МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА

Научный журнал

2017 № 38

Зарегистрирован в Федеральной службе по надзору
в сфере связи и массовых коммуникаций

Свидетельство о регистрации ПИ № ФС 77-33762 от 16 октября 2008 г.

Подписной индекс в объединённом каталоге «Пресса России» 38696

УЧРЕДИТЕЛЬ
Томский государственный университет

Адрес редакции и издателя: 634050, г. Томск, пр. Ленина, 36
E-mail: vestnik_pdm@mail.tsu.ru

В журнале публикуются результаты фундаментальных и прикладных научных исследований отечественных и зарубежных ученых, включая студентов и аспирантов, в области дискретной математики и её приложений в криптографии, компьютерной безопасности, кибернетике, информатике, программировании, теории надёжности, интеллектуальных системах.

Периодичность выхода журнала: 4 номера в год.

Редактор *Н. И. Шидловская*
Верстка *И. А. Панкратовой*

Подписано к печати 12.12.2017. Формат $60 \times 84\frac{1}{8}$. Усл. п. л. 15,48. Тираж 300 экз.
Заказ № 2909. Цена свободная. Дата выхода в свет 26.12.2017.

Отпечатано на оборудовании
Издательского Дома Томского государственного университета
634050, г. Томск, пр. Ленина, 36
Тел.: 8(3822)53-15-28, 52-98-49

СОДЕРЖАНИЕ

CONTENTS

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

УДК 519.714.5

ОДИН ПОДХОД К ПОСТРОЕНИЮ ТРАНЗИТИВНОГО МНОЖЕСТВА БЛОЧНЫХ ПРЕОБРАЗОВАНИЙ

И. В. Чередник

Московский технологический университет (МИРЭА), г. Москва, Россия

Пусть Ω — произвольное конечное множество и $\mathcal{Q}(\Omega)$ — семейство всех бинарных квазигрупп, определённых на множестве Ω. Преобразование $\Omega^n \to \Omega^n$, $n \geqslant 2$, реализуемое сетью Σ с одной бинарной операцией F, будем обозначать Σ^F. В терминах строения сети Σ доказан критерий биективности всех преобразований из множества $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ и определено каноническое представление таких сетей. Вводится и разрабатывается аппарат разметки сетей, который позволяет сформулировать и обосновать необходимые и достаточные условия для транзитивности множества преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$. Предложен эффективный способ проверки транзитивности множества преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$. Изложен и обоснован алгоритм построения сетей Σ, для которых множество преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ является транзитивным.

Ключевые слова: *сети, квазигруппы, блочные преобразования, транзитивное множество блочных преобразований.*

DOI 10.17223/20710410/38/1

ONE APPROACH TO CONSTRUCTING A TRANSITIVE CLASS OF BLOCK TRANSFORMATIONS

I. V. Cherednik

Moscow Technological University (MIREA), Moscow, Russia

E-mail: p.n.v.k.s@mail.ru

In the paper, a new type of block transformations is determined and investigated. These transformations can be used to construct hash functions and block ciphers. Let Ω be an arbitrary finite set, and $\mathcal{Q}(\Omega)$ be the collection of all binary quasigroups defined on the set Ω. We consider the mappings $\Sigma^F\colon \Omega^n \to \Omega^n$ that are implemented by a network Σ of a width n with one binary operation $F \in \mathcal{Q}(\Omega)$. The network Σ is called bijective for Ω if the mapping Σ^F is bijective for each $F \in \mathcal{Q}(\Omega)$. We show that the network Σ is bijective for all finite sets iff the network Σ is bijective for some finite set Ω such that $|\Omega| \geqslant 2$. Therefore, we say that the network Σ is bijective if it is bijective for a nontrivial finite set. The networks Σ_1, Σ_2 are called equivalent for Ω if the map Σ_1^F coincides with the map Σ_2^F for each $F \in \mathcal{Q}(\Omega)$. Moreover, we say that the networks Σ_1, Σ_2 are equivalent if the networks Σ_1, Σ_2 are equivalent for all finite sets. It is easy to define the elementary networks by analogy with the elementary matrices. We prove that every bijective network Σ is equivalent to a unique

product of elementary networks. This product is called the canonical representation of Σ and its length is denoted by $\|\Sigma\|$. We prove that bijective networks Σ_1, Σ_2 of equal width n are equivalent iff they are equivalent for some finite set Ω such that $|\Omega| \geqslant \|\Sigma_1\| + \|\Sigma_2\| + n$. A bijective network Σ is called transitive for Ω if the set $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ is transitive. We prove that the bijective network Σ is transitive for all sufficiently large finite sets iff Σ is transitive for some finite set Ω such that $|\Omega| \geqslant \|\Sigma\| + n$. In addition, we propose an effective method for verifying the network transitivity for all sufficiently large finite sets, namely the bijective network Σ is transitive for Ω such that $|\Omega| \geqslant \|\Sigma\| + n$ whenever it is transitive for some two-element subset of Ω. In the final section, we expound an algorithm for constructing transitive networks. For a given bijective network Σ of a width n, the algorithm adds $3n - 3$ elementary networks to the canonical representation of Σ without changing the existing contents. As a result of these modifications, we obtain a bijective network $\widehat{\Sigma}$ that is transitive for every sufficiently large finite set Ω ($|\Omega| \geqslant \|\widehat{\Sigma}\| + n$).

Keywords: *network, quasigroup, block transformation, transitive class of block transformations.*

Введение

Произвольная бинарная операция $F\colon \Omega \times \Omega \to \Omega$ называется квазигруппой на множестве Ω, если уравнения вида

$$F(x, b) = c, \quad F(a, y) = c$$

однозначно разрешимы при любых $a, b, c \in \Omega$ [1]. Множество всех квазигрупп, заданных на множестве Ω, будем обозначать $\mathcal{Q}(\Omega)$.

Пусть $\{x_1, \ldots, x_n\}$ — множество переменных и $*$ — символ бинарной операции. Множество всех формул в алфавите $\{x_1, \ldots, x_n, *\}$ будем обозначать $\mathcal{W}$. При сопоставлении символу «$*$» конкретной бинарной квазигруппы $F \in \mathcal{Q}(\Omega)$ формула $w(x_1, \ldots, x_n)$ реализует отображение $w^F\colon \Omega^n \to \Omega$, а набор формул $(w_1, \ldots, w_m) \in \mathcal{W}^m$ — отображение $(w_1^F, \ldots, w_m^F)\colon \Omega^n \to \Omega^m$.

Определение 1. Пусть $(v_1, \ldots, v_k) \in \mathcal{W}^k$ и в наборе $(w_1, \ldots, w_m) \in \mathcal{W}^m$ каждая из формул w_j, $j \in \{1, \ldots, m\}$, либо имеет вид $v_{i_1} * v_{i_2}$, $i_1 \neq i_2$, $i_1, i_2 \in \{1, \ldots, k\}$, либо является некоторой формулой v_i, $i \in \{1, \ldots, k\}$. Тогда будем говорить, что набор формул $(w_1, \ldots, w_m)$ является *результатом преобразования* набора формул $(v_1, \ldots, v_k)$.

Один из способов построения произвольного набора формул $(w_1, \ldots, w_m)$ заключается в последовательном преобразовании набора переменных $(x_1, \ldots, x_n)$. Для исследования свойств отображений одного класса, соответствующего определённому набору формул, введём дополнительное представление процесса преобразований набора формул, которое отличается большей наглядностью.

Определение 2. Пусть $t, n_0, n_1, \ldots, n_t \in \mathbb{N}$ и

$$X_0 = \{x_1^{(0)}, x_2^{(0)}, \ldots, x_{n_0}^{(0)}\}, X_1 = \{x_1^{(1)}, x_2^{(1)}, \ldots, x_{n_1}^{(1)}\}, \ldots, X_t = \{x_1^{(t)}, x_2^{(t)}, \ldots, x_{n_t}^{(t)}\}$$

— семейство попарно непересекающихся конечных непустых множеств. Тогда *квазигрупповой сетью* (далее — просто *сетью*) *длины* t будем называть простой ориентированный граф Σ с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$, содержащий только рёбра вида $(x_i^{(s-1)}, x_j^{(s)})$, $s \in \{1, \ldots, t\}$, с тем ограничением, что степень захода каждой вершины $x_j^{(s)}$, $s \in \{1, \ldots, t\}$, равна 1 или 2. При этом если степень захода вершины $x_j^{(s)}$

равна 2, то рёбра $(x_{i_1}^{(s-1)}, x_j^{(s)})$ и $(x_{i_2}^{(s-1)}, x_j^{(s)})$ имеют различные метки из множества $\{l, r\}$. Число $\max\{n_0, \ldots, n_t\}$ будем называть *шириной* сети Σ. Множества X_0 и X_t называются множествами начальных и конечных вершин соответственно. Подграф Σ_s сети Σ, основанный на множестве вершин $X_{s-1} \cup X_s$, будем называть *s-м слоем* сети Σ. Сеть Σ называется *однослойной*, если она имеет длину 1.

Определение 3. Пусть Σ и Σ' — сети с множествами вершин $X = X_0 \cup X_1 \cup \ldots \ldots \cup X_s$ и $X' = X'_0 \cup X'_1 \cup \ldots \cup X'_t$ соответственно и $X \cap X' = X_s = X'_0$. Тогда естественным образом можно определить сеть длины $s+t$ с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_s \cup \cup X'_1 \cup \ldots \cup X'_t$, которую будем называть *произведением* сетей Σ и Σ' и обозначать $\Sigma \cdot \Sigma'$.

Непосредственно из определений 2 и 3 следует, что произвольная сеть Σ длины t является произведением однослойных сетей: $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t$.

Определение 4. Пусть $(v_1, \ldots, v_n)$ — произвольный набор формул и Σ — однослойная сеть с множеством вершин $\{x_1^0, \ldots, x_n^{(0)}\} \cup \{x_1^{(1)}, \ldots, x_m^{(1)}\}$. Тогда определим набор формул $(w_1, \ldots, w_m)$ по следующим правилам:

— если вершине $x_j^{(1)}$ инцидентно единственное ребро $(x_i^{(0)}, x_j^{(1)})$, то полагаем $w_j = v_i$;
— если вершине $x_j^{(1)}$ инцидентны рёбра $(x_{i_1}^{(0)}, x_j^{(1)})$ и $(x_{i_2}^{(0)}, x_j^{(1)})$ с метками l и r соответственно, то полагаем $w_j = v_{i_1} * v_{i_2}$.

При этом будем говорить, что однослойная сеть Σ *описывает преобразование* набора формул $(v_1, \ldots, v_n)$ в набор формул $(w_1, \ldots, w_m)$. Произвольная сеть Σ является произведением однослойных сетей, являющихся её слоями, и потому естественным образом описывает последовательность преобразований набора формул.

Пусть $F \in \mathcal{Q}(\Omega)$ — произвольная квазигруппа и сеть Σ описывает последовательность преобразований набора переменных $(x_1, \ldots, x_n)$ в набор формул $(w_1, \ldots, w_m)$. Тогда отображение $(w_1^F, \ldots, w_m^F) \colon \Omega^n \to \Omega^m$ будем обозначать Σ^F.

Нетрудно понять, что если $\Sigma = \Sigma_1 \cdot \Sigma_2$, то при выборе любой квазигруппы F справедливо соответствующее равенство отображений $\Sigma^F = \Sigma_1^F \cdot \Sigma_2^F$.

Определение 5. Будем говорить, что сети Σ и Σ' *эквивалентны для множества* Ω, если при выборе любой квазигруппы $F \in \mathcal{Q}(\Omega)$ отображения Σ^F и Σ'^F совпадают. Будем говорить, что сети Σ и Σ' *эквивалентны*, если они эквивалентны для любого множества.

Замечание 1. Если сети Σ и Σ' описывают преобразование набора переменных $(x_1, \ldots, x_n)$ в наборы формул $(w_1, \ldots, w_m)$ и $(w'_1, \ldots, w'_m)$ соответственно, то совпадение указанных наборов формул является достаточным условием для эквивалентности сетей Σ и Σ'.

Определение 6. Сеть Σ будем называть *биективной для множества* Ω, если при выборе любой квазигруппы $F \in \mathcal{Q}(\Omega)$ отображение Σ^F является биективным. Сеть Σ будем называть *биективной*, если она биективна для любого множества.

Очевидно, что для биективности сети Σ с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$ необходимо, чтобы множества начальных и конечных вершин были равномощны, то есть выполнялось равенство $|X_0| = |X_t|$.

Определение 7. Сеть Σ с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$ будем называть *сетью постоянной ширины*, если $|X_0| = |X_1| = \ldots = |X_t|$.

В данной работе рассматриваются только сети постоянной ширины, поэтому будем использовать термин «*сеть*», подразумевая при этом *сеть постоянной ширины*.

Поскольку любая сеть Σ биективна для произвольного одноэлементного множества Ω и, более того, в таком случае любые две сети Σ и Σ' одинаковой ширины представляют одно и то же единственное отображение $\Omega^n \to \Omega^n$, то в дальнейшем всегда будем полагать, что $|\Omega| \geqslant 2$.

1. Строение биективных сетей

Пусть Σ — сеть с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$. Тогда $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t$ и при выборе любой квазигруппы $F \in \mathcal{Q}(\Omega)$ выполняется равенство $\Sigma^F = \Sigma_1^F \cdot \ldots \cdot \Sigma_t^F$. Поэтому Σ биективна для множества Ω в том и только в том случае, когда каждый слой Σ_s, $s \in \{1, \ldots, t\}$, биективен для множества Ω.

Для однослойной сети Σ с множеством вершин $\{x_1^{(0)}, \ldots, x_n^{(0)}\} \cup \{x_1^{(1)}, \ldots, x_n^{(1)}\}$ определим $(0,1)$-матрицу связности $A_\Sigma = (a_{ij})_{n \times n}$ по следующему правилу: $a_{ij} = 1$ в случае, если сеть Σ содержит ребро $(x_i^{(0)}, x_j^{(1)})$, и $a_{ij} = 0$ — в противном случае.

Напомним, что диагональю в матрице размера $n \times n$ называют всякую совокупность из n её попарно неколлинеарных элементов, при этом диагональ называется положительной, если все её элементы положительны [2, 3].

Теорема 1. Однослойная сеть Σ является биективной для множества Ω тогда и только тогда, когда матрица A_Σ обладает единственной положительной диагональю.

Доказательство. Доказательство критерия проведём индукцией по ширине сети. База для сети ширины 1 очевидна. В предположении, что критерий верен для любой однослойной биективной сети ширины строго меньше чем n, докажем его для произвольной однослойной биективной сети Σ ширины n.

Необходимость. По определению сети Σ каждый столбец матрицы A_Σ содержит хотя бы одну единицу. При этом если сеть Σ биективна для множества Ω, то в матрице A_Σ найдётся хотя бы один столбец, содержащий ровно одну единицу, так как в противном случае при выборе произвольной квазигруппы $F \in \mathcal{Q}(\Omega)$ отображение Σ^F будет действовать следующим образом:

$$(a_1, a_2, \ldots, a_n) \mapsto (F(a_{i_1}, a_{j_1}), F(a_{i_2}, a_{j_2}), \ldots, F(a_{i_n}, a_{j_n})).$$

Нетрудно понять, что при выборе квазигруппы $F \in \mathcal{Q}(\Omega)$ со свойством $F(a,a) = b$ для всех $a \in \Omega$ и некоторого фиксированного $b \in \Omega$ (указанным квазигруппам соответствуют латинские квадраты, у которых на главной диагонали стоит только элемент b), система

$$(F(a,a), F(a,a), \ldots, F(a,a)) = (b, b, \ldots, b)$$

будет иметь $|\Omega| \geqslant 2$ решений, и отображение Σ^F не может быть биективным — противоречие. Значит, в матрице A_Σ существует столбец, содержащий одну единицу. Не ограничивая общности, можно считать, что это первый столбец, а единица в нём расположена на пересечении с первой строкой.

Пусть в первой строке матрицы A_Σ содержится r единиц; не ограничивая общности, будем считать, что они стоят на первых r местах. Тогда при выборе произвольной квазигруппы $F \in \mathcal{Q}(\Omega)$ отображение Σ^F будет действовать следующим образом:

$$(a_1, a_2, \ldots, a_n) \mapsto (a_1, F'\{a_1, a_{i_2}\}, \ldots, F'\{a_1, a_{i_r}\}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}),$$

где $\{i_2, \ldots, i_n, j_{r+1}, \ldots, j_n\} \subseteq \{2, \ldots, n\}$, а запись $F'\{a_i, a_j\}$ означает либо $F(a_i, a_j)$, либо $F(a_j, a_i)$, либо просто a_i. Поскольку сеть Σ является биективной для множества Ω, очевидно, что действие отображения Σ^F можно уточнить:

$$(a_1, a_2, \ldots, a_n) \mapsto (a_1, F\{a_1, a_{i_2}\}, \ldots, F\{a_1, a_{i_r}\}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}),$$

где запись $F\{a_i, a_j\}$ означает либо $F(a_i, a_j)$, либо $F(a_j, a_i)$. Биективность отображения Σ^F равносильна однозначной разрешимости при любых $b_1, \ldots, b_n \in \Omega$ системы уравнений

$$(a_1, F\{a_1, a_{i_2}\}, \ldots, F\{a_1, a_{i_r}\}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}) = (b_1, b_2, \ldots, b_n)$$

относительно $a_1, \ldots, a_n \in \Omega$. Это, в свою очередь, равносильно однозначной разрешимости при любых $b_1, \ldots, b_n \in \Omega$ следующей системы:

$$(F\{b_1, a_{i_2}\}, \ldots, F\{b_1, a_{i_r}\}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}) = (b_2, \ldots, b_r, b_{r+1}, \ldots, b_n),$$

которую можно переписать в эквивалентном виде

$$(a_{i_2}, \ldots, a_{i_r}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}) = (F^{-1}\{b_1, b_2\}, \ldots, F^{-1}\{b_1, b_r\}, b_{r+1}, \ldots, b_n),$$

где запись $F^{-1}\{b_1, b_j\}$ означает либо решение уравнения $F(b_1, x) = b_j$, либо решение уравнения $F(x, b_1) = b_j$. Однозначная разрешимость последней системы при любых $b_1, \ldots, b_n \in \Omega$ равносильна однозначной разрешимости системы

$$(a_{i_2}, \ldots, a_{i_r}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}) = (b'_2, \ldots, b'_r, b_{r+1}, \ldots, b_n)$$

при любых $b'_2, \ldots, b'_r, b_{r+1}, \ldots, b_n \in \Omega$.

Обозначим подграф сети Σ, основанный на множестве вершин $\{x_2^{(0)}, \ldots, x_n^{(0)}\} \cup \cup \{x_2^{(1)}, \ldots, x_n^{(1)}\}$, через Σ'. Легко видеть, что Σ' является однослойной сетью ширины $(n-1)$ и однозначная разрешимость системы

$$(a_{i_2}, \ldots, a_{i_r}, F'\{a_{i_{r+1}}, a_{j_{r+1}}\}, \ldots, F'\{a_{i_n}, a_{j_n}\}) = (b'_2, \ldots, b'_r, b_{r+1}, \ldots, b_n)$$

при любых $b'_2, \ldots, b'_r, b_{r+1}, \ldots, b_n \in \Omega$ на самом деле означает биективность отображения Σ'^F. Таким образом, показано, что при выборе любой квазигруппы $F \in \mathcal{Q}(\Omega)$ биективность отображения Σ^F равносильна биективности отображения Σ'^F. Другими словами, сеть Σ является биективной для множества Ω в том и только в том случае, когда сеть Σ' является биективной для множества Ω.

По предположению индукции матрица $A_{\Sigma'} = A_\Sigma \left(\begin{smallmatrix} 2 & \cdots & n \\ 2 & \cdots & n \end{smallmatrix}\right)$ содержит единственную положительную диагональ, которая однозначно продолжается до единственной положительной диагонали матрицы A_Σ.

Д о с т а т о ч н о с т ь. Пусть матрица A_Σ имеет единственную положительную диагональ. Тогда в матрице A_Σ существует столбец, в котором содержится ровно одна единица, так как в противном случае в результате последовательного вычёркивания всех строк матрицы A_Σ, содержащих ровно одну единицу, вместе с соответствующими им столбцами останется подматрица $A_\Sigma \begin{pmatrix} i_1 & \cdots & i_l \\ j_1 & \cdots & j_l \end{pmatrix}$, $2 \leqslant l \leqslant n$, которая содержит в каждой строке не менее двух единиц и в каждом столбце ровно по две единицы. Нетрудно понять, что такая матрица содержит в каждой строке и в каждом столбце ровно две единицы и, следовательно, имеет не менее двух положительных диагоналей, которые продолжаются до различных положительных диагоналей матрицы A_Σ, — противоречие. Значит, в матрице A_Σ существует столбец, содержащий одну единицу, и, не ограничивая общности, можно считать, что это первый столбец, а единица в нём расположена на пересечении с первой строкой.

Пусть в первой строке матрицы A_Σ содержится r единиц, будем считать, что они стоят на первых r местах. Тогда, аналогично предыдущей части доказательства,

нетрудно показать, что сеть Σ является биективной для множества Ω в том и только в том случае, когда сеть Σ' является биективной для множества Ω. При этом матрица $A_{\Sigma'} = A_\Sigma \begin{pmatrix} 2 & \dots & n \\ 2 & \dots & n \end{pmatrix}$ имеет единственную положительную диагональ, поскольку всякая положительная диагональ матрицы A_Σ содержит единственную единицу из первого столбца и является продолжением некоторой положительной диагонали матрицы $A_\Sigma \begin{pmatrix} 2 & \dots & n \\ 2 & \dots & n \end{pmatrix}$, и по предположению индукции сеть Σ' является биективной для множества Ω. ■

Следствие 1. Следующие утверждения являются равносильными:

1) сеть Σ является биективной для некоторого множества Ω;
2) сеть Σ является биективной.

Доказательство. $1 \Rightarrow 2$. Если сеть Σ с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$ является биективной для некоторого множества Ω, то каждый её слой Σ_s, $s \in \{1, \ldots, t\}$, является биективным для данного множества. Из теоремы 1 следует, что биективность однослойных сетей Σ_s, $s \in \{1, \ldots, t\}$, для множества Ω не зависит от природы этого множества и его мощности, а зависит только от строения указанных однослойных сетей. Значит, каждый слой Σ_s, $s \in \{1, \ldots, t\}$, является биективным для всех множеств и, следовательно, сама сеть Σ является биективной для всех множеств.

$2 \Rightarrow 1$. Очевидно. ■

Ввиду следствия 1 в определении биективной сети достаточно требовать биективность только для одного множества, мощность которого больше чем 1.

Следствие 2. Пусть Σ — биективная однослойная сеть. Тогда:

1) сеть Σ содержит вершины со степенью захода 1;
2) сеть Σ содержит вершины со степенью исхода 1.

Доказательство.

1) Согласно теореме 1, матрица A_Σ имеет единственную положительную диагональ, и в доказательстве теоремы 1 показано, что существует столбец матрицы A_Σ, в котором содержится ровно одна единица.

2) Согласно критерию биективности однослойной сети, в матрице A_Σ отсутствуют нулевые строки. Кроме того, невозможно, чтобы каждая строка матрицы A_Σ содержала более одной единицы, так как в этом случае каждый столбец матрицы A_Σ будет содержать ровно две единицы, а сама матрица A_Σ будет иметь не менее двух положительных диагоналей, что противоречит биективности сети Σ. ■

Пример 1. Легко проверить, что преобразование набора переменных $(x_1, \ldots, x_8)$ в набор формул $(x_2 * (x_4 * x_5), x_1, x_4 * x_5, x_2 * x_3, x_5, (x_2 * x_3) * (x_6 * x_7), x_7 * x_8, x_5 * x_6)$ может быть описано (при подходящей разметке дуг) сетью, приведённой на рис. 1. Также нетрудно убедиться, что это преобразование набора переменных может быть описано сетью, приведённой на рис. 2. Значит, эти сети эквивалентны.

