.

Рыбников К.А. Введение в комбинаторный анализ. М.: Изд-во Моск. ун-та. 1985. 308 с.

Саломаа А. Криптография с открытым ключем. М.: Мир,1986. 318 с.

Сикорский Р. Булевы алгебры. М.: МИР. 1969. 376 с.

Смарт Н. Криптография. М.: Техносфера, 2006. 525 с.

Трахтенброт Б.А., Барздынь Я.М. Конечные автоматы. М.: Наука, 1970. 400 с.

Уилсон Р. Введение в теорию графов. М.: Мир, 1977. 208 с.

Успенский В.А., Семенов А.Л. Теория алгоритмов: основные открытия и приложения. М.: Наука, 1987. 288 с.

Фаддеев Д.К. Лекции по алгебре. М.: Наука, 1984. 416 с.

Фролов А.Б., Андреев А.Е., Болотов А.А., Коляда К.В. Прикладные задачи дискретной математики и сложность алгоритмов. М.: Издат-во МЭИ, 1997. 312 с.

Харари Ф. Теория графов. М.: Мир, 1873. 300 с.

Чмора А.Л. Современная прикладная криптография. М.: "Гелиос АФВ", 2001. 480 с.

Шнайер Б. Практическая криптография. М.: Триумф, 2002. 815 с.

Шнайер Б., Фергюсон Н. Практическая криптография. М.: Вильямс, 2005. 424 с.

Яблонский С.В. Введение в дискретную математику. М.: Наука, 1986. 384 с.

Яблонский С.В., Гаврилов Г.П., Набебин А.А. Предполные классы в многозначных логиках. М.: Изд-во МЭИ, 1997. 144 с.

Ященко В.В. Введение в криптографию. М.: МЦНМО, 1998.

Albert A.A. Fundamental concept of higher algebra. 1956.

Appel K., Haken W. Every planar map is colorable. Illinois journal of mathematics: 1976. V.20. No 2. P.218–297; 1977. V.21. No 3. P.429–490; 1977. V.21. No 3. P.491–567.

Berlencamp E.R. Algebraic coding theory. California,1984.

Fillmore J.P., Marks M.L. Linear recursive sequences. SIAM Rev., 10, 1968. P.342–353.

Garey M.R., Johnson D.S. Computers and intractability. A guide to the theory of NP–completeness. W.H.Freeman, San Francisko, 1979.

Kahn D. The codebreakers. N.Y., 1967.

Koblitz N. A course in number theory and cryptography. Springer–Verlag, 1994. 238 p.

Koblitz N. Algebraic aspects of cryptography. Springer–Verlag, 1997. 206 p.

McEliece R.J. Finite fields for computer scientists and engineers. 1987. 207 p.

Menezes A., van Oorshot P., Vanstone S. Handbook of applied cryptography. CRC Press. 1996. 780 p.

Internet address: www.cacr.math.uwaterloo.ca/hac

Renji T. On finite automaton one–key cryptosystem. / Fast software encryption. Cambridge security workshop proceedings, Springer Verlag, 1994. P.135–148.

ОГЛАВЛЕНИЕ

Лучшие книги по всем областям знаний – в издательстве «Научный мир»

Набебин А.А., Кораблин Ю.П. **Математическая логика и теория алгоритмов.** Учебное пособие. – М.: Научный мир, 2008. – 343 с.

Набебин А.А. **Сборник заданий по дискретной математике.** – М.: Научный мир, 2009. – 280 с.

Шиханович Ю.А. **Начальные главы математического анализа в полуформальном изложении.** Учебное пособие. – М.: Научный мир, 2010. – 288 с.

Шиханович Ю.А. **Введение в математику.** Учебное пособие. – М.: Научный мир, 2005. – 383 с.

Шиханович Ю.А. **Минимум для теории алгоритмов для нематематиков.** – М.: Научный мир, 2009. – 160 с.

Хайрер Э., Ваннер Г. **Математический анализ в свете его истории.** – М.: Научный мир, 2008. – 396 с.

Свешников А.Г., Альшин А.Б., Корпусов М.О. **Нелинейный функциональный анализ и его приложения к уравнениям в частных производных.** – М.: Научный мир, 2008. – 400 с.

Рублев В.С. **Основы теории алгоритмов.** Учебное пособие. – М.: Научный мир, 2008. – 136 с. +илл.

Казарян В.П., Лолаев Т.П. **Математика и культура.** – М.: Научный мир, 2004. – 288 с.

Дмитриев В.И. **Математика (для абитуриентов) /УДК 373.167.1.** – М.: Научный мир, 2005. – 336 с., 141 илл.

Бениаминов Е.М., Ефимова Е.А. **Элементы универсальной алгебры и ее приложений в информатике.** – М.: Научный мир, 2004. – 168 с.

Бениаминов Е.М. **Алгебраические методы в теории без данных и представлений знаний.** – М.: Научный мир, 2003. – 184 с.

Елизарова Т.Г. **Квазигазодинамические уравнения и методы расчета вязких течений. Лекции по математическим моделям и численным методам в газовой динамике.** – М.: Научный мир, 2007. – 352 с.

Гросс Д.Х.Э. **Микроканоническая термодинамика. Пер. с англ. (серия: Фундаментальные основы нанотехнологий: лучшие зарубежные учебники).** – М.: Научный мир, 2010. – 304 с.

Мансури Г.А. **Принципы нанотехнологии. Исследование конденсированных веществ малых систем на молекулярном уровне.** – М.: Научный мир, 2008. – 320 с.

Книги издательства «Научный мир»
можно приобрести по издательским ценам
в отделе реализации издательства по адресу:
119992, Москва, ул. Знаменка, д. 11/11
Проезд до станции метро «Боровицкая»

Заказы следует направлять
по факсу: (495) 691-28-47 или E-mail: naumir@benran.ru
E-mail: naumir@naumir.ru

Адрес издательства «Научный мир»:
127055, Москва, Тихвинский переулок, д.10/12, корп. 4
Тел. (499) 973-26-70; (499) 973-25-13
E-mail: naumir@naumir.ru
Internet: http://www.naumir.ru

Учебное издание

Набебин Алексей Александрович

ДИСКРЕТНАЯ МАТЕМАТИКА

Корректор И.Г. Ерохина

«Научный мир»
Тел./факс: (499) 973-25-13, (495) 691-28-47.
E-mail: naumir@benran.ru; naumir@naumir.ru

Internet: http://www.naumir.ru

Подписано к печати 12.01.2010. Формат 60x90/16. Гарнитура Таймс.
Печать офсетная. Печ. л. 32. Тираж 1000 экз. Заказ 51.

Издание отпечатано в типографии ООО «Галлея-Принт»
Москва, ул. 5-я Кабельная, 2-б