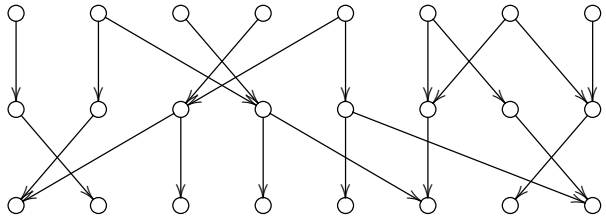

Рис. 1

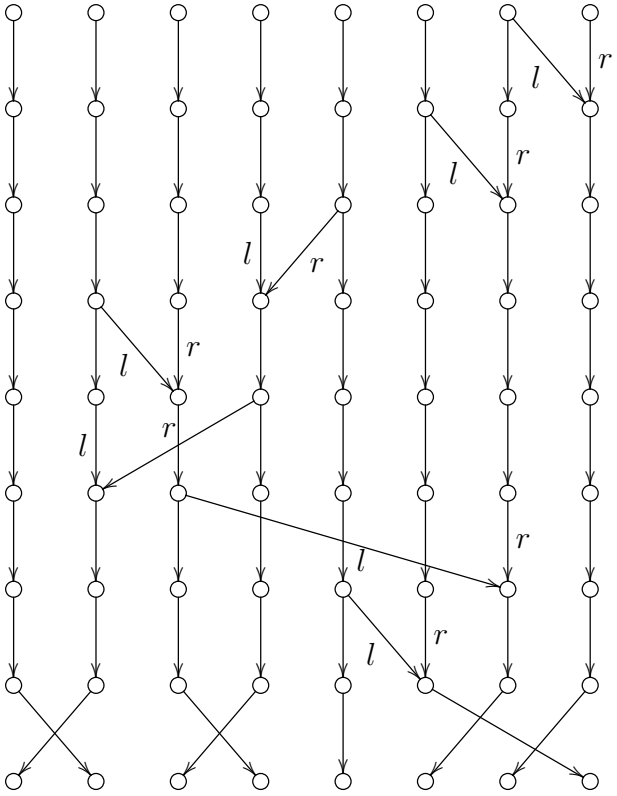

Рис. 2

Определение 8. Пусть Σ — однослойная сеть с множеством вершин $X_0 \cup X_1$. Вершину $x_i^{(0)} \in X_0$ сети Σ будем называть *неподвижной*, если Σ содержит ребро $(x_i^{(0)}, x_i^{(1)})$. Сеть Σ будем называть *элементарной*, если все вершины из множества X_0 неподвижны и ровно одна вершина из множества X_1 имеет степень захода 2.

Элементарную сеть с множеством вершин $X_0 \cup X_1$, которая содержит рёбра $(x_i^{(0)}, x_i^{(1)})$ и $(x_j^{(0)}, x_i^{(1)})$, будем обозначать $\Sigma_i^{\{i,j\}}$. В случае, когда ребро $(x_i^{(0)}, x_i^{(1)})$ имеет метку l, обозначение можно уточнить как $\Sigma_i^{(i,j)}$, а если оно имеет метку r — как $\Sigma_i^{(j,i)}$.

Произвольная элементарная сеть всегда является биективной. Ещё одним важным примером биективных сетей являются сети с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$, у которых степень захода каждой вершины $x_j^{(s)}$, $s \in \{1, \ldots, t\}$, равна 1. Такие сети будем называть *перестановочными*. Произвольная перестановочная сеть определяет

отображение $\Omega^n \to \Omega^n$, не зависящее от выбора квазигруппы $F \in \mathcal{Q}(\Omega)$ и действующее на множестве Ω^n как перестановка координат вектора. Отсюда следует, что любая перестановочная сеть эквивалентна однослойной перестановочной сети. Произвольная перестановочная сеть эквивалентна произведению перестановочных сетей, у каждой из которых ровно две вершины не являются неподвижными, — это следует из известного результата о представлении произвольной перестановки в виде произведения транспозиций.

Элементарные и перестановочные сети являются примерами простейших биективных сетей, однако, как показывает следующая теорема, этих примитивов достаточно для реализации произвольной биективной сети.

Теорема 2. Сеть Σ является биективной в том и только в том случае, когда она эквивалентна произведению

$$\Sigma_{R1} \cdot \ldots \cdot \Sigma_{Rt} \cdot \Pi_R \text{ (или } \Pi_L \cdot \Sigma_{L1} \cdot \ldots \cdot \Sigma_{Lt}),$$

где $\Sigma_{R1}, \ldots, \Sigma_{Rt}$ $(\Sigma_{L1}, \ldots, \Sigma_{Lt})$ — элементарные сети; Π_R (Π_L) — однослойная перестановочная сеть. При этом количество элементарных сетей в произведении равно количеству вершин сети Σ со степенью захода 2.

Доказательство. Достаточность очевидна. Для доказательства необходимости потребуются два вспомогательных утверждения.

Лемма 1. Пусть Σ — произвольная биективная однослойная сеть, а Π_1, Π_2 — такие однослойные перестановочные сети, что корректно определить произведения $\Pi_1 \cdot \Sigma$ и $\Sigma \cdot \Pi_2$. Тогда справедливы следующие утверждения:

1) сеть $\Pi_1 \cdot \Sigma$ эквивалентна однослойной сети;
2) сеть $\Sigma \cdot \Pi_2$ эквивалентна однослойной сети;
3) сеть Σ эквивалентна произведению $\Sigma_R \cdot \Pi_R$ однослойной биективной сети Σ_R, у которой все вершины неподвижны, и однослойной перестановочной сети Π_R;
4) сеть Σ эквивалентна произведению $\Pi_L \cdot \Sigma_L$ однослойной биективной сети Σ_L, у которой все вершины неподвижны, и однослойной перестановочной сети Π_L.

Доказательство. Доказательство утверждений 1 и 2, с использованием замечания 1, очевидно.

Утверждение 3: согласно критерию биективности однослойной сети, матрица A_Σ обладает единственной ненулевой диагональю, расположенной на местах $(1, i_1), \ldots, (n, i_n)$. Перестановка $\begin{pmatrix} 1 & \ldots & n \\ i_1 & \ldots & i_n \end{pmatrix}$ определяет перестановочную сеть Π_R. Сеть $\Sigma \cdot \Pi_R^{-1}$ описывает преобразование формул, соответствующее некоторой однослойной сети Σ_R, у которой все вершины неподвижны. Таким образом, сети Σ и $\Sigma_R \cdot \Pi_R$ описывают преобразование набора переменных в один набор формул и, согласно замечанию 1, они эквивалентны.

Доказательство утверждения 4 аналогично. ■

Лемма 2. Пусть Σ — биективная однослойная сеть, у которой все вершины неподвижны и хотя бы одна имеет степень захода 2. Тогда сеть Σ эквивалентна произведению элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_t$, где t — число вершин сети Σ со степенью захода 2.

Доказательство. Пусть Σ — однослойная биективная сеть ширины n, у которой все вершины неподвижны. Тогда количество вершин сети Σ со степенью захода 2 равно $\omega(A_\Sigma) - n$, где $\omega(A_\Sigma)$ — количество ненулевых элементов матрицы A_Σ.

Доказательство утверждения проведём индукцией по параметру $\omega(A_\Sigma) - n$. База при $\omega(A_\Sigma) - n = 1$ очевидна: Σ является элементарной сетью ширины n.

Предположим, что утверждение верно для любой однослойной биективной сети Σ ширины n при условии $\omega(A_\Sigma) - n < t$. Докажем утверждение для произвольной однослойной биективной сети Σ ширины n с условием $\omega(A_\Sigma) - n = t$.

Нетрудно понять, что $\omega(A_\Sigma) < 2n$, так как в противном случае каждый столбец матрицы A_Σ содержит ровно две единицы и в сети Σ отсутствуют вершины со степенью захода 1, что противоречит следствию 2. Таким образом, $t = \omega(A_\Sigma) - n < n$ и в матрице A_Σ в точности t столбцов содержат две единицы; не ограничивая общности, будем считать, что это первые t столбцов.

Поскольку все вершины сети Σ неподвижны, то элементы матрицы A_Σ, расположенные на местах $(1,1),\ldots,(n,n)$, очевидно, образуют положительную диагональ. При этом все $2t$ единиц, расположенных в первых t столбцах, не могут находиться на пересечении с первыми t строками, так как в противном случае матрица $A_\Sigma \begin{pmatrix} 1 & \ldots & t \\ 1 & \ldots & t \end{pmatrix}$ имеет две различные трансверсали, которые продолжаются до различных трансверсалей матрицы A_Σ, что противоречит биективности сети Σ. Значит, хотя бы одна из последних $(n-t)$ строк матрицы A_Σ содержит не менее двух единиц; не ограничивая общности, будем считать, что это строка с номером $(t+1)$.

Пусть в $(t+1)$-й строке матрицы A_Σ присутствует единица на месте $s \in \{1,\ldots,t\}$. Тогда обозначим через Σ' однослойную сеть, полученную из сети Σ удалением ребра $(x_{t+1}^{(0)}, x_s^{(1)})$. Очевидно, что $A_{\Sigma'} = A_\Sigma - E_{(t+1)s}$, где $E_{(t+1)s}$ — $(0,1)$-матрица, у которой единственная единица стоит на пересечении $(t+1)$-й строки и s-го столбца. Поскольку матрица A_Σ обладает единственной трансверсалью, а её главная диагональ не содержит нулей, матрица $A_{\Sigma'}$ также обладает единственной трансверсалью, расположенной на главной диагонали. Последнее означает, что сеть Σ' является биективной.

При выборе подходящей элементарной сети $\Sigma_s^{\{s,t+1\}}$ произведение $\Sigma' \cdot \Sigma_s^{\{s,t+1\}}$ описывает такое же преобразование набора переменных $(x_1,\ldots,x_n)$, что и сеть Σ. Значит, согласно замечанию 1, сети Σ и $\Sigma' \cdot \Sigma_s^{\{s,t+1\}}$ являются эквивалентными. При этом $\omega(A_{\Sigma'}) - n = \omega(A_\Sigma) - 1 - n = t - 1$ и по предположению индукции однослойная сеть Σ' эквивалентна произведению элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_{t-1}$.

Таким образом, доказали, что однослойная биективная сеть Σ ширины n эквивалентна произведению элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_{t-1} \cdot \Sigma_s^{\{s,t+1\}}$, длина которого удовлетворяет равенству $t = \omega(A_\Sigma) - n$. ■

Теперь, для завершения доказательства теоремы, остаётся воспользоваться представлением сети Σ в виде произведения собственных слоёв и применить к ним доказанные леммы. ■

Количество вершин сети Σ со степенью захода 2 будем называть *весом сети* Σ или её *сложностью* и обозначать $\|\Sigma\|$.

2. Разметка биективных сетей

Введём понятие разметки сети — инструмента, который позволяет обнаруживать особенности биективной сети, нарушающие её транзитивность, и докажем критерий эквивалентности биективных сетей и единственность указанного в теореме 2 представления.

Учитывая результаты, полученные в п. 1, не ограничивая общности, будем считать, что произвольная биективная сеть Σ представляет собой произведение $\Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi$

с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t \cup X_{t+1}$, где $\Sigma_1, \ldots, \Sigma_t$ — элементарные сети; Π — однослойная перестановочная сеть. Также, не ограничивая общности, будем считать, что $\Omega \subset \mathbb{N}$.

Определение 9. Если для элементов $y_1, y_2, y_3 \in \mathbb{N}$ и частично определённого отображения $F\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ выполняется соотношение $F(y_1, y_2) = y_3$, то будем говорить, что элементы y_1 и y_2 *содержатся в области определения отображения* F, а элемент y_3 *содержится в области значений отображения* F.

Определение 10. Частично определённое отображение $F\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, удовлетворяющее условию

$$(F(y_1, y_2) = F(y_1', y_2')) \Longrightarrow ((y_1, y_2) = (y_1', y_2')) \text{ или } (y_1 \neq y_1', y_2 \neq y_2')$$

при всех допустимых $y_1, y_2, y_1', y_2' \in \mathbb{N}$, будем называть *частично определённым отображением без противоречий* (или *частично определённым непротиворечивым отображением*).

Определение 11. *Разметкой* сети $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi$ будем называть произвольное отображение $\mu\colon X_0 \cup X_1 \cup \ldots \cup X_t \cup X_{t+1} \to \mathbb{N}$. Пусть $F\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ — частично определённое отображение. Тогда разметку μ сети Σ, которая удовлетворяет следующим условиям:

- для всех $s \in \{1, \ldots, t\}$ и $i \in \{1, \ldots, n\}$:
  - если $\deg^- x_i^{(s)} = 1$, то $\mu(x_i^{(s)}) = \mu(x_i^{(s-1)})$;
  - если $\deg^- x_i^{(s)} = 2$ и рёбра $(x_i^{(s-1)}, x_i^{(s)})$, $(x_j^{(s-1)}, x_i^{(s)})$ имеют метки l и r соответственно, то $\mu(x_i^{(s)}) = F(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)}))$;
  - если $\deg^- x_i^{(s)} = 2$ и рёбра $(x_i^{(s-1)}, x_i^{(s)})$, $(x_j^{(s-1)}, x_i^{(s)})$ имеют метки r и l соответственно, то $\mu(x_i^{(s)}) = F(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)}))$;
- если перестановочная сеть Π содержит рёбра $(x_{i_k}^{(t)}, x_k^{(t+1)})$, $k \in \{1, \ldots, n\}$, то выполняются равенства $\mu(x_k^{(t+1)}) = \mu(x_{i_k}^{(t)})$, $k \in \{1, \ldots, n\}$,

будем называть *правильной относительно* F. При этом само отображение F будем называть *правилом* разметки μ.

Определение 12. Пусть μ — разметка сети Σ с правилом F и при этом никакое сужение частичного отображения F не является правилом разметки μ. Тогда будем говорить, что F является *минимальным правилом* разметки μ. Нетрудно понять, что минимальное правило разметки μ определено однозначно. Правильную разметку μ будем называть *непротиворечивой*, если её минимальное правило является непротиворечивым отображением.

Определение 13. Если для разметки μ сети Σ выполняется система равенств $\{\mu(x_{i_j}^{(s_j)}) = v_j : j \in J\}$, то будем говорить, что μ — *разметка сети* Σ *с условиями* $\{\mu(x_{i_j}^{(s_j)}) = v_j : j \in J\}$. Система равенств $\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_n$ называется *начальным условием* разметки μ, при этом говорят, что μ — разметка с начальным условием $(v_1, \ldots, v_n)$.

Каждая правильная разметка сети Σ однозначно определяется своим начальным условием и правилом. В тех случаях, когда при некоторой разметке вершин $x_1^{(0)}, \ldots, x_n^{(0)}$ сети Σ для полного задания правильной разметки не хватает области определения частично определённого отображения F, можно непротиворечивым образом расширить область определения F и тем самым определить разметку с правилом F. Поясним это подробнее.

Пусть задана начальная разметка $\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_n$ сети Σ и $\mathbb{N}\backslash\{v_1, \ldots, v_n\} \supset \{y_1, y_2, \ldots\}$ — счётное множество меток, которые не содержатся ни в области определения, ни в области значений правила F (при этом возможно, что метки $v_1, \ldots, v_n$ также не содержатся ни в области определения, ни в области значений правила F). Тогда продолжим разметку μ сети Σ по следующему правилу:

— для всех $s \in \{1, \ldots, t\}$ при $\Sigma_s = \Sigma_i^{\{i,j\}}$ положим $\mu(x_l^{(s)}) = \mu(x_l^{(s-1)})$, если $l \neq i$, а для разметки вершины $x_i^{(s)}$ возможны следующие варианты:
  – если $\Sigma_s = \Sigma_i^{(i,j)}$ и значение $F\left(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)})\right)$ определено, то пометим вершину $x_i^{(s)}$ меткой $F\left(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)})\right)$, в противном случае пометим вершину $x_i^{(s)}$ ранее не использованной меткой y_s и определим $F\left(\mu(x_i^{(s-1)}), \mu(x_j^{(s-1)})\right) = y_s$;
  – если $\Sigma_s = \Sigma_i^{(j,i)}$ и значение $F\left(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)})\right)$ определено, то пометим вершину $x_i^{(s)}$ меткой $F\left(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)})\right)$, в противном случае пометим вершину $x_i^{(s)}$ ранее не использованной меткой y_s и определим $F\left(\mu(x_j^{(s-1)}), \mu(x_i^{(s-1)})\right) = y_s$;

— если перестановочная сеть Π содержит рёбра $(x_{i_k}^{(t)}, x_k^{(t+1)})$, $k \in \{1, \ldots, n\}$, то положим $\mu(x_k^{(t+1)}) = \mu(x_{i_k}^{(t)})$, $k \in \{1, \ldots, n\}$.

При проведении разметки μ сети Σ описанным способом частично определённое (непротиворечивое) отображение F корректным образом продолжается до частично определённого (непротиворечивого) отображения, которое будем обозначать $F_{\Sigma,\mu}$, при этом построенная разметка μ является правильной относительно $F_{\Sigma,\mu}$.

Определение 14. Описанную процедуру продолжения разметки μ и расширения области определения F будем называть *свободным продолжением начальной разметки* $\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_n$ *и её правила* F *относительно сети* Σ.

Пусть η и μ — разметки сети Σ и для отображения $\sigma_\mu \colon \mathbb{N} \to \mathbb{N}$ справедливо соотношение $\sigma_\mu \circ \eta = \mu$, то есть при всех $s \in \{0, \ldots, t+1\}$ и $i \in \{1, \ldots, n\}$ выполняется равенство $\sigma_\mu(\eta(x_i^{(s)})) = \mu(x_i^{(s)})$. Тогда будем обозначать это условие как $\sigma_\mu \colon \eta \to \mu$.

Определение 15. Правильную разметку η сети Σ с начальным условием $(v_1, \ldots, v_n)$ будем называть *свободной*, если для любой правильной разметки μ сети Σ с начальным условием $(v_1, \ldots, v_n)$ существует отображение σ_μ, удовлетворяющее условию $\sigma_\mu \colon \eta \to \mu$.

Непосредственно из определения свободной разметки следует, что при условии существования свободная разметка сети Σ с начальным условием $(v_1, \ldots, v_n)$ определена однозначно с точностью до обратимого переобозначения меток.

Теорема 3. Пусть разметка η получена в результате свободного продолжения начальной разметки $\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n$ и пустого правила G относительно сети Σ. Тогда η — свободная разметка сети Σ с начальным условием $(v_1, \ldots, v_n)$, а отображение $G_{\Sigma,\eta}$ — её минимальное правило.

Доказательство. Из определения процедуры свободного продолжения начальной разметки $\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n$ и её пустого правила G относительно сети Σ следует, что разметка η является непротиворечивой и её минимальное правило $G_{\Sigma,\eta}$ удовлетворяет условию

$$(G_{\Sigma,\eta}(z_1, z_2) = G_{\Sigma,\eta}(z_1', z_2')) \Longrightarrow ((z_1, z_2) = (z_1', z_2')) \tag{1}$$

при всех допустимых $z_1, z_2, z_1', z_2' \in \mathbb{N}_0$. Пусть μ — произвольная правильная разметка сети Σ с теми же начальными условиями, что и разметка η. Тогда для доказательства существования отображения $\sigma_\mu\colon \eta \to \mu$ достаточно показать, что при совпадении меток $\eta(x_i^{(s)}) = \eta(x_j^{(r)})$ также выполняется равенство $\mu(x_i^{(s)}) = \mu(x_j^{(r)})$.

Не ограничивая общности, будем считать, что $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t$. Докажем утверждение индукцией по длине произведения $\Sigma_1 \cdot \ldots \cdot \Sigma_t$. База индукции при $t = 1$ очевидна.

Пусть теперь $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_{t-1} \cdot \Sigma_t$ — сеть длины $t > 1$ и $\Sigma_t = \Sigma_k^{\{k,l\}}$. Рассмотрим все возможные варианты для пары вершин $x_i^{(s)}$ и $x_j^{(r)}$:

1) Если $r, s < t$, то истинность утверждения следует из предположения индукции.
2) Если $r < t, s = t$ и $k \neq i$, то выполняется равенство $\eta(x_i^{(s-1)}) = \eta(x_j^{(r)})$, и остаётся воспользоваться предположением индукции.
3) Если $r = t, s < t$ и $k \neq j$, то выполняется равенство $\eta(x_i^{(s)}) = \eta(x_j^{(r-1)})$, и остаётся воспользоваться предположением индукции.
4) Если $r = s = t$ и $k \notin \{i, j\}$, то выполняется равенство $\eta(x_i^{(s-1)}) = \eta(x_j^{(r-1)})$, и остаётся воспользоваться предположением индукции.
5) В случае, когда $s = t$ и $\Sigma_t = \Sigma_i^{\{i,l\}}$, не ограничивая общности, будем считать, что $\Sigma_t = \Sigma_i^{(i,l)}$. Из определения процедуры свободного продолжения разметки следует, что $\eta(x_i^{(s)}) \notin \{v_1, \ldots, v_n\}$, а минимальное правило разметки η удовлетворяет условию (1). Значит, равенство меток $\eta(x_i^{(s)}) = \eta(x_j^{(r)})$ влечёт за собой совпадение упорядоченных наборов меток $(\eta(x_i^{(s-1)}), \eta(x_l^{(s-1)}))$ и, не ограничивая общности, $(\eta(x_j^{(r')}), \eta(x_{l'}^{(r')}))$, где $r' \in \{0, \ldots, r-1\}$ — наибольшее со свойством $\eta(x_j^{(r')}) \neq \eta(x_j^{(r)})$. По предположению индукции упорядоченные наборы меток $(\mu(x_i^{(s-1)}), \mu(x_l^{(s-1)}))$ и $(\mu(x_j^{(r')}), \mu(x_{l'}^{(r')}))$ также совпадают и, следовательно, выполняются равенства $\mu(x_i^{(s)}) = \mu(x_j^{(r'+1)}) = \mu(x_j^{(r)})$.
6) Доказательство случая, когда $r = t$ и $\Sigma_t = \Sigma_j^{\{j,l\}}$, аналогично доказательству 5.

Теорема доказана. ■

Замечание 2. Поскольку свободная разметка сети Σ с начальным условием $(v_1, \ldots, v_n)$ определена однозначно с точностью до обратимого переобозначения меток, то, не ограничивая общности, можно считать, что произвольная свободная разметка η сети Σ может быть получена при помощи свободного продолжения соответствующей начальной разметки $\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n$ и её пустого правила G относительно сети Σ.

Следующая теорема фактически оправдывает название свободной разметки.

Теорема 4. Пусть η — свободная разметка сети Σ, μ — правильная разметка сети Σ, и возможно определить отображение σ_μ по правилу $\sigma_\mu(\eta(x_i^{(0)})) = \mu(x_i^{(0)})$, $i \in \{1, \ldots, n\}$. Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu\colon \eta \to \mu$.

Доказательство. Для доказательства теоремы 4 достаточно показать, что при совпадении меток $\eta(x_i^{(s)}) = \eta(x_j^{(r)})$ также выполняется равенство $\mu(x_i^{(s)}) = \mu(x_j^{(r)})$. Это устанавливается индукцией по длине сети Σ аналогично доказательству теоремы 3. ■

Следствие 3. В условиях теоремы 4, если G и F — минимальные правила разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$.

Среди множества всех свободных разметок сети Σ особенным образом выделяются два типа свободной разметки.

Определение 16. Свободную разметку сети Σ с начальным условием $(v_1, \ldots, v_1)$ будем называть *минимальной свободной разметкой сети* Σ и обозначать $\eta_{\min}$. Свободную разметку сети Σ с начальным условием $(v_1, \ldots, v_n)$ будем называть *максимальной свободной разметкой сети* Σ, если все метки $v_1, \ldots, v_n$ — различны, и обозначать $\eta_{\max}$.

Следствие 4. Для произвольной свободной разметки η сети Σ существует отображение $\sigma_\eta \colon \eta \to \eta_{\min}$.

Следствие 5. Для произвольной свободной разметки η сети Σ существует отображение $\sigma_\eta \colon \eta_{\max} \to \eta$.

Замечание 3. На множестве всех свободных разметок фиксированной сети Σ можно естественным образом ввести отношение эквивалентности $\sim$: класс $[\eta]_\sim$ содержит все свободные разметки сети Σ, которые могут быть получены из свободной разметки η с помощью обратимого переобозначения меток. При этом фактор-множество всех свободных разметок по данному отношению эквивалентности является частично упорядоченным множеством с единственными минимальным и максимальным элементами $[\eta_{\min}]_\sim$ и $[\eta_{\max}]_\sim$ соответственно.

В п. 1 доказано, что произвольная биективная сеть эквивалентна произведению элементарных и перестановочной сетей. Однако открытым остался вопрос об однозначности такого представления:

— во-первых, какие элементарные сети в указанном произведении допустимо менять местами?
— во-вторых, однозначен ли состав представления биективной сети в виде произведения элементарных и перестановочной сетей?

Ответ на первый вопрос дает следующее очевидное утверждение.

Утверждение 1. Пусть $\Sigma_{i_1}^{\{i_1,j_1\}}$ и $\Sigma_{i_2}^{\{i_2,j_2\}}$ — различные элементарные сети одной ширины. Тогда следующие утверждения эквивалентны:

1) сети $\Sigma_{i_1}^{\{i_1,j_1\}} \cdot \Sigma_{i_2}^{\{i_2,j_2\}}$ и $\Sigma_{i_2}^{\{i_2,j_2\}} \cdot \Sigma_{i_1}^{\{i_1,j_1\}}$ описывают одинаковые преобразования набора переменных;
2) сети $\Sigma_{i_1}^{\{i_1,j_1\}} \cdot \Sigma_{i_2}^{\{i_2,j_2\}}$ и $\Sigma_{i_2}^{\{i_2,j_2\}} \cdot \Sigma_{i_1}^{\{i_1,j_1\}}$ эквивалентны;
3) $i_1 \neq i_2$, $i_1 \neq j_2$ и $i_2 \neq j_1$.

Если для элементарных сетей $\Sigma_{i_1}^{\{i_1,j_1\}}$ и $\Sigma_{i_2}^{\{i_2,j_2\}}$ выполняются условия утверждения 1, то будем говорить, что для них допустима перестановка.

Введённый аппарат разметки позволяет доказать однозначность состава произведения элементарных и перестановочной сетей, представляющего биективную сеть. Для этого потребуется результат, известный как гипотеза Эванса [4] и в общем случае независимо доказанный в работах [5, 6].

Теорема 5 (B. Smetaniuk, 1981). Если частично определённое непротиворечивое отображение $F \colon \Omega \times \Omega \to \Omega$ определено не более чем на $|\Omega| - 1$ наборах, то оно продолжается до квазигруппы на множестве Ω.

Теорема 6. Если сети $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi$ и $\Sigma' = \Sigma'_1 \cdot \ldots \cdot \Sigma'_s \cdot \Pi'$ эквивалентны, то $\Pi = \Pi'$, а произведения элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ и $\Sigma'_1 \cdot \ldots \cdot \Sigma'_s$ имеют одинаковый состав и могут отличаться только допустимой перестановкой множителей. В частности, эквивалентные сети имеют одинаковую сложность.

Доказательство. Пусть Σ — сеть ширины n с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t \cup X_{t+1}$ и максимальная свободная разметка η сети Σ с начальным условием $(v_1, \ldots, v_n)$ получена в результате свободного продолжения начальной разметки $\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n$ и пустого правила относительно сети Σ. Тогда, согласно определению свободного продолжения разметки, в свободной разметке η используется множество меток $\Omega = \{v_1, \ldots, v_n, y_1, \ldots, y_t\}$, а её минимальное правило $G\colon \Omega \times \Omega \to \Omega$ определено на t различных наборах. Далее будем считать, что разметка η сети Σ удовлетворяет условиям

$$\eta(x_1{}^{(0)}) = v_1, \ldots, \eta(x_n{}^{(0)}) = v_n, \quad \eta(x_1{}^{(t+1)}) = y_{i_1}, \ldots, \eta(x_n{}^{(t+1)}) = y_{i_n},$$

где $\{y_{i_1}, \ldots, y_{i_n}\} \subset \{y_1, \ldots, y_t\}$.

Пусть Σ' — сеть с множеством вершин $X'_0 \cup X'_1 \cup \ldots \cup X'_s \cup X'_{s+1}$ и разметка η' сети Σ' получена в результате свободного продолжения начальной разметки $\eta'(x'_1{}^{(0)}) = v_1$, $\ldots$, $\eta'(x'_n{}^{(0)}) = v_n$ и её правила G относительно сети Σ'; для удобства непротиворечивое правило $G_{\Sigma',\eta'}$ будем обозначать G'. Тогда, согласно определению свободного продолжения разметки, в разметке η' используются метки из множества $\Omega' = \{v_1, \ldots, v_n, y_1, \ldots, y_t, \ldots, y_d\}$, а её правило $G'\colon \Omega' \times \Omega' \to \Omega'$ определено на d различных наборах. Далее будем считать, что разметка η' сети Σ' удовлетворяет условиям

$$\eta'(x'_1{}^{(0)}) = v_1, \ldots, \eta'(x'_n{}^{(0)}) = v_n, \quad \eta'(x'_1{}^{(s+1)}) = y_{j_1}, \ldots, \eta'(x'_n{}^{(s+1)}) = y_{j_n},$$

где $\{y_{j_1}, \ldots, y_{j_n}\} \subset \{y_1, \ldots, y_t, \ldots, y_d\}$.

Согласно гипотезе Эванса, непротиворечивое отображение G' продолжается до квазигруппы $\overline{G} \in \mathcal{Q}(\Omega')$, и для эквивалентных сетей Σ и Σ' справедливы соотношения

$$\Sigma^{\overline{G}}(v_1, \ldots, v_n) = \Sigma^{G}(v_1, \ldots, v_n) = (\eta(x_1^{(t+1)}), \ldots, \eta(x_n^{(t+1)})) = (y_{i_1}, \ldots, y_{i_n}),$$
$$\Sigma'^{\overline{G}}(v_1, \ldots, v_n) = \Sigma'^{G'}(v_1, \ldots, v_n) = (\eta'(x'_1{}^{(s+1)}), \ldots, \eta'(x'_n{}^{(s+1)})) = (y_{j_1}, \ldots, y_{j_n}).$$

Поскольку отображения $\Sigma^{\overline{G}}$ и $\Sigma'^{\overline{G}}$ совпадают, то $(y_{i_1}, \ldots, y_{i_n}) = (y_{j_1}, \ldots, y_{j_n})$. Последнее равенство означает, что на самом деле при свободном продолжении начальной разметки $\eta'(x'_1{}^{(0)}) = v_1, \ldots, \eta'(x'_n{}^{(0)}) = v_n$ и её правила G относительно сети Σ' дополнительные метки $y_{t+1}, \ldots, y_d$ не возникают и правило G' совпадает с G. Другими словами, разметка η' сети Σ' является G-правильной.

Далее потребуется вспомогательное утверждение.

Лемма 3. Пусть η — максимальная свободная разметка сети $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t$ с минимальным правилом G, а η' — G-правильная разметка сети $\Sigma' = \Sigma'_1 \cdot \ldots \cdot \Sigma'_s$ с тем же начальным условием. Тогда из равенства $\{\eta(x_1{}^{(t)}), \ldots, \eta(x_n{}^{(t)})\} = \{\eta'(x'_1{}^{(s)}), \ldots, \eta'(x'_n{}^{(s)})\}$ следует, что $t = s$, а произведения элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ и $\Sigma'_1 \cdot \ldots \cdot \Sigma'_s$ отличаются лишь допустимой перестановкой множителей.

Доказательство. Докажем утверждение индукцией по t. База при $t = 1$ очевидна.

Пусть теперь $t > 1$. Не ограничивая общности, будем считать, что максимальная свободная разметка η сети Σ получена в результате свободного продолжения начальной разметки $\eta(x_1^{(0)}) = y_{-1}, \ldots, \eta(x_n^{(0)}) = y_{-n}$ и пустого правила относительно сети Σ. Согласно определению свободного продолжения разметки, минимальное правило G свободной разметки η удовлетворяет условию

$$(G(y_i, y_j) = G(y_d, y_m)) \Longrightarrow ((y_i, y_j) = (y_d, y_m))$$

при всех допустимых $y_i, y_j, y_d, y_m \in \{y_{-n}, \ldots, y_{-1}, y_1, \ldots, y_t\}$; при этом если выполняется равенство $G(y_i, y_j) = y_m$, то $m > \max\{i, j\}$.

Используя указанные свойства отображения G, индукцией по длине сети Σ' можно показать (аналогично доказательству теоремы 3), что для правильной G-разметки η' сети Σ' выполняется свойство

$$\left(\eta'(x_i'^{(d)}) = \eta'(x_j'^{(m)})\right) \Longrightarrow (i = j)\,. \tag{2}$$

Не ограничивая общности, будем считать, что $\Sigma_t = \Sigma_1^{(1,2)}$ и соответственно выполняются равенства

$$y_t = \eta(x_1^{(t)}) = G(\eta(x_1^{(t-1)}), \quad \eta(x_2^{(t-1)})) = G(y_l, y_r).$$

Из условия леммы следует, что $y_t = \eta(x_1^{(t)}) = \eta'(x_{i_1}'^{(s)})$ и $y_r = \eta(x_2^{(t)}) = \eta'(x_{i_2}'^{(s)})$. Согласно (2), метку y_r могут иметь только вершины из множества $\{x_{i_2}'^{(0)}, \ldots, x_{i_2}'^{(s)}\}$, а метку y_t — только вершины из множества $\{x_{i_1}'^{(1)}, \ldots, x_{i_1}'^{(s)}\}$. Выберем наименьшее $l \in \{1, \ldots, s\}$ со свойством $\eta'(x_{i_1}'^{(l)}) = \ldots = \eta'(x_{i_1}'^{(s)}) = y_t$. Тогда для метки $\eta'(x_{i_1}'^{(l)}) = y_t$ в правильной G-разметке η' справедливы равенства

$$\eta'(x_{i_1}'^{(l)}) = y_t = G(y_l, y_r) = G(\eta(x_{i_1}'^{(l-1)}), \eta(x_{i_2}'^{(l-1)})).$$

Таким образом, $\eta'(x_{i_1}'^{(l)}) = \ldots = \eta'(x_{i_1}'^{(s)}) = y_t$, $\eta'(x_{i_2}'^{(l)}) = \ldots = \eta'(x_{i_2}'^{(s)}) = y_r$, при этом все вершины $x_{i_1}'^{(l)}, \ldots, x_{i_1}'^{(s-1)}$ имеют степень исхода 1, поскольку их метка не содержится в области определения правила G. Значит, согласно утверждению 1, произведение $\Sigma_1' \cdot \ldots \cdot \Sigma_s'$ эквивалентно произведению $\Sigma_1' \cdot \ldots \cdot \Sigma_{l-1}' \cdot \Sigma_{l+1}' \cdot \ldots \cdot \Sigma_s' \cdot \Sigma_l'$, и легко видеть, что для сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_{t-1}$ и $\Sigma_1' \cdot \ldots \cdot \Sigma_{l-1}' \cdot \Sigma_{l+1}' \cdot \ldots \cdot \Sigma_s'$ выполняется предположение индукции: разметка η является максимальной свободной разметкой сети $\Sigma_1 \cdot \ldots \cdot \Sigma_{t-1}$, её минимальное правило также является правилом для разметки η' сети $\Sigma_1' \cdot \ldots \cdot \Sigma_{l-1}' \cdot \Sigma_{l+1}' \cdot \ldots \cdot \Sigma_s'$ и выполняется равенство

$$\{\eta(x_1{}^{(t-1)}), \ldots, \eta(x_n{}^{(t-1)})\} = \{\eta'(x_1'^{(s-1)}), \ldots, \eta'(x_n'^{(s-1)})\}.$$

Значит, по предположению индукции $t-1 = s-1$, а произведения элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_{t-1}$ и $\Sigma_1' \cdot \ldots \cdot \Sigma_{l-1}' \cdot \Sigma_{l+1}' \cdot \ldots \cdot \Sigma_s'$ отличаются лишь допустимой перестановкой множителей. В таком случае $\Sigma_t = \Sigma_l'$. ■

Из условия $(\eta(x_1^{(t+1)}), \ldots, \eta(x_n^{(t+1)})) = (y_{i_1}, \ldots, y_{i_n}) = (\eta'(x_1'^{(s+1)}), \ldots, \eta'(x_n'^{(s+1)}))$ следует, что $\{\eta(x_1{}^{(t)}), \ldots, \eta(x_n{}^{(t)})\} = \{y_{i_1}, \ldots, y_{i_n}\} = \{\eta'(x_1'^{(s)}), \ldots, \eta'(x_n'^{(s)})\}$, и при этом минимальное правило G максимальной свободной разметки η сети $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ с начальным условием $(v_1, \ldots, v_n)$ также является правилом разметки η' сети $\Sigma_1' \cdot \ldots \cdot \Sigma_s'$ с начальным условием $(v_1, \ldots, v_n)$.

Согласно лемме 3, произведения элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ и $\Sigma_1' \cdot \ldots \cdot \Sigma_s'$ имеют одинаковый состав и могут отличаться лишь допустимой перестановкой множителей. Значит, для правильных G-разметок η и η' выполняется равенство $(\eta(x_1{}^{(t)}), \ldots, \eta(x_n{}^{(t)})) = (\eta'(x_1'^{(s)}), \ldots, \eta'(x_n'^{(s)}))$. Поскольку выполняется также равенство $(\eta(x_1^{(t+1)}), \ldots, \eta(x_n^{(t+1)})) = (\eta'(x_1'^{(s+1)}), \ldots, \eta'(x_n'^{(s+1)}))$, очевидно, что $\Pi = \Pi'$. Теорема 6 доказана. ■

Следствие 6. Пусть Σ и Σ' — биективные сети ширины n. Тогда следующие утверждения равносильны:

1) сети Σ и Σ' эквивалентны для множества Ω, $|\Omega| \geqslant \|\Sigma\| + \|\Sigma'\| + n$;
2) сети Σ и Σ' эквивалентны.

Доказательство. $1 \Rightarrow 2$. Пусть $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi$ и $\Sigma' = \Sigma'_1 \cdot \ldots \cdot \Sigma'_s \cdot \Pi'$. Тогда, как видно из доказательства теоремы 6, эквивалентности сетей Σ и Σ' для множества Ω, $|\Omega| \geqslant \|\Sigma\| + \|\Sigma'\| + n$, достаточно для того, чтобы произведения $\Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi$ и $\Sigma'_1 \cdot \ldots \cdot \Sigma'_s \cdot \Pi'$ имели одинаковый состав и отличались лишь допустимой перестановкой множителей.

$2 \Rightarrow 1$. Очевидно. ■

Теперь можно уточнить результат теоремы 2 следующим образом.

Следствие 7. Сеть Σ является биективной в том и только в том случае, когда она эквивалентна произведению

$$\Sigma_{R1} \cdot \ldots \cdot \Sigma_{Rt} \cdot \Pi_R \text{ (или } \Pi_L \cdot \Sigma_{L1} \cdot \ldots \cdot \Sigma_{Lt}),$$

где $\Sigma_{R1}, \ldots, \Sigma_{Rt}$ ($\Sigma_{L1}, \ldots, \Sigma_{Lt}$) — элементарные сети, а Π_R (Π_L) — однослойная перестановочная сеть. При этом произведение определено однозначно с точностью до возможной перестановки элементарных сетей, а количество элементарных сетей в произведении равно количеству вершин сети Σ со степенью захода 2.

Ранее мы отмечали и неоднократно пользовались тем, что две сети Σ и Σ' одинаковой ширины, которые описывают преобразование набора переменных в один и тот же набор формул, являются эквивалентными. Теперь можно доказать обратное утверждение.

Теорема 7. Эквивалентные сети Σ и Σ' описывают преобразование набора переменных в один и тот же набор формул.

Доказательство. Из теоремы 6 следует, что представления

$$\Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi \quad \text{и} \quad \Sigma'_1 \cdot \ldots \cdot \Sigma'_s \cdot \Pi'$$

эквивалентных сетей Σ и Σ' имеют одинаковый состав и произведение $\Sigma'_1 \cdot \ldots \cdot \Sigma'_s$ может быть получено из произведения $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ путём конечного числа допустимых перестановок элементарных сетей. Согласно утверждению 1, если в произведении элементарных сетей произведена допустимая перестановка множителей, то полученное произведение описывает такое же преобразование набора переменных, что и исходное произведение. ■

3. Транзитивность сетей

Продолжим развивать аппарат разметки и сформулируем на новом языке условия, достаточные для транзитивности сети.

Определение 17. Биективную сеть Σ будем называть *транзитивной для множества* Ω, если множество отображений $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ является транзитивным.

Нетрудно понять, что сама природа множества Ω в данном определении не играет никакой роли, и поэтому будет корректным говорить, что биективная сеть Σ является транзитивной для множеств мощности q. По-прежнему будем считать, что $\Omega \subset \mathbb{N}$, а для множества $\{1, \ldots, q\}$ будем использовать обозначение Ω_q.

Поскольку действие перестановочной сети не зависит от выбора квазигруппы $F \in \mathcal{Q}(\Omega)$, представляется очевидным следующее

Утверждение 2. Пусть Σ — произвольная транзитивная для множества Ω сеть, Π_1, Π_2 — произвольные перестановочные сети, для которых корректно определить произведения $\Pi_1 \cdot \Sigma$ и $\Sigma \cdot \Pi_2$. Тогда сети $\Pi_1 \cdot \Sigma$ и $\Sigma \cdot \Pi_2$ также являются транзитивными для множества Ω.

Далее, не ограничивая общности, будем считать, что произвольная биективная сеть Σ представляет собой произведение элементарных сетей $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ с множеством вершин $X_0 \cup X_1 \cup \ldots \cup X_t$.

Определение 18. Разметку μ сети Σ с условиями

$$\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_n, \mu(x_1^{(t)}) = w_1, \ldots, \mu(x_n^{(t)}) = w_n$$

будем называть *разметкой сети* Σ *с ограничениями* $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$. При этом будем говорить, что сеть Σ *допускает* разметку μ с ограничениями $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$.

Если биективная сеть Σ транзитивна для некоторого множества Ω, то для любых $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in \Omega^n$ существует такая квазигруппа $F \in \mathcal{Q}(\Omega)$, что $\Sigma^F(v_1, \ldots, v_n) = (w_1, \ldots, w_n)$. В таком случае квазигруппа F определяет правильную и непротиворечивую разметку сети Σ с ограничениями $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$. Другими словами, существование правильной непротиворечивой разметки сети Σ при произвольных ограничениях $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из множества Ω является необходимым условием для того, чтобы сеть Σ была транзитивной для множества Ω.

С другой стороны, известно, что не каждое частично определённое непротиворечивое отображение $F\colon \Omega \times \Omega \to \Omega$ может быть продолжено до квазигруппы на множестве Ω. Поэтому в общем случае условие существования правильных непротиворечивых разметок сети Σ со всеми возможными ограничениями из множества Ω^n не является гарантией транзитивности сети Σ для множества Ω. Однако связь между существованием правильных непротиворечивых разметок сети Σ и её транзитивностью существует.

Теорема 8. Пусть Σ — биективная сеть ширины n и Ω — множество мощности строго больше чем $\|\Sigma\|$. Тогда следующие утверждения эквивалентны:

1) сеть Σ является транзитивной для множества Ω;
2) сеть Σ допускает правильную непротиворечивую разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из множества Ω.

Доказательство. $1 \Rightarrow 2$. Очевидно.

$2 \Rightarrow 1$. Каждой правильной непротиворечивой разметке сети Σ с ограничениями $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из множества Ω соответствует непротиворечивое правило, определённое не более чем на $\|\Sigma\| \leqslant |\Omega| - 1$ наборах. Согласно гипотезе Эванса, данное правило продолжается до квазигруппы на множестве Ω. ■

Следствие 8. Пусть Σ — биективная сеть ширины n и Ω — множество мощности не менее чем $\|\Sigma\| + n$. Тогда следующие утверждения эквивалентны:

1) сеть Σ является транзитивной для множества Ω;

2) сеть Σ допускает правильную непротиворечивую разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из множества Ω;
3) сеть Σ допускает правильную непротиворечивую разметку при любых ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из множества $\mathbb{N}$;
4) сеть Σ является транзитивной для любого множества, мощность которого не менее чем $\|\Sigma\| + n$.

Замечание 4. В общем случае утверждение теоремы 8 нельзя усилить, поскольку гипотеза Эванса также не допускает усиления оценки в общем случае.

Пусть η — произвольная разметка сети $\Sigma = \Sigma_1 \cdot \ldots \cdot \Sigma_t$. Тогда с разметкой η свяжем отношение $G \subset \mathbb{N}^3$, определённое следующим образом: отношение G содержит тройку (y_l, y_r, y_q) в том и только в том случае, когда для некоторого $s \in \{1, \dots, t\}$ выполняются равенства $\Sigma_s = \Sigma_m^{(i,j)}$ и $(\eta(x_i^{(s-1)}), \eta(x_j^{(s-1)}), \eta(x_m^{(s)})) = (y_l, y_r, y_q)$.

Если в отношении G содержатся две тройки, отличающиеся только в одной координате, например (y_l, y_r, y_q) и $(y_l, y_r, y_{q'})$, то, заменив в разметке η все метки $y_{q'}$ на y_q, получим разметку η_1, в которой используется на одну метку меньше, чем в разметке η. Если в отношении G_1, соответствующем разметке η_1, присутствуют две тройки, отличающиеся только в одной координате, то повторим эти действия, и так далее.

Таким образом построим последовательность разметок $\eta = \eta_0, \eta_1, \dots$ сети Σ, в которой каждая следующая разметка использует на одну метку меньше, чем предыдущая. Поэтому указанная последовательность разметок оборвётся на некотором конечном шаге, например с номером k, в том смысле, что в отношении G_k, соответствующем разметке η_k, не найдётся двух троек, отличающихся только в одной координате. Такая разметка η_k будет правильной и непротиворечивой разметкой сети Σ.

Описанную процедуру будем называть *устранением противоречий в разметке* η. При этом будем говорить, что разметка η_d, $d \in \{0, 1, \dots, k\}$, *получена из разметки* η *устранением противоречий.*

Лемма 4. Пусть η — произвольная разметка сети Σ, μ — правильная непротиворечивая разметка сети Σ, и при этом существует отображение $\sigma_\mu \colon \eta \to \mu$. Тогда для любой разметки $\widetilde{\eta}$, полученной из разметки η устранением противоречий, выполняется условие $\sigma_\mu \colon \widetilde{\eta} \to \mu$.

Доказательство. Пусть $\eta = \eta_0, \eta_1, \dots, \eta_k = \widetilde{\eta}$ — последовательность разметок сети Σ, полученная в результате последовательного устранения противоречий в разметке η. Для доказательства утверждения методом математической индукции достаточно показать, что для разметки η_1 также выполняется условие $\sigma_\mu \colon \eta_1 \to \mu$.

При построении разметки η_1 в отношении G, соответствующем разметке η, выбираются две тройки, отличающиеся только в одной координате. Рассмотрим все возможные случаи:

1) Если выбранная пара имеет вид (y_l, y_r, y_q) и $(y_l, y_r, y_{q'})$, то $\sigma_\mu(y_q) = \sigma_\mu(y_{q'})$, поскольку разметка μ является правильной. Значит, для разметки η_1, полученной из разметки η заменой всех меток $y_{q'}$ на y_q, также выполняется условие $\sigma_\mu \colon \eta_1 \to \mu$.
2) Если выбранная пара имеет вид (y_l, y_r, y_q) и $(y_l, y_{r'}, y_q)$, то $\sigma_\mu(y_r) = \sigma_\mu(y_{r'})$, поскольку разметка μ является непротиворечивой. Тогда при замене в разметке η всех меток $y_{r'}$ на y_r получается разметка η_1, для которой, очевидно, выполняется условие $\sigma_\mu \colon \eta_1 \to \mu$.

3) Если выбранная пара имеет вид (y_l, y_r, y_q) и $(y_{l'}, y_r, y_q)$, то $\sigma_\mu(y_l) = \sigma_\mu(y_{l'})$, поскольку разметка μ является непротиворечивой. Тогда при замене в разметке η всех меток $y_{l'}$ на y_l получается разметка η_1, для которой, очевидно, выполняется условие $\sigma_\mu\colon \eta_1 \to \mu$.

Лемма доказана. ■

Следствие 9. Пусть η — произвольная разметка сети Σ. Тогда правильная непротиворечивая разметка $\widetilde{\eta}$, полученная из разметки η устранением противоречий, определена однозначно с точностью до обратимого переобозначения меток.

Доказательство. Пусть $\widetilde{\eta}$ и $\widehat{\eta}$ — две правильные непротиворечивые разметки, полученные из разметки η устранением противоречий. Тогда существуют отображения $\sigma_{\widetilde{\eta}}$ и $\sigma_{\widehat{\eta}}$, удовлетворяющие условиям $\sigma_{\widetilde{\eta}}\colon \eta \to \widetilde{\eta}$ и $\sigma_{\widehat{\eta}}\colon \eta \to \widehat{\eta}$. Согласно лемме 4, отображения $\sigma_{\widetilde{\eta}}$ и $\sigma_{\widehat{\eta}}$ также удовлетворяют условиям $\sigma_{\widetilde{\eta}}\colon \widehat{\eta} \to \widetilde{\eta}$ и $\sigma_{\widehat{\eta}}\colon \widetilde{\eta} \to \widehat{\eta}$, что и доказывает утверждение следствия. ■

Определение 19. Правильную непротиворечивую разметку η сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ будем называть *свободной разметкой сети Σ с ограничениями*, если для любой правильной непротиворечивой разметки μ сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ существует отображение $\sigma_\mu\colon \eta \to \mu$.

Из определения следует, что, при условии существования, свободная разметка сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ определена однозначно с точностью до обратимого переобозначения меток.

Пример 2. Рассмотрим биективную сеть Σ ширины 2

$$\Sigma = \Sigma_1^{(2,1)} \cdot \Sigma_1^{(1,2)} \cdot \Sigma_2^{(1,2)}.$$

Частично определённое правило F

$$F(1,2) = 1,\ F(2,1) = 3,\ F(3,2) = 1$$

определяет правильную, но противоречивую разметку μ сети Σ с ограничениями $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$. При этом свободная разметка η сети Σ с теми же ограничениями $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ определяется правилом G:

$$G(1,2) = 1,\ G(2,1) = 1.$$

Пример показывает, что в определении свободной разметки с ограничениями нельзя отказаться от условия непротиворечивости правильной разметки μ. Другими словами, свободная разметка с ограничениями не является «свободной» в классе всех правильных разметок с теми же ограничениями.

Теорема 9. Если сеть Σ допускает правильную непротиворечивую разметку с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$, то существует свободная разметка сети Σ с указанными ограничениями.

Доказательство. Для удобства дальнейшего изложения будем считать, что $\mathbb{N}\backslash\{v_1, \dots, v_n, w_1, \dots, w_n\} = \{y_1, y_2, \dots\}$. Пусть η_0 — свободная разметка сети Σ с начальным условием $(v_1, \dots, v_n)$, полученная в результате свободного продолжения начальной разметки $\eta_0(x_1^{(0)}) = v_1, \dots, \eta_0(x_n^{(0)}) = v_n$ с использованием меток $y_1, y_2, \dots$

Тогда, согласно теореме 3, для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ существует отображение σ_μ, удовлетворяющее условию $\sigma_\mu\colon \eta_0 \to \mu$. Продолжим указанное отображение σ_μ по правилу $\sigma_\mu(w_i) = w_i$, $i \in \{1, \dots, n\}$, и заменим в разметке η_0 метки $\eta_0(x_1^{(t)}), \dots, \eta_0(x_n^{(t)})$ на $w_1, \dots, w_n$ соответственно. Таким образом, мы построили разметку η_1 сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ и при этом для любой правильной непротиворечивой разметки μ с этими ограничениями существует отображение σ_μ, удовлетворяющее условию $\sigma_\mu\colon \eta_1 \to \mu$.

Проведём процедуру устранения противоречий в разметке η_1 с двумя уточнениями:

— если при устранении противоречия требуется отождествить метки v_i и y_j, то будем заменять метку y_j на v_i;

— если при устранении противоречия требуется отождествить метки w_i и y_j, то будем заменять метку y_j на w_i.

Пусть $\widetilde{\eta}$ — правильная непротиворечивая разметка сети Σ, полученная из разметки η_1 устранением противоречий. Тогда, согласно уточнениям, все метки $\widetilde{\eta}(x_1^{(0)})$, …, $\widetilde{\eta}(x_n^{(0)})$, $\widetilde{\eta}(x_1^{(t)})$, …, $\widetilde{\eta}(x_n^{(t)})$ содержатся в множестве $\{v_1, \dots, v_n, w_1, \dots, w_n\}$. При этом, согласно лемме 4, для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ соответствующее отображение σ_μ удовлетворяет условию $\sigma_\mu\colon \widetilde{\eta} \to \mu$. Значит, на самом деле разметка $\widetilde{\eta}$ является свободной разметкой сети Σ с указанными ограничениями. ■

Следствие 10. Если биективная сеть Σ ширины n является транзитивной для множества Ω_q при $q \geqslant 2n$, то при любых $v_1, \dots, v_n, w_1, \dots, w_n \in \mathbb{N}$ существует свободная разметка сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$.

Следующая теорема фактически оправдывает название свободной разметки с ограничениями.

Теорема 10. Пусть η — свободная разметка сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$, μ — правильная непротиворечивая разметка сети Σ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$. Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu\colon \eta \to \mu$.

Доказательство. Для удобства дальнейшего изложения будем считать, что $\mathbb{N}\backslash\{v_1, \dots, v_n, w_1, \dots, w_n\} = \{y_1, y_2, \dots\}$. Пусть η_0 — свободная разметка сети Σ с начальным условием $(v_1, \dots, v_n)$, полученная в результате свободного продолжения начальной разметки $\eta_0(x_1^{(0)}) = v_1, \dots, \eta_0(x_n^{(0)}) = v_n$ с использованием меток $y_1, y_2, \dots$ Тогда, согласно теореме 4, для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$, указанное отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu\colon \eta_0 \to \mu$. Заменив в разметке η_0 метки $\eta_0(x_1^{(t)}), \dots, \eta_0(x_n^{(t)})$ на $w_1, \dots, w_n$ соответственно, получим разметку η_1 сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$, при этом для любой правильной непротиворечивой размет-

ки μ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu \colon \eta_1 \to \mu$.

Проведём процедуру устранения противоречий в разметке η_1 с двумя уточнениями:

— если при устранении противоречия требуется отождествить метки v_i и y_j, то будем заменять метку y_j на v_i;

— если при устранении противоречия требуется отождествить метки w_i и y_j, то будем заменять метку y_j на w_i.

В доказательстве теоремы 9 показано, что правильная непротиворечивая разметка сети Σ, полученная из разметки η_1 устранением противоречий, является свободной разметкой сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$. Без потери общности можно считать, что при устранении противоречий в разметке η_1 получается свободная разметка η. Поскольку для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$, отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu \colon \eta_1 \to \mu$, то, согласно лемме 4, данное продолжение также удовлетворяет условию $\sigma_\mu \colon \eta \to \mu$. ■

Следствие 11. В условиях теоремы 10, если G и F — минимальные правила разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$.

Достаточным условием для существования правильных непротиворечивых разметок сети Σ при всех возможных ограничениях из $\mathbb{N}$ является существование правильных непротиворечивых разметок сети Σ при всех возможных ограничениях из Ω_{2n}. Справедливо и более сильное утверждение.

Теорема 11. Сеть Σ допускает правильные непротиворечивые разметки при всех возможных ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из $\mathbb{N}$ в том и только в том случае, когда сеть Σ допускает правильные непротиворечивые разметки при всех возможных ограничениях $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из Ω_2.

Доказательство. Необходимость очевидна, докажем достаточность. Пусть $\mathbb{N} \backslash \{v_1, \dots, v_n, w_1, \dots, w_n\} = \{y_1, y_2, \dots\}$ и η_0 — свободная разметка сети Σ с начальным условием $(v_1, \dots, v_n)$, полученная в результате свободного продолжения начальной разметки $\eta_0(x_1^{(0)}) = v_1, \dots, \eta_0(x_n^{(0)}) = v_n$ с использованием меток $y_1, y_2, \dots$ Тогда, согласно теореме 4, для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из Ω_2, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu \colon \eta_0 \to \mu$. Заменив в разметке η_0 метки $\eta_0(x_1^{(t)}), \dots, \eta_0(x_n^{(t)})$ на $w_1, \dots, w_n$ соответственно, получим разметку η_1 сети Σ с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$, при этом для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из Ω_2, при которых возможно определить

отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu \colon \eta_1 \to \mu$.

Проведём процедуру устранения противоречий в разметке η_1 с двумя уточнениями:

— если при устранении противоречия требуется отождествить метки v_i и y_j, то будем заменять метку y_j на v_i;
— если при устранении противоречия требуется отождествить метки w_i и y_j, то будем заменять метку y_j на w_i.

Пусть $\widetilde{\eta}$ — правильная непротиворечивая разметка сети Σ, полученная из разметки η_1 устранением противоречий. Тогда, согласно уточнениям, все метки $\widetilde{\eta}(x_1^{(0)}), \dots, \widetilde{\eta}(x_n^{(0)})$, $\widetilde{\eta}(x_1^{(t)}), \dots, \widetilde{\eta}(x_n^{(t)})$ содержатся в множестве $\{v_1, \dots, v_n, w_1, \dots, w_n\}$. При этом, согласно лемме 4, для любой правильной непротиворечивой разметки μ с ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из Ω_2, при которых возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$, $i \in \{1, \dots, n\}$, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu \colon \widetilde{\eta} \to \mu$.

Методом от противного покажем, что правильная непротиворечивая разметка $\widetilde{\eta}$ будет разметкой с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$. Согласно сделанным уточнениям, в ограничениях разметки $\widetilde{\eta}$ могли появиться противоречия только следующих типов:

— $\widetilde{\eta}(x_i^{(0)}) = v_j \neq v_i$;
— $\widetilde{\eta}(x_i^{(0)}) = w_j \neq v_i$;
— $\widetilde{\eta}(x_i^{(t)}) = v_j \neq w_i$;
— $\widetilde{\eta}(x_i^{(t)}) = w_j \neq w_i$.

Разберём первый случай: $\widetilde{\eta}(x_i^{(0)}) = v_j \neq v_i$. Согласно условию теоремы, существует правильная непротиворечивая разметка μ с ограничениями $(\delta_{v_1,v_j} + 1, \dots, \delta_{v_n,v_j} + 1)$, $(\delta_{w_1,v_j} + 1, \dots, \delta_{w_n,v_j} + 1) \in \Omega_2^n$ и при этом возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \delta_{v_i,v_j} + 1$, $\sigma_\mu(w_i) = \delta_{w_i,v_j} + 1$, $i \in \{1, \dots, n\}$. Значит, отображение σ_μ продолжается таким образом, что удовлетворяет условию $\sigma_\mu \colon \widetilde{\eta} \to \mu$. Получили противоречие, поскольку

$$\sigma_\mu(\widetilde{\eta}(x_i^{(0)})) = \sigma_\mu(v_j) = 2 \neq 1 = \delta_{v_i,v_j} + 1 = \mu(x_i^{(0)}).$$

Отсутствие противоречий остальных типов устанавливается аналогичным образом. ■

Ввиду следствия 8 становится очевидным следующее утверждение.

Следствие 12. Пусть Σ — биективная сеть ширины n и Ω — множество мощности не менее чем $\|\Sigma\| + n$. Тогда следующие утверждения эквивалентны:

1) сеть Σ является транзитивной для множества Ω;
2) сеть Σ допускает правильную непротиворечивую разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из множества Ω;
3) сеть Σ допускает правильную непротиворечивую разметку элементами множества Ω при любых ограничениях $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из множества $\Omega_2 \subset \Omega$;
4) множество преобразований $\{\Sigma^F : F \in \mathcal{Q}(\Omega)\}$ действует транзитивным образом на подмножестве $\Omega_2^n \subset \Omega^n$.

Пример 3. На первый взгляд может показаться, что результат теоремы 11 можно усилить следующим образом: если сеть Σ допускает правильную непротиворечивую

разметку при всех возможных ограничениях $\begin{pmatrix} 1 & \dots & 1 \\ w_1 & \dots & w_n \end{pmatrix}$ из Ω_2, то она допускает правильную непротиворечивую разметку с любыми ограничениями $\begin{pmatrix} \bar{v}_1 & \dots & \bar{v}_n \\ \bar{w}_1 & \dots & \bar{w}_n \end{pmatrix}$ из Ω_2 и, как следствие, допускает правильную непротиворечивую разметку с любыми ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из $\mathbb{N}$. Однако такое предположение является неверным.

Рассмотрим биективную сеть $\Sigma = \Sigma_2^{(1,2)} \cdot \Sigma_2^{(2,1)} \cdot \Sigma_1^{(1,2)}$ ширины 2. Нетрудно проверить, что сеть Σ допускает правильные непротиворечивые разметки при всех возможных ограничениях $\begin{pmatrix} 1 & 1 \\ w_1 & w_2 \end{pmatrix}$ из Ω_2, но при этом не допускает правильной непротиворечивой разметки с ограничениями $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ (рис. 3).

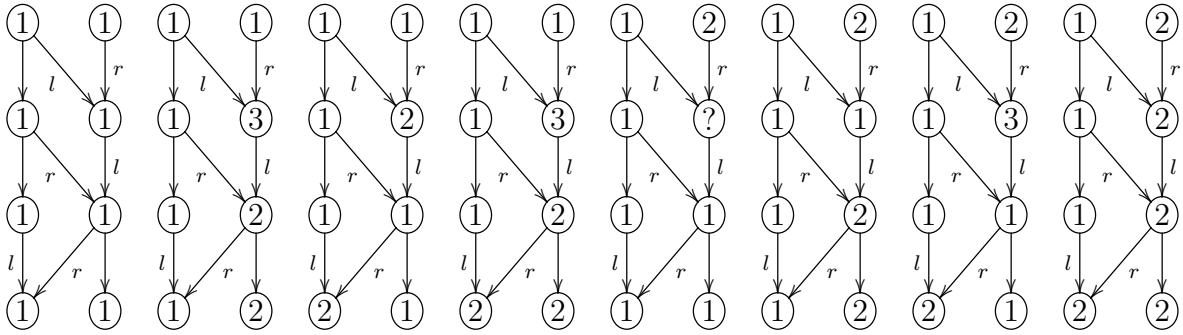

Рис. 3

На самом деле результат теоремы 11 можно усилить, воспользовавшись очевидным соображением: если сеть Σ допускает правильную непротиворечивую разметку с ограничениями $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из Ω_2 и π — подстановка на множестве Ω_2, то сеть Σ допускает правильную непротиворечивую разметку с ограничениями $\begin{pmatrix} \pi(v_1) & \dots & \pi(v_n) \\ \pi(w_1) & \dots & \pi(w_n) \end{pmatrix}$. Другими словами, для проверки того, что сеть Σ допускает правильную непротиворечивую разметку при любых ограничениях из Ω_2, достаточно убедиться в существовании 2^{2n-1} правильных непротиворечивых разметок сети Σ с ограничениями из Ω_2. Данное усиление результата теоремы 11 примечательно тем, что оно не улучшаемо в общем случае — сеть Σ допускает правильную непротиворечивую разметку при любых ограничениях $\begin{pmatrix} v_1 & v_2 \\ w_1 & w_2 \end{pmatrix}$ из Ω_2, за исключением $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ и $\begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}$.

В дальнейшем нам потребуются естественные обобщения понятия свободной разметки с ограничениями и соответствующих результатов.

Определение 20. Правильную непротиворечивую разметку η сети Σ будем называть *свободной разметкой сети Σ с условиями*

$$\eta(x_1^{(0)}) = v_1,\ \dots,\ \eta(x_n^{(0)}) = v_n,\ \eta(x_{i_1}^{(t)}) = w_{i_1},\ \dots,\ \eta(x_{i_k}^{(t)}) = w_{i_k},$$

если для любой правильной непротиворечивой разметки μ сети Σ с аналогичными условиями $\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_n, \mu(x_{i_1}^{(t)}) = w_{i_1}, \ldots, \mu(x_{i_k}^{(t)}) = w_{i_k}$ существует такое отображение σ_μ, что $\sigma_\mu \colon \eta \to \mu$.

Аналогично теоремам 9 и 10 доказываются следующие утверждения.

Теорема 12. Если существует правильная непротиворечивая разметка μ сети Σ с условиями $\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_n, \mu(x_{i_1}^{(t)}) = w_{i_1}, \ldots, \mu(x_{i_k}^{(t)}) = w_{i_k}$, то существует единственная, с точностью до переобозначений, свободная разметка η сети Σ с теми же условиями $\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_{i_1}^{(t)}) = w_{i_1}, \ldots, \eta(x_{i_k}^{(t)}) = w_{i_k}$.

Теорема 13. Пусть для свободной разметки η с условиями

$$\eta(x_1^{(0)}) = v_1, \ \ldots, \ \eta(x_n^{(0)}) = v_n, \ \eta(x_{i_1}^{(t)}) = w_{i_1}, \ \ldots, \ \eta(x_{i_k}^{(t)}) = w_{i_k}$$

и правильной непротиворечивой разметки μ с условиями

$$\mu(x_1^{(0)}) = \bar{v}_1, \ \ldots, \ \mu(x_n^{(0)}) = \bar{v}_n, \ \mu(x_{i_1}^{(t)}) = \bar{w}_{i_1}, \ \ldots, \ \mu(x_{i_k}^{(t)}) = \bar{w}_{i_k}$$

возможно определить отображение σ_μ по правилу $\sigma_\mu(v_i) = \bar{v}_i$, $\sigma_\mu(w_i) = \bar{w}_i$. Тогда отображение σ_μ допускает продолжение, удовлетворяющее условию $\sigma_\mu \colon \eta \to \mu$.

Следствие 13. В условиях теоремы 13, если G и F — минимальные правила разметок η и μ соответственно, то при всех допустимых $z_i, z_j \in \mathbb{N}$ выполняется равенство $\sigma_\mu(G(z_i, z_j)) = F(\sigma_\mu(z_i), \sigma_\mu(z_j))$. В частности, если метка $\mu(x_i^{(s)})$ не содержится в области определения F, то метка $\eta(x_i^{(s)})$ не содержится в области определения G.

4. Построение транзитивных сетей

Здесь нам потребуется продолжение результата теоремы 2 о представлении биективной сети в виде произведения элементарных и перестановочной сетей.

Утверждение 3. Произвольная биективная сеть Σ эквивалентна произведению

$$\Pi_L \left(\Sigma_{11} \cdot \ldots \cdot \Sigma_{1k_1}\right) \cdot \ldots \cdot \left(\Sigma_{n1} \cdot \ldots \cdot \Sigma_{nk_n}\right) \Pi_R,$$

в котором Π_L, Π_R — перестановочные сети, а $\Sigma_{11}, \ldots, \Sigma_{1k_1}, \ldots, \Sigma_{n1}, \ldots, \Sigma_{nk_n}$ — элементарные сети, удовлетворяющие следующему условию: при всех $s \in \{2, \ldots, n\}$, $r \in \{1, \ldots, k_s\}$ вершины $x_1^{(k_1+\ldots+k_{s-1}+r)}, \ldots, x_{s-1}^{(k_1+\ldots+k_{s-1}+r)}$ сети Σ_{sr} имеют степень захода 1.

Доказательство. Согласно теореме 2, биективная сеть Σ эквивалентна произведению

$$\Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi_R,$$

где $\Sigma_1, \ldots, \Sigma_t$ — элементарные сети, а Π_R — перестановочная сеть. Пусть $i_1, \ldots, i_n$ — такая последовательность номеров от 1 до n, что при всех $s \in \{2, \ldots, n\}$ и $l \in \{1, \ldots, t\}$ для вершин сети $\Sigma_1 \cdot \ldots \cdot \Sigma_t$ выполняется условие

$$(\deg^- x_{i_s}^{(l)} = \ldots = \deg^- x_{i_s}^{(t)} = 1) \implies (\deg^- x_{i_{s-1}}^{(l)} = \ldots = \deg^- x_{i_{s-1}}^{(t)} = 1).$$

Другими словами, $i_1, \ldots, i_n$ — порядок «окончания преобразований» вершин произведения $\Sigma_1 \cdot \ldots \cdot \Sigma_t$. Нетрудно понять, что при выборе перестановочной сети Π_L, соответствующей подстановке $\begin{pmatrix} i_1 & \ldots & i_n \\ 1 & \ldots & n \end{pmatrix}$, произведение $\Pi_L^{-1} \cdot \Sigma_1 \cdot \ldots \cdot \Sigma_t \cdot \Pi_L$ эквивалентно произведению элементарных сетей

$$\left(\Sigma_{11} \cdot \ldots \cdot \Sigma_{1k_1}\right) \cdot \ldots \cdot \left(\Sigma_{n1} \cdot \ldots \cdot \Sigma_{nk_n}\right),$$

где $\Sigma_{11},\ldots,\Sigma_{1k_1},\ldots,\Sigma_{n1},\ldots,\Sigma_{nk_n}$ — элементарные сети, удовлетворяющие следующему условию: при всех $s\in\{2,\ldots,n\}$, $r\in\{1,\ldots,k_s\}$ вершины $x_1^{(k_1+\ldots+k_{s-1}+r)}$, $\ldots$, $x_{s-1}^{(k_1+\ldots+k_{s-1}+r)}$ сети Σ_{sr} имеют степень захода 1.

Утверждение можно считать доказанным, поскольку исходная сеть Σ эквивалентна произведению $\Pi_L\,(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{n1}\cdot\ldots\cdot\Sigma_{nk_n})\,(\Pi_L^{-1}\Pi_R)$. ■

Ввиду следствия 7 корректно дать следующее

Определение 21. Указанное в утверждении 3 произведение

$$\Pi_L\,(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{n1}\cdot\ldots\cdot\Sigma_{nk_n})\,\Pi_R$$

будем называть *каноническим представлением* биективной сети Σ. При этом сети $\Sigma_{s1}\cdot\ldots\cdot\Sigma_{sk_s}$, $s\in\{1,\ldots,n\}$, будем называть *слоями* канонического представления биективной сети Σ. Слой $\Sigma_{s1}\cdot\ldots\cdot\Sigma_{sk_s}$ называется *вырожденным*, если $k_s=0$.

Замечание 5. В каноническом представлении произвольной биективной сети вырожденные слои присутствуют в том и только в том случае, когда для некоторого s все вершины $x_s^{(i)}$, $i\geqslant 0$, имеют степень захода 1. При этом пустыми могут быть только первые несколько подряд идущих слоев.

Не ограничивая общности, всюду далее будем считать, что произвольная биективная сеть Σ равна своему каноническому представлению

$$(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{n1}\cdot\ldots\cdot\Sigma_{nk_n})$$

с множеством вершин $X_0\cup X_{11}\cup\ldots\cup X_{1k_1}\cup\ldots\cup X_{n1}\cup\ldots\cup X_{nk_n}$.

Далее предложен алгоритм модификации канонического представления произвольной биективной сети, в результате работы которого получается сеть, действующая транзитивным образом для всех достаточно больших множеств. Предварительно докажем вспомогательное утверждение о свойстве процедуры свободного продолжения разметки.

Лемма 5. Пусть $\Sigma=(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{n1}\cdot\ldots\cdot\Sigma_{nk_n})$ и μ — разметка сети

$$\Sigma_{s-1}=(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{(s-1)1}\cdot\ldots\cdot\Sigma_{(s-1)k_{s-1}})$$

с минимальным правилом $F_{\Sigma_{s-1}}$. Тогда если среди $x_s^{(k_1+\ldots+k_{s-1})},\ldots,x_n^{(k_1+\ldots+k_{s-1})}$ найдётся вершина $x_i^{(k_1+\ldots+k_{s-1})}$, метка которой не содержится в области определения $F_{\Sigma_{s-1}}$, то при свободном продолжении разметки μ до разметки сети

$$\Sigma_s=(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{(s-1)1}\cdot\ldots\cdot\Sigma_{(s-1)k_{s-1}})\,(\Sigma_{s1}\cdot\ldots\cdot\Sigma_{sk_s})$$

среди $x_s^{(k_1+\ldots+k_s)},\ldots,x_n^{(k_1+\ldots+k_s)}$ найдётся вершина $x_j^{(k_1+\ldots+k_s)}$, метка которой не содержится в области определения F_{Σ_s} — минимального правила разметки μ сети Σ_s.

Доказательство. Докажем утверждение индукцией по длине слоя $(\Sigma_{s1}\cdot\ldots\cdot\Sigma_{sk_s})$. База при $k_s=0$ очевидна.

Пусть теперь $k_s\geqslant 1$ и $\Sigma_{sk_s}=\Sigma_l^{\{l,r\}}$, $l\in\{s,\ldots,n\}$. Тогда по предположению индукции среди вершин $x_s^{(k_1+\ldots+k_s-1)},\ldots,x_n^{(k_1+\ldots+k_s-1)}$ найдётся вершина $x_j^{(k_1+\ldots+k_s-1)}$, метка которой не содержится в области определения $F_{\Sigma'_s}$ — минимального правила разметки μ сети $\Sigma'_s=(\Sigma_{11}\cdot\ldots\cdot\Sigma_{1k_1})\cdot\ldots\cdot(\Sigma_{(s-1)1}\cdot\ldots\cdot\Sigma_{(s-1)k_{s-1}})\cdot(\Sigma_{s1}\cdot\ldots\cdot\Sigma_{s(k_s-1)})$. Рассмотрим два возможных случая:

1) Если одна из вершин $x_l^{(k_1+\ldots+k_s-1)}$ или $x_r^{(k_1+\ldots+k_s-1)}$ имеет метку, не содержащуюся в области определения $F_{\Sigma'_s}$, то, согласно определению процедуры свободного продолжения разметки, вершина $x_l^{(k_1+\ldots+k_s)}$ будет иметь метку, не содержащуюся в области определения F_{Σ_s}.
2) Если метки обеих вершин $x_l^{(k_1+\ldots+k_s-1)}$ и $x_r^{(k_1+\ldots+k_s-1)}$ содержатся в области определения $F_{\Sigma'_s}$, то очевидно, что метка вершины $x_j^{(k_1+\ldots+k_s-1)}$ отличается от меток вершин $x_l^{(k_1+\ldots+k_s-1)}$ и $x_r^{(k_1+\ldots+k_s-1)}$. Значит, вершина $x_j^{(k_1+\ldots+k_s-1)}$ и соответственно вершина $x_j^{(k_1+\ldots+k_s)}$ будут иметь метку, не содержащуюся в области определения F_{Σ_s}.

Лемма доказана. ■

Алгоритм построения транзитивной сети

Вход: произвольная биективная сеть $\Sigma = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1k_1}) \cdot \ldots \cdot (\Sigma_{n1} \cdot \ldots \cdot \Sigma_{nk_n})$.

Шаг $s \leqslant n-2$**.** Пусть первые $(s-1)$ слоев канонического представления сети Σ уже модифицированы таким образом, что сеть

$$\widehat{\Sigma}_{s-1} = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(s-1)1} \cdot \ldots \cdot \Sigma_{(s-1)\widehat{k}_{s-1}})$$

допускает свободную разметку η при любых условиях

$$\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_1^{(\widehat{k}_1)}) = w_1, \ldots, \eta(x_{s-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}) = w_{s-1} \in \Omega_2$$

и при этом среди $x_s^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}, \ldots, x_n^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}$ существует вершина $x_i^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}$, метка которой $\eta(x_i^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})})$, независимо от условий $v_1, \ldots, v_n, w_1, \ldots, w_{s-1} \in \Omega_2$, не содержится в области определения $G_{\widehat{\Sigma}_{s-1}}$ — минимального правила разметки η сети $\widehat{\Sigma}_{s-1}$.

Пусть μ — свободная разметка сети $\widehat{\Sigma}_{s-1}$ с условиями

$$\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \ldots, \mu(x_{s-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}) = v_1 \in \Omega_2.$$

Тогда, согласно сделанному предположению, метка $\mu(x_i^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})})$ не содержится в области определения минимального правила $F_{\widehat{\Sigma}_{s-1}}$. Продолжим разметку μ сети $\widehat{\Sigma}_{s-1}$ свободным образом до разметки сети

$$\widehat{\Sigma}'_s = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(s-1)1} \cdot \ldots \cdot \Sigma_{(s-1)\widehat{k}_{s-1}}) \cdot (\Sigma_{s1} \cdot \ldots \cdot \Sigma_{sk_s}).$$

Согласно лемме 5, среди вершин $x_s^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1}+k_s)}, \ldots, x_n^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1}+k_s)}$ существует такая вершина $x_j^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1}+k_s)}$, что метка $\mu(x_j^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1}+k_s)})$ не содержится в области определения $F_{\widehat{\Sigma}'_s}$ — минимального правила разметки μ сети $\widehat{\Sigma}'_s$. Поскольку разметка μ по построению является свободной разметкой сети $\widehat{\Sigma}'_s$ с условиями

$$\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \ldots, \mu(x_{s-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}) = v_1 \in \Omega_2,$$

то, согласно следствию 13, для любой свободной разметки η сети $\widehat{\Sigma}'_s$ с условиями

$$\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_1^{(\widehat{k}_1)}) = w_1, \ldots, \eta(x_{s-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}) = w_{s-1} \in \Omega_2,$$

независимо от условий $v_1, \ldots, v_n, w_1, \ldots, w_{s-1} \in \Omega_2$, метка $\eta(x_j^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1}+k_s)})$ не содержится в области определения $G_{\widehat{\Sigma}'_s}$ — минимального правила разметки η сети $\widehat{\Sigma}'_s$.

Рассмотрим два возможных варианта модификации s-го слоя $\Sigma_{s1} \cdot \ldots \cdot \Sigma_{sk_s}$:

1) Если $j = s$, то выберем произвольные $l, m \in \{s+1, \ldots, n\}$, $l \neq m$, и модифицируем s-й слой $\Sigma_{s1} \cdot \ldots \cdot \Sigma_{sk_s}$ следующим образом:

$$\Sigma_{s1} \cdot \ldots \cdot \Sigma_{s\widehat{k}_s} = \Sigma_{s1} \cdot \ldots \cdot \Sigma_{sk_s} \cdot \Sigma_l^{(s,l)} \cdot \Sigma_s^{(s,l)} \cdot \Sigma_m^{(l,m)}.$$

2) Если $j \neq s$, то выберем произвольный $m \in \{s+1, \ldots, n\}$, $m \neq j$, и модифицируем s-й слой $\Sigma_{s1} \cdot \ldots \cdot \Sigma_{sk_s}$ следующим образом:

$$\Sigma_{s1} \cdot \ldots \cdot \Sigma_{s\widehat{k}_s} = \Sigma_{s1} \cdot \ldots \cdot \Sigma_{sk_s} \cdot \Sigma_s^{(s,j)} \cdot \Sigma_m^{(s,m)} \cdot \Sigma_s^{(j,s)}.$$

В каждом из этих случаев свободная разметка η сети $\widehat{\Sigma}'_s$ с произвольными условиями $\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_1^{(\widehat{k}_1)}) = w_1, \ldots, \eta(x_{s-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{s-1})}) = w_{s-1} \in \Omega_2$ продолжается до свободной разметки η сети

$$\widehat{\Sigma}_s = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(s-1)1} \cdot \ldots \cdot \Sigma_{(s-1)\widehat{k}_{s-1}})(\Sigma_{s1} \cdot \ldots \cdot \Sigma_{s\widehat{k}_s})$$

с любым условием $\eta(x_s^{(\widehat{k}_1+\ldots+\widehat{k}_s)}) = w_s \in \Omega_2$, при этом метка $\eta(x_m^{(\widehat{k}_1+\ldots+\widehat{k}_s)})$, независимо от выбора условий $v_1, \ldots, v_n, w_1, \ldots, w_{s-1}, w_s \in \Omega_2$, не содержится в области определения $G_{\widehat{\Sigma}_s}$ — минимального правила разметки η сети $\widehat{\Sigma}_s$.

Шаг $n-1$**.** Пусть первые $(n-2)$ слоев канонического представления сети Σ уже модифицированы таким образом, что сеть

$$\widehat{\Sigma}_{n-2} = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(n-2)1} \cdot \ldots \cdot \Sigma_{(n-2)\widehat{k}_{n-2}})$$

допускает свободную разметку η при любых условиях

$$\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_1^{(\widehat{k}_1)}) = w_1, \ldots, \eta(x_{n-2}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}) = w_{n-2} \in \Omega_2$$

и при этом среди $x_{n-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}, x_n^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}$ существует вершина $x_i^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}$, метка которой $\eta(x_i^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})})$, независимо от условий $v_1, \ldots, v_n, w_1, \ldots, w_{n-2} \in \Omega_2$, не содержится в области определения $G_{\widehat{\Sigma}_{n-2}}$ — минимального правила разметки η сети $\widehat{\Sigma}_{n-2}$.

Пусть μ — свободная разметка сети $\widehat{\Sigma}_{n-2}$ с условиями

$$\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \ldots, \mu(x_{n-2}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}) = v_1 \in \Omega_2.$$

Тогда, согласно сделанному предположению, метка $\mu(x_i^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})})$ не содержится в области определения минимального правила $F_{\widehat{\Sigma}_{n-2}}$. Продолжим разметку μ сети $\widehat{\Sigma}_{n-2}$ свободным образом до разметки сети

$$\widehat{\Sigma}'_{n-1} = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(n-2)1} \cdot \ldots \cdot \Sigma_{(n-2)\widehat{k}_{n-2}}) \cdot (\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)k_{n-1}}).$$

Согласно лемме 5, среди вершин $x_{n-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2}+k_{n-1})}, x_n^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2}+k_{n-1})}$ существует такая вершина $x_j^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2}+k_{n-1})}$, что метка $\mu(x_j^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2}+k_{n-1})})$ не содержится в области определения минимального правила $F_{\widehat{\Sigma}'_{n-1}}$. Поскольку разметка μ по построению является свободной разметкой сети $\widehat{\Sigma}'_{n-1}$ с условиями

$$\mu(x_1^{(0)}) = v_1, \ldots, \mu(x_n^{(0)}) = v_1, \mu(x_1^{(\widehat{k}_1)}) = v_1, \ldots, \mu(x_{n-2}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}) = v_1 \in \Omega_2,$$

то, согласно следствию 13, для любой свободной разметки η сети $\widehat{\Sigma}'_{n-1}$ с условиями

$$\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_1^{(\widehat{k}_1)}) = w_1, \ldots, \eta(x_{n-2}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}) = w_{n-2} \in \Omega_2,$$

независимо от условий $v_1, \ldots, v_n, w_1, \ldots, w_{n-2} \in \Omega_2$, метка $\eta(x_j^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2}+k_{n-1})})$ не содержится в области определения минимального правила $G_{\widehat{\Sigma}'_{n-1}}$.

Рассмотрим два возможных варианта модификации слоя $\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)k_{n-1}}$:

1) Если $j = n-1$, то модифицируем $(n-1)$-й слой $\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)k_{n-1}}$ следующим образом:

$$\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)\widehat{k}_{n-1}} = \Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)k_{n-1}} \cdot \Sigma_n^{(n-1,n)} \cdot \Sigma_{n-1}^{(n-1,n)} \cdot \Sigma_n^{(n,n-1)}.$$

2) Если $j = n$, то модифицируем $(n-1)$-й слой $\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)k_{n-1}}$ следующим образом:

$$\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)\widehat{k}_{n-1}} = \Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)k_{n-1}} \cdot \Sigma_{n-1}^{(n,n-1)} \cdot \Sigma_n^{(n,n-1)} \cdot \Sigma_{n-1}^{(n-1,n)}.$$

В каждом из указанных случаев свободная разметка η сети $\widehat{\Sigma}'_{n-1}$ с произвольными условиями

$$\eta(x_1^{(0)}) = v_1, \ldots, \eta(x_n^{(0)}) = v_n, \eta(x_1^{(\widehat{k}_1)}) = w_1, \ldots, \eta(x_{n-2}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-2})}) = w_{n-2} \in \Omega_2$$

продолжается до свободной разметки η сети

$$\widehat{\Sigma} = (\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(n-2)1} \cdot \ldots \cdot \Sigma_{(n-2)\widehat{k}_{n-2}}) \cdot (\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)\widehat{k}_{n-1}})$$

с любыми условиями $\eta(x_{n-1}^{(\widehat{k}_1+\ldots+\widehat{k}_{n-1})}) = w_{n-1}, \eta(x_n^{(\widehat{k}_1+\ldots+\widehat{k}_{n-1})}) = w_n \in \Omega_2$.

Выход: модифицируя каноническое представление исходной сети Σ, мы построили «почти» каноническое представление

$$(\Sigma_{11} \cdot \ldots \cdot \Sigma_{1\widehat{k}_1}) \cdot \ldots \cdot (\Sigma_{(n-1)1} \cdot \ldots \cdot \Sigma_{(n-1)\widehat{k}_{n-1}})$$

новой биективной сети $\widehat{\Sigma}$, сложность которой равна $\|\Sigma\| + 3n - 3$. При этом сеть $\widehat{\Sigma}$ допускает свободную разметку с произвольными ограничениями $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из Ω_2.

Теорема 14. Пусть Σ — произвольная биективная сеть ширины n. Тогда её модификация $\widehat{\Sigma}$ транзитивна для любого множества Ω, мощность которого не менее чем $\|\Sigma\| + 4n - 3$.

Доказательство. Модификация $\widehat{\Sigma}$ произвольной биективной сети Σ допускает свободную разметку с произвольными ограничениями $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из Ω_2. Значит, согласно теореме 11, модификация $\widehat{\Sigma}$ допускает свободную разметку при любых ограничениях $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из $\mathbb{N}$.

Поскольку выполняется соотношение $\|\widehat{\Sigma}\| = \|\Sigma\| + 3n - 3$, то для проведения свободной разметки сети $\widehat{\Sigma}$ с произвольными ограничениями $\begin{pmatrix} v_1 & \ldots & v_n \\ w_1 & \ldots & w_n \end{pmatrix}$ из $\mathbb{N}$ потребуется не более чем $\|\Sigma\| + 4n - 3$ различных меток, и при выборе любого множества Ω,

мощность которого не менее чем $\|\Sigma\| + 4n - 3$, можно считать, что сеть $\widehat{\Sigma}$ допускает свободную разметку элементами множества Ω при произвольных ограничениях $\begin{pmatrix} v_1 & \dots & v_n \\ w_1 & \dots & w_n \end{pmatrix}$ из Ω. Последнее утверждение, согласно следствию 8, равносильно транзитивности сети $\widehat{\Sigma}$ для множества Ω. ■

Следствие 14. Для любого $n \geqslant 2$ существует сеть $\widehat{\Sigma}$ ширины n и веса $3n - 3$, транзитивная для всех множеств, мощность которых не менее чем $4n - 3$.

Сеть на рис. 4 служит иллюстрацией к описанному алгоритму и следствию 14, её разметка при любых ограничениях $\begin{pmatrix} v_1 & v_2 & v_3 & v_4 \\ w_1 & w_2 & w_3 & w_4 \end{pmatrix}$ из $\mathbb{N}$ является правильной и непротиворечивой.

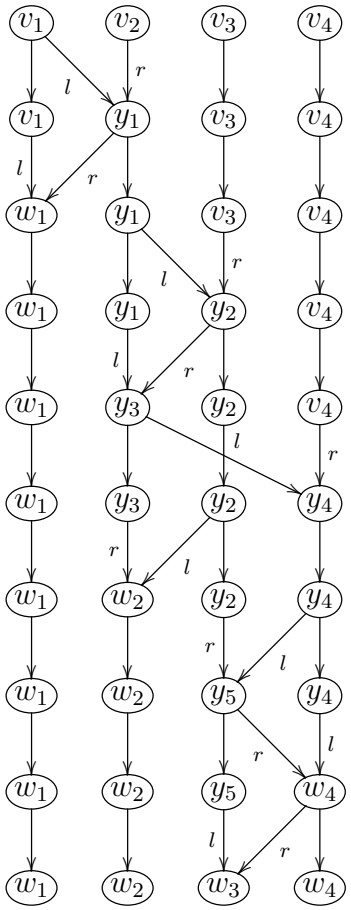

Рис. 4

Автор выражает благодарность профессору А. В. Черемушкину за постановку задачи и внимание к проводимым исследованиям.

ЛИТЕРАТУРА

1. *Белоусов В. Д.* Основы теории квазигрупп и луп. М.: Наука, 1967.
2. *Минк Х.* Перманенты. М.: Мир, 1982.
3. *Сачков В. Н., Тараканов В. Е.* Комбинаторика неотрицательных матриц. М.: ТВП, 2000. 448 с.

4. *Evans T.* Embedding incomplete latin squares // Amer. Math. Monthly. 1960. No. 67. P. 959–961.
5. *Smetaniuk B.* A new construction on latin squares I. A proof of the Evans conjecture // Ars Combinatoria. 1981. No. 11. P. 155–172.
6. *Anderson L. D. and Hilton A. J. W.* Thank Evans! // Proc. London Math. Soc. 1983. No. 47. P. 507–522.

REFERENCES

1. *Belousov V. D.* Osnovy teorii kvazigrupp i lup [Foundations of the Quasigroups and Loops theory]. Moscow, Nauka Publ., 1967. (in Russian)
2. *Mink Kh.* Permanenty [Permanents]. Moscow, Mir Publ., 1982. (in Russian)
3. *Sachkov V. N., Tarakanov V. E.* Kombinatorika neotritsatel'nykh matrits [Combinatorics of Nonnegative Matrices]. Moscow, TVP Publ., 2000. 448 p. (in Russian)
4. *Evans T.* Embedding incomplete latin squares. Amer. Math. Monthly, 1960, no. 67, pp. 959–961.
5. *Smetaniuk B.* A new construction on latin squares I. A proof of the Evans conjecture. Ars Combinatoria, 1981, no. 11, pp. 155–172.
6. *Anderson L. D. and Hilton A. J. W.* Thank Evans! Proc. London Math. Soc., 1983, no. 47, pp. 507–522.

УДК 512.55

О ПОЧТИ НИЛЬПОТЕНТНЫХ МНОГООБРАЗИЯХ АНТИКОММУТАТИВНЫХ МЕТАБЕЛЕВЫХ АЛГЕБР

О. В. Шулежко*, Н. П. Панов**

*Ульяновский государственный педагогический университет имени И. Н. Ульянова, г. Ульяновск, Россия

**Ульяновский государственный университет, г. Ульяновск, Россия

Представлены новые результаты, касающиеся многообразий антикоммутативных метабелевых алгебр над полем нулевой характеристики. Получены числовые характеристики многообразия всех антикоммутативных метабелевых алгебр. Для любого целого $m \geqslant 2$ доказано существование почти нильпотентного многообразия экспоненты m. Определены и изучены два почти нильпотентных многообразия подэкспоненциального роста, доказано, что других таких многообразий в исследуемом классе алгебр нет. Все результаты получены с помощью комбинаторных методов.

Ключевые слова: *многообразие линейных алгебр, нильпотентность, рост коразмерностей.*

DOI 10.17223/20710410/38/2

ON ALMOST NILPOTENT VARIETIES OF ANTICOMMUTATIVE METABELIAN ALGEBRAS

O. V. Shulezhko*, N. P. Panov**

*Ilya Ulyanov State Pedagogical University, Ulyanovsk, Russia

**Ulyanovsk State University, Ulyanovsk, Russia

E-mail: ol.shulezhko@gmail.com, nppanov@yandex.ru

Let Φ be a field of characteristic zero. We consider variety of anticommutative metabelian algebras, denoted $\mathbf{MA}$, in which the anticommutativity identity $x_1x_2 \equiv \equiv -x_2x_1$ and the metabelian identity $(x_1x_2)(x_3x_4) \equiv 0$ are satisfied. The associativity of multiplication is not assumed. Numerical invariants of the variety of all anticommutative metabelian algebras are obtained: the sequence of codimensions is $c_n(\mathbf{MA}) = n!/2$. An algorithm for computing the multiplicities of $m_\lambda(\mathbf{MA})$ for $n > 2$ is presented. We define a series of anticommutative metabelian algebras C_m for any integer $m \geqslant 2$ and prove the existence of almost nilpotent variety with PI-exponent of m. Moreover, two almost nilpotent varieties of subexponential growth are studied. The first variety is the well-known variety of all metabelian Lie algebras, denoted $\mathbf{A}^2$, the second — the almost nilpotent variety $\mathbf{V}_{\text{anti}}$ generated by the anticommutative metabelian algebra G, $\mathbf{V}_{\text{anti}} = \text{var}(G)$, which is defined in our investigation. In case of varieties of anticommutative metabelian algebras, it is shown that there are only two almost nilpotent varieties of subexponential growth: $\mathbf{A}^2$ and $\mathbf{V}_{\text{anti}}$. The proofs are based on the theory of irreducible modules, Young diagram and tableau, and some basic notions of the representation theory for the symmetric group. All results are obtained by means of combinatorial methods.

Keywords: *polynomial identity, variety, almost nilpotent, codimension growth.*

Введение

На протяжении всей работы основное поле Φ имеет нулевую характеристику. Линейной алгеброй, или алгеброй над полем, называют векторное пространство, на котором задана бинарная билинейная операция. Многообразие алгебр — совокупность линейных алгебр, удовлетворяющих фиксированному набору тождественных соотношений. Дополнительные сведения о теории алгебр с тождествами, а также все неопределяемые далее понятия и обозначения можно найти в [1, 2].

Обозначим через $F(X)$ свободную алгебру со счётным множеством образующих $X = \{x_1, x_2, \ldots\}$. Далее свободные образующие будем обозначать также другими латинскими буквами. В работе исследуются алгебры, в которых выполняется тождество антикоммутативности

$$x_1x_2 \equiv -x_2x_1 \tag{1}$$

и (по аналогии с алгебрами Ли) тождество метабелевости

$$(x_1x_2)(x_3x_4) \equiv 0. \tag{2}$$

Так как ассоциативность умножения не предполагается, в произведениях необходимо следить за расстановкой скобок. Договоримся опускать скобки в случае их левонормированной расстановки, например $(xy)z = xyz$. В алгебре A умножение справа на образующую a обозначим с помощью оператора $R_a : A \to A$, например $abR_c^2 = abcc$. Умножение справа на свободную образующую обозначим соответствующей заглавной буквой, например $xyY = xyy$. Условимся также помечать образующие с помощью специальных символов, например черты, для обозначения кососимметризации. Например,

$$\overline{x}_1\overline{x}_2 \ldots \overline{x}_n = \sum_{p\in S_n} (-1)^p x_{p(1)}x_{p(2)} \ldots x_{p(n)},$$

где S_n — симметрическая группа и $(-1)^p$ равно ± 1 в зависимости от чётности подстановки p.

В относительно свободной алгебре $F(X, \mathbf{V}) = F(X)/\mathrm{Id}(\mathbf{V})$ многообразия $\mathbf{V}$ рассмотрим так называемую полилинейную часть $P_n(\mathbf{V})$ — пространство полилинейных элементов степени n, $n \geqslant 1$, от образующих $x_1, \ldots, x_n$. Известно, что над основным полем нулевой характеристики любое тождество эквивалентно некоторой системе полилинейных тождеств, поэтому исследование строения полилинейных частей $P_n(\mathbf{V})$ позволяет получить информацию о многообразии $\mathbf{V}$. Пространство $P_n(\mathbf{V})$ как ΦS_n-модуль симметрической группы S_n имеет единственное с точностью до изоморфизма разложение в прямую сумму неприводимых подмодулей, соответствующих диаграммам Юнга разбиений $\lambda = (\lambda_1, \ldots, \lambda_l)$, $1 \leqslant l \leqslant n$, числа n, $\lambda \vdash n$. При этом кохарактер $\chi_n(\mathbf{V}) = \chi(P_n(\mathbf{V}))$ равен сумме неприводимых характеров χ_λ, $\lambda \vdash n$, взятых с кратностями $m_\lambda(\mathbf{V})$:

$$\chi_n(\mathbf{V}) = \sum_{\lambda\vdash n} m_\lambda(\mathbf{V})\chi_\lambda.$$

Число слагаемых в этой сумме называют кодлиной многообразия $\mathbf{V}$ и обозначают $l_n(\mathbf{V}) = \sum\limits_{\lambda\vdash n} m_\lambda(\mathbf{V})$.

Асимптотическое поведение последовательности коразмерностей $\{c_n(\mathbf{V})\}_{n\geqslant 1}$ определяет рост многообразия $\mathbf{V}$. Рост называют подэкспоненциальным, если для любого действительного $\alpha > 1$ найдётся такое натуральное n_0, что для всех $n \geqslant n_0$ выполнено неравенство $c_n(\mathbf{V}) < \alpha^n$. Говорят, что многообразие $\mathbf{V}$ имеет полиномиальный

рост, если коразмерности $c_n(\mathbf{V})$, $n \geqslant 1$, удовлетворяют ограничению $c_n(\mathbf{V}) \leqslant \gamma n^t$ для некоторых действительных чисел $\gamma, t \geqslant 0$. Если рост многообразия $\mathbf{V}$ выше подэкспоненциального, но последовательность коразмерностей экспоненциально ограничена, то оно экспоненциального роста. При этом предел $\lim\limits_{n\to\infty} \sqrt[n]{c_n(\mathbf{V})} = \beta > 1$, в случае его существования, называют экспонентой многообразия $\mathbf{V}$, $\exp(\mathbf{V}) = \beta$. Многообразие $\mathbf{V}$ является нильпотентным, если найдётся такое натуральное n_0, что для всех $n \geqslant n_0$ выполняется равенство $c_n(\mathbf{V}) = 0$. Говорят, что нильпотентное многообразие имеет нулевой рост. Ненильпотентное многообразие, все собственные подмногообразия которого нильпотентные, называют почти нильпотентным.

Известно, что единственным почти нильпотентным многообразием ассоциативных алгебр является многообразие $\mathbf{AC}$ всех коммутативных ассоциативных алгебр. В случае алгебр Ли почти нильпотентным является многообразие $\mathbf{A^2}$ всех метабелевых алгебр Ли. При изучении алгебр Лейбница были найдены два примера почти нильпотентных многообразий и доказано, что других нет [3]. Следует отметить, что все перечисленные многообразия имеют незначительный полиномиальный рост. В общем случае оказалось, что существуют достаточно экзотические примеры почти нильпотентных многообразий. В работе [4] впервые удалось построить почти нильпотентное многообразие экспоненты два. В случае нулевой характеристики основного поля доказано существование почти нильпотентных коммутативных метабелевых многообразий любой целой экспоненты, а именно: для любого натурального числа $m \geqslant 2$ существует почти нильпотентное многообразие, экспонента которого равна m [5]. В этом же классе алгебр оказалось только два почти нильпотентных многообразия подэкспоненциального роста [6]. В [7] дано описание всех почти нильпотентных многообразий подэкспоненциального роста в классе левонильпотентных ступени два алгебр, то есть алгебр, в которых выполнено тождество $x(yz) \equiv 0$. Обзор известных почти нильпотентных многообразий в этих классах алгебр выполнен в работе [8]. Удалось также установить существование континуального множества метабелевых почти нильпотентных многообразий полиномиального роста [9].

Настоящая работа является логическим продолжением серии исследований, посвящённых почти нильпотентным многообразиям в различных классах алгебр над полем нулевой характеристики. В ней представлены числовые характеристики многообразия всех антикоммутативных метабелевых алгебр. В данном классе алгебр для любого натурального $m \geqslant 2$ докажем существование почти нильпотентного многообразия экспоненты m, а также определим ровно два почти нильпотентных многообразия подэкспоненциального роста. Многообразие алгебр, удовлетворяющих тождествам (1) и (2), обозначим через $\mathbf{MA}$ и перейдём к изложению полученных результатов.

1. Числовые характеристики многообразия антикоммутативных метабелевых алгебр

При помощи определения множества антикоммутативных метабелевых алгебр A_n, $n = 2, 3, \ldots$, получим значения коразмерностей $c_n(\mathbf{MA})$. Пусть алгебра A_n задана образующими z_{ij}, e_{ij}, $1 \leqslant i, j \leqslant n$, и следующими определяющими соотношениями. Для любых i, j, k, l, $1 \leqslant i, j, k, l \leqslant n$,

1) $e_{ij}e_{kl} = 0$;
2) $z_{ij}z_{kl} = 0$;
3) $z_{ij}e_{kl} = -e_{kl}z_{ij} = \delta_{jk}z_{il}$, где δ_{jk} — символ Кронекера.

Теорема 1. Базис пространства $P_n(\mathbf{MA})$, $n \geqslant 2$, образуют элементы вида

$$x_{i_1}x_{i_2}\dots x_{i_n},\ i_1 > i_2, \tag{3}$$

и выполняется равенство

$$c_n(\mathbf{MA}) = \frac{n!}{2}. \tag{4}$$

Доказательство. С помощью тождеств (1) и (2) любой полилинейный моном степени n, $n \geqslant 2$, может быть записан в виде (3). Следовательно, каждый элемент пространства $P_n(\mathbf{MA})$ является линейной комбинацией $n!/2$ мономов вида (3). Докажем от противного их линейную независимость. Пусть имеет место тождество

$$\sum_{i=1}^{n!/2} \alpha_i x_{i_1}x_{i_2}\dots x_{i_n} \equiv 0,\ i_1 > i_2,$$

в котором не все коэффициенты $\alpha_i \in \Phi$ равны нулю. Проверим его справедливость в алгебре A_n. Пусть для некоторого s, $1 \leqslant s \leqslant n!/2$, $\alpha_s \neq 0$, тогда выполним подстановку $x_{s_1} = z_{11}$, $x_{s_2} = e_{12}$, …, $x_{s_n} = e_{(n-1)n}$ и получим равенство $\alpha_s z_{1n} = 0$, откуда $\alpha_s = 0$. Таким образом, все мономы вида (3) в пространстве $P_n(\mathbf{MA})$ линейно независимы и образуют базис размерности $n!/2$. ■

Заметим очевидное равенство $c_1(\mathbf{MA}) = 1$ и для $n \geqslant 2$ рассмотрим разложение ΦS_n-модуля $P_n(\mathbf{MA})$ в прямую сумму неприводимых подмодулей с кратностями $m_\lambda(\mathbf{MA})$, $\lambda \vdash n$. В силу тождества антикоммутативности $m_{(n)}(\mathbf{MA}) = 0$. Следовательно, модуль $P_2(\mathbf{MA})$ является одномерным и неприводимым, $m_{(1,1)}(\mathbf{MA}) = 1$ и $\chi_2(\mathbf{MA}) = \chi_{(1,1)}$. В следующей теореме сформулируем алгоритм вычисления кратностей $m_\lambda(\mathbf{MA})$ для $n > 2$ и диаграмм Юнга $\lambda \vdash n$ из двух и более строк.

Теорема 2. Зафиксируем $n > 2$. Выберем диаграмму Юнга, соответствующую разбиению $\lambda \vdash n$, которая содержит две или более строк. Из двух разных строк диаграммы λ удалим по одной клетке так, чтобы в результате удаления получилась диаграмма μ разбиения числа $n-2$, $\mu \vdash n-2$. Для всех таких диаграмм μ по формуле крюков найдем соответствующие значения размерностей d_μ. Тогда кратность $m_\lambda(\mathbf{MA})$ определяется как сумма всех значений d_μ.

Доказательство. Пусть H — подгруппа симметрической группы S_n, состоящая из тождественной подстановки и цикла $(n-1\ n)$. Обозначим через G подгруппу S_n, $n > 2$, $G = HS_{n-2} \cong H \times S_{n-2}$. Определим пространство полилинейных многочленов $Q_n = \operatorname{span}\{x_n x_{n-1} x_{\sigma(1)}\dots x_{\sigma(n-2)} : \sigma \in S_{n-2}\}$ и рассмотрим ΦG-модуль

$$Q_n(\mathbf{MA}) = \frac{Q_n}{Q_n \cap \mathrm{Id}(\mathbf{MA})}.$$

Заметим, что в относительно свободной алгебре $F(X, \mathbf{MA})$ ΦH-модуль, порождённый мономом $x_n x_{n-1}$, является неприводимым, одномерным и соответствует диаграмме Юнга разбиения $(1,1) \vdash 2$. При этом известно, что регулярный модуль ΦS_n разлагается в прямую сумму неприводимых подмодулей с кратностями, совпадающими с размерностями соответствующих неприводимых подмодулей. Из конструкции ΦG-модуля $Q_n(\mathbf{MA})$ получим индуцированный ΦS_n-модуль $P_n(\mathbf{MA})$, который разлагается в прямую сумму неприводимых подмодулей с диаграммами Юнга, заданными по правилу Литтлвуда — Ричардсона. В разложении модуля $P_n(\mathbf{MA})$ каждая диаграмма

$\lambda \vdash n$ каждого неприводимого подмодуля получена присоединением клеток диаграммы $(1,1) \vdash 2$ к двум различным строкам некоторых диаграмм $\mu \vdash n-2$, причём за счёт присоединения двух новых клеток диаграммы μ расширяются вправо или вниз. Следовательно, в разложении $P_n(\mathbf{MA})$ кратность $m_\lambda(\mathbf{MA})$ определяется как сумма значений размерностей d_μ модулей с диаграммами $\mu \vdash n-2$, которые получены удалением двух клеток из разных строк диаграммы λ. В частности, $m_{(1^n)}(\mathbf{MA}) = d_{(1^{n-2})} = 1$. Теорема доказана. ■

Пример 1. Для $n = 4$ вычислим $m_\lambda(\mathbf{MA})$ и d_λ, $\lambda \vdash n$,

$$\begin{array}{ll} m_{(4)} = 0, & d_{(4)} = 1, \\ m_{(3,1)} = d_{(2)} = 1, & d_{(3,1)} = 3, \\ m_{(2,2)} = d_{(1,1)} = 1, & d_{(2,2)} = 2, \\ m_{(2,1,1)} = d_{(2)} + d_{(1,1)} = 2, & d_{(2,1,1)} = 3, \\ m_{(1,1,1,1)} = d_{(1,1)} = 1, & d_{(1,1,1,1)} = 1. \end{array}$$

Сложим выписанные значения и получим равенство (4):

$$c_4(\mathbf{MA}) = 0 \cdot 1 + 1 \cdot 3 + 1 \cdot 2 + 2 \cdot 3 + 1 \cdot 1 = 12 = 4!/2.$$

2. Почти нильпотентные многообразия экспоненциального роста

При помощи построения серии антикоммутативных метабелевых алгебр C_m, $m \geqslant 2$, докажем, что в данном классе алгебр для любого целого $m \geqslant 2$ существует почти нильпотентное многообразие экспоненты m. Для каждого $m \geqslant 2$ определим алгебру C_m образующими $z_1, z_2, c_1, \ldots, c_m$ и следующими определяющими соотношениями:

1) $uv + vu = 0,\ u, v \in C_m$;
2) $c_i c_j = 0,\ 1 \leqslant i, j \leqslant m$;
3) $z_i c_j = 0,\ i = 1, 2,\ 1 \leqslant j \leqslant m$;
4) $z_1 z_2 w(R_{c_1}, \ldots, R_{c_m}) z_i = 0,\ i = 1, 2,\ \deg w \geqslant 0$;
5) $(z_1 z_2 w(R_{c_1}, \ldots, R_{c_m}))(z_1 z_2 w'(R_{c_1}, \ldots, R_{c_m})) = 0,\ \deg w \geqslant 0,\ \deg w' \geqslant 0$;
6) $z_1 z_2 (R_{c_1} \ldots R_{c_m})^k c_{i_1} \ldots c_{i_s} c_{i_{s+1}} \ldots c_{i_t} + z_1 z_2 (R_{c_1} \ldots R_{c_m})^k c_{i_1} \ldots c_{i_{s+1}} c_{i_s} \ldots c_{i_t} = 0$, $k \geqslant 0,\ 1 \leqslant s < t \leqslant m,\ 1 \leqslant i_1, \ldots, i_t \leqslant m$,

где слова w, w' — ассоциативные мономы от операторов R_{c_j}, $1 \leqslant j \leqslant m$. Заметим, что последнее определяющее соотношение влечёт равенство

$$z_1 z_2 (R_{c_1} \ldots R_{c_m})^k w(R_{c_1}, \ldots, R_{c_m}) = 0,$$

в котором $k \geqslant 0$ и ассоциативный моном $w(R_{c_1}, \ldots, R_{c_m})$, $2 \leqslant \deg w \leqslant m$, содержит по меньшей мере два одинаковых R_{c_i}, $1 \leqslant i \leqslant m$. Таким образом, базис алгебры C_m представлен левонормированными элементами

$$z_1, z_2,\ c_1,\ \ldots,\ c_m,\ z_1 z_2 (R_{c_1} \ldots R_{c_m})^k,\ z_1 z_2 (R_{c_1} \ldots R_{c_m})^k c_{i_1} c_{i_2} \ldots c_{i_t}$$

для всех $k \geqslant 0$, $1 \leqslant t < m$, $1 \leqslant i_1 < i_2 < \cdots < i_t \leqslant m$.

Нетрудно проверить, что в алгебре C_m, как и в коммутативной метабелевой алгебре B_m из работы [5], выполняются тождественные соотношения

$$x_0 y_0 X^3 \equiv 0; \tag{5}$$

$$x_0 y_0 X^2 Z_1 \ldots Z_s Y^2 \equiv 0, \tag{6}$$

где остаток от деления s на m отличен от $m-2$. Так как основное поле Φ имеет нулевую характеристику, алгебра C_m удовлетворяет следующим линеаризациям тождеств (5) и (6):

$$x_0y_0X_1XX_1 \equiv -x_0y_0X_1X_1X - x_0y_0XX_1X_1; \tag{7}$$

$$x_0y_0XX_1wY^2 \equiv -x_0y_0X_1XwY^2; \tag{8}$$

$$x_0y_0X^2wYY_1 \equiv -x_0y_0X^2wY_1Y; \tag{9}$$

$$x_0y_0XX_1wYY_1 \equiv -x_0y_0X_1XwYY_1 - x_0y_0XX_1wY_1Y - x_0y_0X_1XwY_1Y, \tag{10}$$

где $w = Z_1 \dots Z_s$ и s отлично от $m-2$ по модулю m.

Утверждение 1. Пусть $w = w(X_1, \dots, X_m)$ — ассоциативный моном с условием

$$\deg w - \min_{1 \leqslant i \leqslant m} \{\deg_{X_i} w\} \cdot m \geqslant 2m-1. \tag{11}$$

Тогда в алгебре C_m выполняется тождество $x_0y_0w(X_1, \dots, X_m) \equiv 0$.

Доказательство. Пусть w зависит не от всех X_i, $1 \leqslant i \leqslant m$. Без потери общности примем $\deg_{X_m} w = 0$ и из (11) получим неравенство $\deg w \geqslant 2m-1$. Тогда среди первых m букв слова w найдутся две одинаковые, например пара X_{m-1}, между которыми расположено не более $m-2$ других букв. С помощью тождеств (10) и (7) представим x_0y_0w в виде линейной комбинации мономов, в каждом из которых выбранные буквы X_{m-1} размещены на соседних позициях. Так как $\deg w \geqslant 2m-1$, в каждом из полученных мономов справа от X_{m-1}^2 находится по меньшей мере $m-1$ букв. Если в некотором мономе среди них найдётся X_{m-1}, то, применив, возможно несколько раз, тождество (9), получим три подряд одинаковые буквы и соответственно (5) равный нулю элемент. Иначе среди данных $m-1$ букв общее число различных не превышает $m-2$ и найдутся две одинаковые. В силу кососимметричности перестановки букв (9) все такие мономы также равны нулю.

Для удобства обозначим $\theta = \min\limits_{1 \leqslant i \leqslant m} \{\deg_{X_i} w\}$. Пусть теперь $\theta \geqslant 1$, $\deg w \geqslant 3m-1$. В слове w в любой последовательности букв X_i, $1 \leqslant i \leqslant m$, длины $m+1$ найдутся две одинаковые буквы, между которыми расположено не более $m-1$ других букв. Аналогично предыдущему случаю выберем пару таких букв и с помощью тождеств (10) и (7) запишем x_0y_0w в виде линейной комбинации мономов, содержащих выбранные буквы на соседних позициях. С помощью приведённых выше рассуждений отбросим из полученной линейной комбинации все нулевые мономы. В оставшихся мономах с помощью тождеств (8), (9) упорядочим буквы X_i, $1 \leqslant i \leqslant m$, следующим образом.

При $\theta > 1$ получим мономы вида

$$x_0y_0w'(X_1 \dots X_m)^{r_1}(X_1 \dots \widehat{X_s} \dots X_mX_s^2X_1 \dots \widehat{X_s} \dots X_m)(X_1 \dots X_m)^{r_2}w'', \tag{12}$$

где $r_1, r_2 \geqslant 0$, $r_1 + r_2 + 2 \leqslant \theta$, обозначение $\widehat{X_s}$ значит пропуск X_s и w', w'' — ассоциативные мономы, зависящие не от всех X_i, $1 \leqslant i \leqslant m$. По условию (11) хотя бы одно из слов w', w'' имеет длину не менее m и поэтому среди любых m букв содержит две одинаковые. То есть в силу тождеств (8), (9) все полученные мономы также равны нулю.

Если $\theta = 1$, аналогичные рассуждения приводят к линейной комбинации одночленов вида

$$x_0y_0w_1X_1 \dots \widehat{X_s} \dots X_mX_s^2w_2; \tag{13}$$

$$x_0y_0w_3X_s^2X_1 \dots \widehat{X_s} \dots X_mw_4, \tag{14}$$

в которых слова $w_1, \dots, w_4$ зависят не от всех X_i, $1 \leqslant i \leqslant m$. Покажем равенство нулю мономов вида (13). Если слово w_1 (или w_2) длины не менее m, то с помощью тождества (8) (или (9)) в слове w_1 (или w_2) разместим подряд две одинаковые буквы и в силу (6) получим, что моном (13) равен нулю. Пусть теперь $n = 3m - 1$ и каждое из слов w_1, w_2 длины $m - 1$. Возможны следующие два случая: если $\deg_{X_s} w_2 \geqslant 1$ (в частности, при $m = 2$), то в силу (9), (5) моном (13) равен нулю. Так как $\theta = 1$, при $\deg_{X_s} w_2 = 0$ среди $m-1$ букв слова w_2 найдутся две одинаковые, и в силу тождества (9) моном (13) равен нулю. Равенство нулю мономов вида (14) получим в результате аналогичных рассуждений. ■

Пусть $L_n = \operatorname{span}\{x_0 y_0 x_{\sigma(1)} \dots x_{\sigma(n)} : \sigma \in S_n\}$ — пространство полилинейных левонормированных мономов с фиксированным произведением $x_0 y_0$ во внутренних скобках, $\mathbf{W}_m = \operatorname{var}(C_m)$ — многообразие, порождённое алгеброй C_m, $m \geqslant 2$. Тогда рассмотрим ΦS_n-модуль

$$L_n(\mathbf{W}_m) = \frac{L_n}{L_n \cap \operatorname{Id}(\mathbf{W}_m)}$$

с кохарактером $\chi_n^L(\mathbf{W}_m)$, который разлагается в сумму неприводимых характеров с кратностями m_λ^L, $\lambda \vdash n$.

Так как характеристика поля Φ равна нулю, то, используя условие эквивалентности полилинейных и соответствующих полиоднородных тождеств, докажем следующее утверждение.

Утверждение 2. Если $m_\lambda^L \neq 0$, $n \geqslant 3m - 1$, то $\lambda \vdash n$ удовлетворяет следующим условиям:

1) $\lambda_1' = m$;
2) $n - \lambda_m \cdot m < 2m - 1$;
3) $\lambda_1 - \lambda_m \leqslant 2$.

Доказательство. Зафиксируем стандартную таблицу Юнга T_λ, $\lambda \vdash n \geqslant 3m-1$, и обозначим соответственно R_{T_λ}, C_{T_λ} стабилизаторы строк и столбцов таблицы T_λ. Обозначим через

$$f = e_{T_\lambda}(x_0 y_0 x_1 \dots x_n) = \sum_{p \in R_{T_\lambda}} p \sum_{q \in C_{T_\lambda}} (-1)^q q(x_0 y_0 x_1 \dots x_n)$$

полилинейный элемент, порождающий неприводимый подмодуль модуля $L_n(\mathbf{W}_m)$. Пусть полиоднородный элемент h получен из f в результате отождествления свободных образующих, индексы которых находятся в одних и тех же строках таблицы T_λ.

Докажем первое условие. Пусть сначала $\lambda_1' < m$, тогда каждый моном в h удовлетворяет условиям утверждения 1, где $\min\limits_{1 \leqslant i \leqslant m} \{\deg_{X_i} w\} = 0$, $\deg w \geqslant 3m - 1$. Следовательно, в алгебре C_m выполняется тождество $h \equiv 0$ и $m_\lambda^L = 0$ при $\lambda_1' < m$. Пусть теперь $\lambda_1' > m$. Тогда f равен сумме полилинейных элементов, кососимметрических по более чем m образующим x_i, $1 \leqslant i \leqslant m$, и, следовательно, тождественно равных нулю в алгебре C_m. Таким образом, $m_\lambda^L = 0$ при $\lambda_1' > m$, и первое условие доказано.

Второе условие является прямым следствием утверждения 1.

Докажем третье условие. По доказательству утверждения 1 ненулевой полиоднородный элемент h полистепени $(\lambda_1, \dots, \lambda_m)$ по образующим $x_1, \dots, x_m$ по модулю тождеств многообразия $\mathbf{W}_m$ равен линейной комбинации ненулевых мономов $x_0 y_0 w$ вида (12), в которых по условию 2 мономы w', w'', зависящие не от всех X_i, $1 \leqslant i \leqslant m$, удовлетворяют неравенствам $\deg w' \leqslant m - 1$, $\deg w'' \leqslant m - 1$. В силу тождеств (8)

и (9) каждое из слов w', w'' не может содержать повторяющиеся буквы. Так как буквы в словах w', w'' могут совпадать, то выполняется неравенство $\lambda_1 - \lambda_m \leqslant 2$. ■

Утверждение 3 [10]. Пусть T — таблица Юнга, соответствующая разбиению $\lambda \vdash n$, и $M = M_1 \oplus \cdots \oplus M_k$ — ΦS_n-модуль, где M_i — изоморфные неприводимые подмодули с характером χ_λ. Тогда k равно максимальному числу линейно независимых элементов $g \in M$, таких, что $\sigma \cdot g = g$ для любого элемента $\sigma \in R_T$.

Утверждение 4. Существует такая константа $C = C(m)$, что для всех $\lambda \vdash n$, $n \geqslant 3m - 1$, выполняется неравенство $m_\lambda^L \leqslant C$.

Доказательство. Зафиксируем диаграмму $\lambda = (\lambda_1, \ldots, \lambda_m)$, $\lambda \vdash n$. В разложении модуля $L_n(\mathbf{W}_m)$ не равная нулю кратность m_λ^L определяется как размерность пространства полилинейных элементов, удовлетворяющих условиям утверждения 3. В силу эквивалентности полилинейных и соответствующих полиоднородных тождеств и по условию 1 утверждения 2 для оценки данной размерности достаточно по модулю тождеств многообразия $\mathbf{W}_m$ оценить размерность пространства $L_{\lambda_1,\ldots,\lambda_m}$ полиоднородных элементов $x_0 y_0 w(X_1, \ldots, X_m)$ полистепени $(\lambda_1, \ldots, \lambda_m)$ от образующих $x_1, \ldots, x_m$. С помощью тождеств (5)–(10) любой ненулевой элемент пространства $L_{\lambda_1,\ldots,\lambda_m}$ может быть представлен линейной комбинацией ненулевых мономов вида (12), в которых каждое из слов w', w'' состоит из различных букв и имеет длину не более $m - 1$. Нетрудно убедиться, что с помощью тождеств (5)–(10) все такие мономы могут быть приведены к виду

$$x_0 y_0 w'(X_1 \ldots X_m)(X_m X_1 \ldots X_{m-1})(X_1 \ldots X_m)^{r_1+r_2} w'', \tag{15}$$

где буквы в словах w', w'' упорядочены.

Так как мономы (15), отличающиеся только порядком букв в соответствующих словах w', w'', по модулю тождеств (8), (9) с точностью до знака равны, то для диаграммы $\lambda = (\lambda_1, \ldots, \lambda_m)$ число различных мономов (15) равно количеству способов выбора различных букв в любом из слов w', w''. В результате сравнения получим, что наибольшее число различных мономов (15), всего $2^m - 1$, соответствует прямоугольной диаграмме λ. То есть при $\lambda_1 = \ldots = \lambda_m$ имеем

$$m_\lambda^L \leqslant \dim L_{\lambda_1,\ldots,\lambda_m} \leqslant \sum_{i=0}^{m-1} \binom{m}{i} = 2^m - 1.$$

Утверждение доказано. ■

Обозначим $c_n^L(\mathbf{W}_m) = \dim L_n(\mathbf{W}_m)$ и установим связь между коразмерностями $c_n^L(\mathbf{W}_m)$ и $c_n(\mathbf{W}_m)$, $n \geqslant 1$.

Утверждение 5. Для любого $n \geqslant 1$ выполняется равенство

$$c_{n+2}(\mathbf{W}_m) = \frac{(n+1)(n+2)}{2} c_n^L(\mathbf{W}_m).$$

Доказательство аналогично проведённому в работе [5] для коммутативного случая.

Теорема 3. Для любого целого $m \geqslant 2$ существует почти нильпотентное антикоммутативное метабелево многообразие экспоненты m.

Доказательство. Рассмотрим ΦS_n-модуль $L'_n(\mathbf{W}_m) \cong L_n(\mathbf{W}_m)$, $n \geqslant 3m - 1$, элементы которого вместо произведения $x_0 y_0$ содержат $x_{n+2} x_{n+1}$. Аналогично доказательству теоремы 2 из ΦS_n-модуля $L'_n(\mathbf{W}_m)$ получим индуцированный ΦS_{n+2}-модуль

$P_{n+2}(\mathbf{W}_m)$. В силу утверждений 2 и 4 в разложении кохарактера $\chi_{n+2}(\mathbf{W}_m)$ в сумму неприводимых характеров все ненулевые кратности $m_\lambda(\mathbf{W}_m)$ ограничены сверху константой и отвечают диаграммам $\lambda \vdash n+2$, у которых вне прямоугольника $m \times \lambda_m$ находится не более $2m$ клеток. Известно, что при достаточно больших n размерность d_λ с диаграммой $\lambda \vdash n$ такого вида удовлетворяет неравенствам $n^\beta m^n \leqslant d_\lambda \leqslant n^\alpha m^n$ для фиксированных α, β [11]. Следовательно, $\exp(\mathbf{W}_m) = m$. При этом полилинейные части всех ненильпотентных подмногообразий многообразия $\mathbf{W}_m$ также удовлетворяют указанным ограничениям. Остаётся заметить, что в любом ненильпотентном многообразии существует почти нильпотентное подмногообразие [4, теорема 1]. ■

3. Почти нильпотентные многообразия подэкспоненциального роста

Далее в исследуемом классе алгебр определим ровно два почти нильпотентных многообразия подэкспоненциального роста. Первое такое многообразие — это хорошо известное многообразие всех метабелевых алгебр Ли $\mathbf{A}^2$. Приведём необходимые в дальнейшем свойства многообразия $\mathbf{A}^2$. Полилинейная часть $P_n(\mathbf{A}^2)$, $n \geqslant 2$, имеет следующие числовые характеристики:

$$c_n(\mathbf{A}^2) = n-1, \quad \chi_n(\mathbf{A}^2) = \chi_{(n-1,1)}, \quad l_n(\mathbf{A}^2) = 1.$$

Утверждение 6. Многообразие $\mathbf{A}^2$ не является подмногообразием многообразия $\mathbf{V} \subset \mathbf{MA}$ тогда и только тогда, когда в многообразии $\mathbf{V}$ для некоторого $k \geqslant 1$ выполнено тождество $x_0 X^k \equiv 0$.

Доказательство. Так как $x_0 X^k \notin \mathrm{Id}(\mathbf{A}^2)$, $k \geqslant 1$, то остаётся доказать необходимость. Пусть $\mathbf{A}^2 \not\subset \mathbf{V}$, тогда из разложения полилинейной части $P_n(\mathbf{A}^2)$, $n \geqslant 2$, следует, что в многообразии $\mathbf{V}$ должно выполняться некоторое тождество, соответствующее диаграмме Юнга $(n-1,1)$. Полиоднородные элементы, построенные по стандартным таблицам данной диаграммы, имеют вид $\overline{x}_1 X_1^i \overline{x}_2 X_1^{n-2-i}$, $i = 0, \ldots, n-2$. Поэтому тождество в многообразии $\mathbf{V}$ может быть записано как

$$\sum_{i=0}^{n-2} \alpha_i \overline{x}_1 X_1^i \overline{x}_2 X_1^{n-2-i} \equiv 0, \ \alpha_i \in \Phi.$$

С помощью тождества антикоммутативности перепишем его в виде

$$\left(2\alpha_0 + \sum_{i=1}^{n-2} \alpha_i\right) x_2 X_1^{n-1} \equiv 0.$$

При этом необходимо принять $2\alpha_0 + \sum\limits_{i=1}^{n-2} \alpha_i \neq 0$, иначе получим, что исходное тождество выполнено в многообразии $\mathbf{A}^2 \subset \mathbf{MA}$. Таким образом, для некоторого $n \geqslant 2$ тождество $x_2 X_1^{n-1} \equiv 0$ выполнено в многообразии $\mathbf{V}$. Заметим, что для $n = 1$ доказательство очевидно. ■

Второе почти нильпотентное многообразие подэкспоненциального роста $\mathbf{V}_{\mathrm{anti}}$ порождается следующей антикоммутативной метабелевой алгеброй G, $\mathbf{V}_{\mathrm{anti}} = \mathrm{var}(G)$. Определим неассоциативную алгебру G бесконечным числом образующих $e_1, e_2, \ldots$ и следующими определяющими соотношениями:

1) $we_i = -e_i w$, $\deg w \geqslant 1$;
2) $w_1 w_2 = 0$, $\deg w_1 \geqslant 2$, $\deg w_2 \geqslant 2$;
3) $e_{i_{p(1)}} e_{i_{p(2)}} \ldots e_{i_{p(n)}} = (-1)^p e_{i_1} e_{i_2} \ldots e_{i_n}$, $p \in S_n$, $n \geqslant 2$.

Легко видеть, что в алгебре G в силу определяющего соотношения 3 любой базисный элемент степени два или более по образующей e_i, $i \geqslant 1$, равен нулю. Обозначим G^2 идеал алгебры G, который образован всеми произведениями её элементов. Заметим, что в силу тождества метабелевости G^2 является алгеброй с нулевым умножением.

Утверждение 7. В многообразии $\mathbf{V}_{\text{anti}}$ выполнены следующие тождества:

$$xyzt + xytz \equiv 0; \tag{16}$$

$$xyzt + zyxt + xtzy + ztxy \equiv 0. \tag{17}$$

Доказательство. Так как данные тождества являются полилинейными, достаточно доказать их справедливость для базисных элементов алгебры G. Если в первое тождество вместо z или t подставить базисный элемент G^2, то при любых x, y оба слагаемых равны нулю. В результате подстановки элементов e_i получим верное равенство по определяющему соотношению 3.

В тождество (17) вместо одной из свободных образующих подставим базисный элемент $g \in G^2$, а вместо остальных — различные e_i, $i \geqslant 1$. В силу тождеств антикоммутативности и метабелевости останется пара ненулевых слагаемых, в каждом из которых g находится на первом месте, а перестановки образующих e_i отличаются на одну транспозицию, то есть имеют разную чётность. Из определяющего соотношения 3 полученная сумма равна нулю. Тождество (17) также обращается в верное равенство при подстановке различных e_i, $i \geqslant 1$. Достаточно заметить, что из монома $xyzt$ слагаемые $zyxt$, $xtzy$ получены одной транспозицией образующих, а одночлен $ztxy$ — двумя транспозициями. ■

С помощью утверждения 7 докажем следующие равенства для числовых характеристик полилинейной части $P_n(\mathbf{V}_{anti})$.

Утверждение 8. При $n \geqslant 3$ выполняются равенства

$$c_n(\mathbf{V}_{\text{anti}}) = n, \quad l_n(\mathbf{V}_{\text{anti}}) = 2, \quad \chi_n(\mathbf{V}_{\text{anti}}) = \chi_{(2,n-2)} + \chi_{(1^n)}.$$

Доказательство. Оценим значение коразмерности $c_n(\mathbf{V}_{\text{anti}})$, $n \geqslant 3$, сверху. Для этого покажем, что любой элемент пространства $P_n(\mathbf{V}_{\text{anti}})$ может быть записан в виде линейной комбинации n элементов двух видов:

$$x_{n-1}x_{n-2}x_nx_{n-3}\dots x_1; \tag{18}$$

$$x_nx_ix_{i_1}x_{i_2}\dots x_{i_{n-2}},\ 1 \leqslant i \leqslant n-1,\ i_1 > i_2 > \dots > i_{n-2}. \tag{19}$$

В любом мономе пространства $P_n(\mathbf{V}_{\text{anti}})$ в силу тождества (16) образующие, начиная с третьей позиции, могут быть упорядочены по убыванию индексов, поэтому возможны три случая записи остальных мономов:

1) $x_{n-1}x_ix_nx_{n-2}\dots$, $1 \leqslant i < n-2$;
2) $x_{n-2}x_ix_nx_{n-1}\dots$, $1 \leqslant i < n-2$;
3) $x_ix_jx_nx_{n-1}\dots$, $1 \leqslant j < i < n-2$.

Воспользуемся в каждом из случаев тождеством (17):

$$x_{n-1}x_ix_nx_{n-2}\dots \equiv -x_nx_ix_{n-1}x_{n-2}\dots - x_{n-1}x_{n-2}x_nx_i\dots - x_nx_{n-2}x_{n-1}x_i\dots,$$

$$x_{n-2}x_ix_nx_{n-1}\dots \equiv -x_nx_ix_{n-2}x_{n-1}\dots - x_{n-2}x_{n-1}x_nx_i\dots - x_nx_{n-1}x_{n-2}x_i\dots,$$

$$x_ix_jx_nx_{n-1}\dots \equiv -x_nx_jx_ix_{n-1}\dots - x_ix_{n-1}x_nx_j\dots - x_nx_{n-1}x_ix_j\dots$$

Получили линейные комбинации элементов вида (18) и (19), и в третьем разложении слагаемое $x_i x_{n-1} x_n x_j \dots$ соответствует первому случаю. Таким образом, имеем неравенство $c_n(\mathbf{V}_{\text{anti}}) \leqslant n$. Оценим $c_n(\mathbf{V}_{\text{anti}})$ снизу. Для этого рассмотрим полиоднородные элементы $h_{(2,1^{n-2})}$ и $h_{(1^n)}$, линеаризации которых порождают неприводимые подмодули с диаграммами $(2,1^{n-2})$ и (1^n). Пусть $h_{(2,1^{n-2})} = \overline{x}_1\overline{x}_2\dots\overline{x}_{n-1}x_1$, тогда вместо x_1 подставим сумму $g+e_1$, $g \in G^2$, а вместо x_i подставим e_i, $i \geqslant 2$. Из определяющих соотношений алгебры G получим ненулевой элемент

$$\overline{(g+e_1)}\overline{e}_2\dots\overline{e}_{n-1}(g+e_1) = \overline{g}\,\overline{e}_2\dots\overline{e}_{n-1}e_1 = 2g\overline{e}_2\dots\overline{e}_{n-1}e_1.$$

Следовательно, соответствующий неприводимый подмодуль не является нулевым и $m_{(2,1^{n-2})}(\mathbf{V}_{\text{anti}}) \geqslant 1$.

В многочлен $h_{(1^n)} = \overline{x}_1\overline{x}_2\dots\overline{x}_n$ вместо x_i, $1 \leqslant i \leqslant n$, подставим e_i. Результат подстановки равен $n!e_1e_2\dots e_n$, и неприводимый подмодуль с диаграммой (1^n) также не является нулевым, $m_{(1^n)}(\mathbf{V}_{\text{anti}}) = 1$. Таким образом, для $c_n(\mathbf{V}_{\text{anti}})$ имеем следующую оценку снизу:

$$c_n(\mathbf{V}_{\text{anti}}) \geqslant m_{(2,1^{n-2})}(\mathbf{V}_{\text{anti}})d_{(2,1^{n-2})} + m_{(1^n)}(\mathbf{V}_{\text{anti}})d_{(1^n)} = (n-1)m_{(2,1^{n-2})}(\mathbf{V}_{\text{anti}}) + 1 \geqslant n.$$

Так как $c_n(\mathbf{V}_{\text{anti}}) \leqslant n$, то $c_n(\mathbf{V}_{\text{anti}}) = n$, $m_{(2,1^{n-2})}(\mathbf{V}_{\text{anti}}) = 1$, и получили искомые равенства. ■

Отметим очевидные равенства для $n = 1, 2$:

$$\begin{aligned} c_1(\mathbf{V}_{\text{anti}}) &= 1, \quad l_1(\mathbf{V}_{\text{anti}}) = 1, \quad \chi_1(\mathbf{V}_{\text{anti}}) = \chi_{(1)}, \\ c_2(\mathbf{V}_{\text{anti}}) &= 1, \quad l_2(\mathbf{V}_{\text{anti}}) = 1, \quad \chi_2(\mathbf{V}_{\text{anti}}) = \chi_{(1,1)}. \end{aligned}$$

Утверждение 9. Многообразие $\mathbf{V}_{\text{anti}}$ не является подмногообразием многообразия $\mathbf{V} \subset \mathbf{MA}$ тогда и только тогда, когда в многообразии $\mathbf{V}$ выполнено тождество $x_0\overline{x}_1\overline{x}_2\dots\overline{x}_m \equiv 0$ для некоторого $m \geqslant 2$.

Доказательство. Так как $x_0\overline{x}_1\overline{x}_2\dots\overline{x}_m \notin \mathrm{Id}(G) = \mathrm{Id}(\mathbf{V}_{\text{anti}})$, $m \geqslant 1$, то остаётся доказать необходимость. Из доказанного в утверждении 8 разложения характера $\chi_n(\mathbf{V}_{\text{anti}})$ в сумму неприводимых характеров следует, что в многообразии $\mathbf{V}_{\text{anti}}$ выполняется некоторое тождество, соответствующее одной из диаграмм (1^n), $(2,1^{n-2})$ при $n \geqslant 3$. Пусть в многообразии $\mathbf{V}$ выполнено тождество $\overline{x}_1\overline{x}_2\dots\overline{x}_n \equiv 0$. Тогда, заменив образующую x_1 на произведение x_0x_1, получим следствие $x_0x_1\overline{x}_2\dots\overline{x}_n \equiv 0$. Запишем кососимметризацию данного следствия по образующим $x_1, x_2, \dots, x_n$:

$$\sum_{\sigma\in H_n} (-1)^\sigma x_0\overline{x}_{\sigma(1)}\overline{x}_{\sigma(2)}\dots\overline{x}_{\sigma(n)} \equiv 0.$$

Здесь H_n — группа всех подстановок, которые оставляют на месте единицу. Из соотношения

$$\overline{x}_1\overline{x}_2\dots\overline{x}_n = (-1)^\theta\overline{x}_{\theta(1)}\overline{x}_{\theta(2)}\dots\overline{x}_{\theta(n)},\ \theta \in S_n,$$

получим в многообразии $\mathbf{V}$ тождество $(n-1)!x_0\overline{x}_1\overline{x}_2\dots\overline{x}_n \equiv 0$.

Для диаграммы $(2,1^{n-2})$ тождество может быть записано от $n-1$ образующих $x_1, \dots, x_{n-1}$ следующим образом:

$$\sum_{i=1}^{n-1} \alpha_i\overline{x}_1\dots\overline{x}_i x_1\overline{x}_{i+1}\dots\overline{x}_{n-1} \equiv 0,$$

где последнее слагаемое имеет вид $\alpha_{n-1}\overline{x}_1\dots\overline{x}_{n-1}x_1$. С помощью тождества антикоммутативности перепишем полилинейный элемент при $i=1$:

$$\overline{x}_1x_1\overline{x}_2\dots\overline{x}_{n-1}\equiv\sum_{j=2}^{n-1}(-1)^jx_1\overline{x}_2\dots\overline{x}_jx_1\overline{x}_{j+1}\dots\overline{x}_{n-1}.$$

В исходное тождество вместо x_1 подставим сумму $x_0x_1+x_n$ и по модулю тождеств многообразия $\mathbf{MA}$ получим

$$\sum_{i=2}^{n-1}\left((-1)^i\alpha_1+2\alpha_i\right)x_0x_1\overline{x}_2\dots\overline{x}_ix_n\overline{x}_{i+1}\dots\overline{x}_{n-1}\equiv 0.$$

К данному тождеству применим кососимметризацию по $x_1,\dots,x_n$. Для удобства обозначим через $\widetilde{H}_n$ группу всех подстановок, которые оставляют на месте 1 и n, и рассмотрим результат кососимметризации слагаемого с коэффициентом $(-1)^k\alpha_1+2\alpha_k$, $2\leqslant k\leqslant n-1$:

$$\sum_{\sigma\in\widetilde{H}_n}(-1)^\sigma x_0\overline{x}_{\sigma(1)}\overline{x}_{\sigma(2)}\dots\overline{x}_{\sigma(k)}\overline{x}_{\sigma(n)}\overline{x}_{\sigma(k+1)}\dots\overline{x}_{\sigma(n-1)}=$$
$$=\sum_{\sigma\in\widetilde{H}_n}x_0\overline{x}_1\overline{x}_2\dots\overline{x}_k\overline{x}_n\overline{x}_{k+1}\dots\overline{x}_{n-1}=(-1)^{n-1-k}(n-2)!x_0\overline{x}_1\overline{x}_2\dots\overline{x}_n.$$

Таким образом, тождество в многообразии $\mathbf{V}$ для диаграммы $(2,1^{n-2})$ может быть записано как

$$\left((-1)^{n-1}(n-2)\alpha_1+\sum_{i=2}^{n-1}(-1)^{n-1-i}2\alpha_i\right)x_0\overline{x}_1\overline{x}_2\dots\overline{x}_n\equiv 0,$$

где сумма коэффициентов в скобках отлична от нуля, так как иначе оно выполнено в многообразии $\mathbf{V}_{\text{anti}}\subset\mathbf{MA}$. То есть для некоторого $n\geqslant 3$ в многообразии $\mathbf{V}$ выполняется тождество $x_0\overline{x}_1\overline{x}_2\dots\overline{x}_n\equiv 0$. Доказательства в случаях $n=1,2$ очевидны. Утверждение доказано. ■

Следствие 1. Многообразие $\mathbf{V}_{\text{anti}}$ является почти нильпотентным.

Доказательство. По утверждению 9 в любом собственном подмногообразии многообразия $\mathbf{V}_{\text{anti}}$ выполнено тождество $x_0\overline{x}_1\overline{x}_2\dots\overline{x}_m\equiv 0$. Подставим в него вместо x_0 произведение x_0y_0 и воспользуемся тождеством (16). Получим $m!x_0y_0x_1\dots x_m\equiv 0$. ■

Главную роль в доказательстве основного результата играет следующая теорема.

Теорема 4 [12]. Пусть действительное число $\alpha>1$ и $\{\lambda^{(n)}\}_{n\geqslant 1}$ — последовательность разбиений, $\lambda^{(n)}\vdash n$, таких, что $\lambda_1^{(n)},\lambda_1'^{(n)}\leqslant n/\alpha$. Тогда для любого действительного β, $1<\beta<\alpha$, найдётся такое натуральное число n_0, что $d_{\lambda^{(n)}}\geqslant\beta^n$ для всех $n\geqslant n_0$.

Сформулируем и докажем основной результат.

Теорема 5. Если рост многообразия $\mathbf{V}\subset\mathbf{MA}$ не выше подэкспоненциального, то или $\mathbf{A}^2\subseteq\mathbf{V}$, или $\mathbf{V}_{\text{anti}}\subseteq\mathbf{V}$, или многообразие $\mathbf{V}$ является нильпотентным.

Доказательство. Пусть многообразие $\mathbf{V}\subset\mathbf{MA}$ не является нильпотентным, $\mathbf{A}^2\not\subseteq\mathbf{V}$ и $\mathbf{V}_{\text{anti}}\not\subseteq\mathbf{V}$. Тогда в силу утверждений 6 и 9 в многообразии $\mathbf{V}$ выполняются тождества $x_0X^k\equiv 0$ и $x_0\overline{x}_1\overline{x}_2\dots\overline{x}_m\equiv 0$ для некоторых $k\geqslant 1$ и $m\geqslant 2$. Так как многообразие $\mathbf{V}$ не является нильпотентным, то ΦS_n-модуль $P_n(\mathbf{V})$, $n\geqslant 1$, содержит ненулевой неприводимый подмодуль $M_{\lambda^{(n)}}$, соответствующий разбиению $\lambda^{(n)}\vdash n$.

Докажем, что последовательность разбиений $\{\lambda^{(n)}\}_{n\geqslant 1}$ удовлетворяет условиям теоремы 4. Для этого рассмотрим общий вид мономов в записи полиоднородного элемента $g_{\lambda^{(n)}}$, линеаризация которого порождает $M_{\lambda^{(n)}}$. Так как $\deg_{x_1} g_{\lambda^{(n)}} = \lambda_1^{(n)}$, то ненулевые мономы могут быть следующих двух видов:

$$y_1X_1^{\gamma_1}y_2X_1^{\gamma_2}\dots y_iX_1^{\gamma_i}\dots y_sX_1^{\gamma_s},\quad x_1X_1^{\gamma_0}y_1X_1^{\gamma_1}\dots y_iX_1^{\gamma_i}\dots y_sX_1^{\gamma_s},$$

где через y_i обозначены необязательно различные образующие x_j, $j>1$; $s=n-\lambda_1^{(n)}$, $\gamma_i\geqslant 0$, $\sum\limits_i \gamma_i=\lambda_1^{(n)}$. В многообразии $\mathbf{V}$ выполнено тождество $x_0X^k\equiv 0$, поэтому $\gamma_i<k$ для всех $0\leqslant i\leqslant s$. Следовательно, для одночленов первого вида имеем неравенство $\lambda_1^{(n)}<ks=k(n-\lambda_1^{(n)})$, откуда $\lambda_1^{(n)}<n/\alpha_1$ для $\alpha_1=(k+1)/k$. Во втором случае для оценки $\lambda_1^{(n)}$ достаточно принять $s\geqslant 1$, тогда $\lambda_1^{(n)}<k(s+1)\leqslant 2ks$. Значит, $\lambda_1^{(n)}<n/\alpha_2$ для $\alpha_2=(2k+1)/(2k)$.

В многочлене $g_{\lambda^{(n)}}$ рассмотрим кососимметрический набор из $\lambda'^{(n)}_1$ различных образующих, где $\lambda'^{(n)}_1$ — длина первого столбца диаграммы $\lambda^{(n)}$. Для удобства обозначим все остальные образующие через y_j, $1\leqslant j\leqslant s$, тогда $g_{\lambda^{(n)}}$ равен линейной комбинации многочленов вида $\overline{x}_1\dots\overline{x}_{\theta_0}y_1\overline{x}_{\theta_0+1}\dots\overline{x}_{\theta_0+\theta_1}y_2\dots y_s\overline{x}_{(\theta_0+\dots+\theta_{s-1}+1)}\dots\overline{x}_{(\theta_0+\dots+\theta_s)}$, где $s=n-\lambda'^{(n)}_1$, $\theta_i\geqslant 0$, $\sum\limits_{i=0}^{s}\theta_i=\lambda'^{(n)}_1$. В силу тождества $x_0\overline{x}_1\overline{x}_2\dots\overline{x}_m\equiv 0$ имеем неравенства $\theta_i<m+1$, $0\leqslant i\leqslant s$. Аналогично рассмотренному случаю достаточно принять $s\geqslant 1$, тогда $\lambda'^{(n)}_1<(s+1)(m+1)<4sm$, поэтому $\lambda'^{(n)}_1<n/\alpha_3$ для $\alpha_3=(4m+1)/(4m)$.

Таким образом, последовательность разбиений $\{\lambda^{(n)}\}_{n\geqslant 1}$ удовлетворяет неравенствам $\lambda_1^{(n)},\lambda'^{(n)}_1\leqslant n/\alpha$ для $\alpha=\min\{\alpha_1,\alpha_2,\alpha_3\}$, $\alpha>1$, и по теореме 4 найдутся такое действительное β, $1<\beta<\alpha$, и натуральное число n_0, что $c_n(\mathbf{V})\geqslant d_{\lambda^{(n)}}\geqslant\beta^n$ для всех $n\geqslant n_0$. Получили противоречие, так как по условию рост многообразия $\mathbf{V}$ не выше подэкспоненциального. ■

Описание всех почти нильпотентных многообразий подэкспоненциального роста в классе антикоммутативных метабелевых алгебр получим в качестве следствия из теоремы 5.

Следствие 2. Пусть многообразие $\mathbf{V}\subset\mathbf{MA}$ является почти нильпотентным подэкспоненциального роста, тогда или $\mathbf{V}=\mathbf{A}^2$, или $\mathbf{V}=\mathbf{V}_{\text{anti}}$.

Авторы выражают благодарность профессору С. П. Мищенко за постоянное внимание к работе, поддержку и ценные замечания.

ЛИТЕРАТУРА

1. *Бахтурин Ю. А.* Тождества в алгебрах Ли. М.: Наука, 1985. 448 с.
2. *Giambruno A. and Zaicev M.* Polynomial Identities and Asymptotic Methods. Providence, RI: AMS, 2005. 352 p.
3. *Фролова Ю. Ю., Шулежко О. В.* Почти нильпотентные многообразия алгебр Лейбница // Прикладная дискретная математика. 2015. № 2(28). С. 30–36.
4. *Mishchenko S. and Valenti A.* An almost nilpotent variety of exponent 2 // Israel J. Mathematics. 2014. V. 199. No. 1. P. 241–257.
5. *Мищенко С. П., Шулежко О. В.* О почти нильпотентных многообразиях в классе коммутативных метабелевых алгебр // Вестник Самарского государственного университета. Естественнонаучная серия. 2015. № 3(125). С. 21–28.
6. *Чанг Н. Т. К., Фролова Ю. Ю.* Почти нильпотентные коммутативные метабелевы многообразия, рост которых не выше экспоненциального // Международная конф. Мальцевские чтения: тез. докл. Новосибирск, 2014. С. 119.

7. *Mishchenko S. and Valenti A.* On almost nilpotent varieties of subexponential growth // J. Algebra. 2015. V. 423. No. 1. P. 902–915.
8. *Шулежко О. В.* О почти нильпотентных многообразиях в различных классах линейных алгебр // Чебышевский сборник. 2015. Т. 16. № 1. С. 67–88.
9. *Мищенко С. П.* Метабелевы почти нильпотентные многообразия полиномиального роста // Материалы Междунар. конф. по алгебре, анализу и геометрии. Казань: Изд-во Академии наук ТР, 2016. С. 247–248.
10. *Зайцев М. В., Мищенко С. П.* О кодлине многообразий линейных алгебр // Математические заметки. 2006. Т. 79. № 4. С. 553–559.
11. *Рацеев С. М.* Рост многообразий алгебр Лейбница с нильпотентным коммутантом // Математические заметки. 2007. Т. 82. № 1. С. 108–117.
12. *Giambruno A. and Mishchenko S.* Degrees of irreducible characters of the symmetric group and exponential growth // Proc. AMS. 2016. V. 144. No. 3. P. 943–953.

REFERENCES

1. *Bakhturin Yu. A.* Tozhdestva v algebrakh Li [Identities of Lie Algebras]. Moscow, Nauka Publ., 1985. 448 p. (in Russian)
2. *Giambruno A. and Zaicev M.* Polynomial Identities and Asymptotic Methods. Providence, RI, AMS, 2005. 352 p.
3. *Frolova Yu. Yu. and Shulezhko O. V.* Pochti nil'potentnye mnogoobraziya algebr Leybnitsa [Almost nilpotent varieties of Leibniz algebras]. Prikladnaya Diskretnaya Matematika, 2015, no. 2(28), pp. 30–36. (in Russian)
4. *Mishchenko S. and Valenti A.* An almost nilpotent variety of exponent 2. Israel J. Mathematics, 2014, vol. 199, no. 1, pp. 241–257.
5. *Mishchenko S. P. and Shulezhko O. V.* O pochti nil'potentnykh mnogoobraziyakh v klasse commutativnykh metabelevykh algebr [On almost nilpotent varieties in the class of commutative metabelian algebras]. Vestnik Samarskogo Gosudarstvennogo Universiteta. Estestvenno-Nauchnaya Seriya, 2015, no. 3(125), pp. 21–28. (in Russian)
6. *Chang N. T. K. and Frolova Yu. Yu.* Pochti nil'potentnye commutativnye metabelevy mnogoobraziya, rost kotorykh ne vyshe eksponentsial'nogo [Almost nilpotent commutative metabelian varieties with not greater than exponential growth rate]. Proc. Intern. Conf. "Mal'tsevskie Chteniya", Novosibirsk, 2014, p. 119. (in Russian)
7. *Mishchenko S. and Valenti A.* On almost nilpotent varieties of subexponential growth. J. Algebra, 2015, vol. 423, no. 1, pp. 902–915.
8. *Shulezhko O. V.* O pochti nil'potentnykh mnogoobraziyakh v razlichnykh klassakh lineynykh algebr [On almost nilpotent varieties in different classes of linear algebras]. Chebyshevskiy Sbornik, 2015, vol. 16, no. 1, pp. 67–88. (in Russian)
9. *Mishchenko S. P.* Metabelevy pochti nil'potentnye mnogoobraziya polinomial'nogo rosta [Almost nilpotent metabelian varieties of polynomial growth]. Proc. Intern. Conf., Kazan, 2016, pp. 247–248. (in Russian)
10. *Zaycev M. V. and Mishchenko S. P.* Colength of varieties of linear algebras. Mathematical Notes, 2006, vol. 79, no. 4, pp. 511–517.
11. *Ratseev S. M.* The growth of varieties of Leibniz algebras with nilpotent commutator subalgebra. Mathematical Notes, 2007, vol. 82, no. 1, pp. 96–103.
12. *Giambruno A. and Mishchenko S.* Degrees of irreducible characters of the symmetric group and exponential growth. Proc. AMS, 2016, vol. 144, no. 3, pp. 943–953.

UDC 512.53 DOI 10.17223/20710410/38/3

ON IRREDUCIBLE ALGEBRAIC SETS OVER LINEARLY ORDERED SEMILATTICES II[1]

A. N. Shevlyakov

Sobolev Institute of Mathematics, Omsk, Russia;
Omsk State Technical University, Omsk, Russia

Equations over finite linearly ordered semilattices are studied. It is assumed that the order of a semilattice is not less than the number of variables in an equation. For any equation $t(X) = s(X)$, we find irreducible components of its solution set. We also compute the average number $\overline{\mathrm{Irr}}(n)$ of irreducible components for all equations in n variables. It turns out that $\overline{\mathrm{Irr}}(n)$ and the function $\frac{4}{9}n!$ are asymptotically equivalent.

Keywords: *irreducible components, algebraic sets, semilattices.*

Introduction

This paper is the sequel of [1], and we recall below the general problems studied in both papers.

Following [2], one can define a notion of an equation over a linearly ordered semilattice $L_l = \{a_1, a_2, \ldots, a_l\}$ (the formal definition of an equation is given below in the paper). A set Y is *algebraic* if it is the solution set for a system of equations over L_l. Let us consider an equation $t(X) = s(X)$ in n variables over L_l, and let Y be the solution set for $t(X) = s(X)$. One can find algebraic sets $Y_1, Y_2, \ldots, Y_m$ such that $Y = \bigcup\limits_{i=1}^{m} Y_i$. One can decompose each Y_i into a union of other algebraic sets, etc. This process terminates after a finite number of steps and gives a decomposition of Y into a union of *irreducible* algebraic sets Y_i (the sets Y_i are called the *irreducible components* of Y). Roughly speaking, irreducible algebraic sets are "atoms" which form any algebraic set. The size and the number of such "atoms" are important characteristics of the semilattice L_l, since there are connections between irreducible algebraic sets and universal theory of linearly ordered semilattices [2]. Moreover, the number of irreducible components was involved in the estimation of lower bounds of algorithm complexity (see [3] for more details).

In the previous paper [1], we studied equations $t(X) = s(X)$ with $n > l$, i.e. the number of variables occurring in $t(X) = s(X)$ is more than the order of the semilattice L_l. In [1], we also studied algebraic sets and irreducible components and computed the average number of irreducible components of the solution sets for equations in n variables.

In this paper, we assume $n \leqslant l$ (i.e. the order of the semilattice L_l is not less than the number of variables in $t(X) = s(X)$) and study the similar problems. Precisely, for any equation $t(X) = s(X)$ in n variables, we study the number and properties of its solution set irreducible components, and for all equations in n variables, we count the average number $\overline{\mathrm{Irr}}(n)$ of irreducible components of the solution sets.

Note that the cases $n > l$ and $n \leqslant l$ need a completely different techniques, and we can not directly use the results of [1] in the current paper. Moreover, almost all the results of [1] do not hold for the current case.

[1]The author was supported by the RSF-grant 17-11-01117.

1. Main definitions

Let $L_l = \{a_1, a_2, \ldots, a_l\}$ be the linearly ordered semilattice of l elements and $a_1 < a_2 < \ldots < a_l$. The multiplication in L_l is defined by $a_i \cdot a_j = a_{\min(i,j)}$. Obviously, the linear order on L_l can be expressed by the multiplication as follows

$$a_i \leqslant a_j \Leftrightarrow a_i a_j = a_i.$$

A *term* $t(X)$ *in variables from* $X = \{x_1, x_2, \ldots, x_n\}$ is a commutative word in letters x_i.

Let $\mathrm{Var}(t)$ be the set of all variables occurring in a term $t(X)$. Following [2], an *equation* is an equality of some terms $t(X) = s(X)$. Below we consider inequalities $t(X) \leqslant s(X)$ as equations, since $t(X) \leqslant s(X)$ is the short form of $t(X)s(X) = t(X)$. Notice that we consider equations as *ordered pairs* of terms, i.e. the expressions $t(X) = s(X)$ and $s(X) = t(X)$ are *different* equations. Let $Eq(n)$ denote the set of all equations in variables from $X = \{x_1, x_2, \ldots, x_n\}$. We assume that each equation $t(X) = s(X)$ in $Eq(n)$ contains the occurrences of all variables $x_1, x_2, \ldots, x_n$. An equation $t(X) = s(X)$ in $Eq(n)$ is said to be a (k_1, k_2)-*equation* if $|\mathrm{Var}(t) \setminus \mathrm{Var}(s)| = k_1$ and $|\mathrm{Var}(s) \setminus \mathrm{Var}(t)| = k_2$. For example, $x_1x_2 = x_1x_3x_4$ is a $(1, 2)$-equation. Let $Eq(k_1, k_2, n)$ be the set of all (k_1, k_2)-equations in $Eq(n)$. Obviously,

$$Eq(n) = \bigcup_{(k_1,k_2)\in K_n} Eq(k_1, k_2, n), \tag{1}$$

where $K_n = \{(k_1, k_2) : k_1 + k_2 \leqslant n\} \setminus \{(0, n), (n, 0)\}$.

Each equation $t(X) = s(X)$ in $Eq(k_1, k_2, n)$ is uniquely defined by k_1 variables in the left part and by k_2 other variables in the right part (the other $n - k_1 - k_2$ variables should occur in both parts of the equation). Thus,

$$\#Eq(k_1, k_2, n) = \binom{n}{k_1}\binom{n-k_1}{k_2}.$$

By (1), one can compute that $\#Eq(n) = 3^n - 2$.

Remark 1. Recall that we consider only equations $t(X) = s(X)$ with $n \leqslant l$, i.e. the number of variables occurring in $t(X) = s(X)$ is not more than the order of the semilattice L_l.

A point $P \in L_l^n$ is a *solution* of an equation $t(X) = s(X)$ if $t(P)$ and $s(P)$ define the same element in the semilattice L_l. By the properties of linearly ordered semilattices, a point $P = (p_1, p_2, \ldots, p_n)$ is a solution of $t(X) = s(X)$ iff there exist variables x_i in $\mathrm{Var}(t)$ and x_j in $\mathrm{Var}(s)$ such that $p_i = p_j$ and $p_i \leqslant p_k$ for all k, $1 \leqslant k \leqslant n$. The set of all solutions of an equation $t(X) = s(X)$ is denoted by $\mathrm{V}(t(X) = s(X))$.

An arbitrary set of equations is called a *system*. The set $\mathrm{V}(\mathbf{S})$ of all solutions of a system $\mathbf{S} = \{t_i(X) = s_i(X) : i \in I\}$ is defined as $\bigcap\limits_{i\in I} \mathrm{V}(t_i(X) = s_i(X))$. A subset Y of the set L_l^n is called *algebraic over* L_l if there exists a system $\mathbf{S}$ in n variables with $\mathrm{V}(\mathbf{S}) = Y$. An algebraic set Y is *irreducible* if Y is not a proper finite union of other algebraic sets.

Proposition 1 [1, Proposition 2.2]. Any algebraic set Y over L_l is a finite union of irreducible sets, that is,

$$Y = Y_1 \cup Y_2 \cup \ldots \cup Y_m, \tag{2}$$

where $Y_i \not\subseteq Y_j$ for all i and j such that $i \neq j$, and this decomposition is unique up to a permutation of components.

The subsets Y_i from the union (2) are called the *irreducible components* of Y.

Let Y be an algebraic set over L_l defined by a system $\mathbf{S}(X)$. One can define an equivalence relation $\sim_Y$ over the set of all terms in variables X as follows

$$t(X) \sim_Y s(X) \Leftrightarrow t(P) = s(P) \text{ for any point } P \in Y.$$

The set of all $\sim_Y$-equivalence classes is called *the coordinate semilattice of* Y and denoted by $\Gamma(Y)$ (see [2] for more details). The following statement describes the coordinate semilattices of irreducible algebraic sets.

Proposition 2 [1, Proposition 2.3]. A set Y is irreducible over L_l iff $\Gamma(Y)$ is embedded into L_l.

There are different algebraic sets over L_l with isomorphic coordinate semilattices. Such sets are called *isomorphic*. For example, the following sets

$$Y_1 = \mathrm{V}(\{x_1 \leqslant x_2 \leqslant x_3\}),\ Y_2 = \mathrm{V}(\{x_3 \leqslant x_2 \leqslant x_1\})$$

have the isomorphic coordinate semilattices

$$\Gamma(Y_1) = \langle x_1, x_2, x_3 \mid x_1 \leqslant x_2 \leqslant x_3 \rangle \cong L_3,$$
$$\Gamma(Y_2) = \langle x_1, x_2, x_3 \mid x_3 \leqslant x_2 \leqslant x_1 \rangle \cong L_3.$$

Thus, Y_1 and Y_2 are isomorphic.

2. Example

Let $n = 3$, $l = 3$. We have exactly $Eq(3) = 3^3 - 2 = 25$ equations in three variables over L_3. The Table on the page 52 contains the information about such equations over L_3. The second column contains systems which define irreducible components of the solution set for an equation in the first column. A cell of the table contains $\uparrow$ if an information in this cell is similar to the cell above.

Notice that $\mathrm{V}(x_1 = x_2 \leqslant x_3)$ does not define an irreducible component for $Y = \mathrm{V}(x_1x_2 = = x_1x_3)$, since $\mathrm{V}(x_1 = x_2 \leqslant x_3)$ is included into the solution set of another irreducible component $\mathrm{V}(x_1 \leqslant x_2 \leqslant x_3)$. Similarly, $\mathrm{V}(x_3 = x_1 \leqslant x_2)$ is not an irreducible component for Y, since it is contained in the irreducible component $\mathrm{V}(x_1 \leqslant x_3 \leqslant x_2)$.

It turns out that the number of irreducible components does not depend on the semilattice order l. One can directly compute the average number of irreducible components of algebraic sets defined by equations in three variables:

$$\overline{\mathrm{Irr}}(3) = \frac{6 + 2(2+2+2+2+2+2+3+3+3+4+4+4)}{25} = \frac{72}{25} = 2.88. \qquad (3)$$

Equations	Irreducible components (IC)	Number of IC
$x_1x_2x_3 = x_1x_2x_3$	$x_1 \leqslant x_2 \leqslant x_3 \cup x_1 \leqslant x_3 \leqslant x_2 \cup$ $\cup x_2 \leqslant x_1 \leqslant x_3 \cup x_2 \leqslant x_3 \leqslant x_2 \cup$ $\cup x_3 \leqslant x_1 \leqslant x_2 \cup x_3 \leqslant x_2 \leqslant x_1$	6
$x_1 = x_1x_2x_3,$ $x_1x_2x_3 = x_1$	$x_1 \leqslant x_2 \leqslant x_3 \cup x_1 \leqslant x_3 \leqslant x_1$	2
$x_2 = x_1x_2x_3,$ $x_1x_2x_3 = x_2$	$\uparrow$	2
$x_3 = x_1x_2x_3,$ $x_1x_2x_3 = x_3$	$\uparrow$	2
$x_1 = x_2x_3,$ $x_2x_3 = x_1$	$x_1 = x_2 \leqslant x_3 \cup x_1 = x_3 \leqslant x_2$	2
$x_2 = x_1x_3,$ $x_1x_3 = x_2$	$\uparrow$	2
$x_3 = x_1x_2,$ $x_1x_2 = x_3$	$\uparrow$	2
$x_1x_2 = x_1x_3,$ $x_1x_3 = x_1x_2$	$x_1 \leqslant x_2 \leqslant x_3 \cup x_1 \leqslant x_3 \leqslant x_2 \cup$ $\cup x_2 = x_3 \leqslant x_1$	3
$x_1x_2 = x_2x_3,$ $x_2x_3 = x_1x_2$	$\uparrow$	3
$x_1x_3 = x_2x_3,$ $x_2x_3 = x_1x_3$	$\uparrow$	3
$x_1x_2 = x_1x_2x_3,$ $x_1x_2x_3 = x_1x_2$	$x_1 \leqslant x_2 \leqslant x_3 \cup x_1 \leqslant x_3 \leqslant x_2 \cup$ $\cup x_2 \leqslant x_1 \leqslant x_3 \cup x_2 \leqslant x_3 \leqslant x_1$	4
$x_1x_3 = x_1x_2x_3,$ $x_1x_2x_3 = x_1x_3$	$\uparrow$	4
$x_2x_3 = x_1x_2x_3,$ $x_1x_2x_3 = x_2x_3$	$\uparrow$	4

3. Decompositions of algebraic sets

Let Y denote the solution set for an equation $t(X) = s(X)$ over the semilattice $L_l = \{a_1, a_2, \ldots, a_l\}$. The table on the page 52 shows that any irreducible component sorts the variables X into some order. The following definition formalizes this property of irreducible components.

Let σ be a permutation of the set $\{1, 2, \ldots, n\}$; σ sorts the set X as follows: $\{x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\}$, i.e. $\sigma(i)$ is the i-th variable in the sorted set X. A permutation σ is called a *permutation of the first* (*second*) *kind* if $x_{\sigma(1)} \in \mathrm{Var}(t) \cap \mathrm{Var}(s)$ (respectively, $x_{\sigma(2)} \in \in \mathrm{Var}(t) \setminus \mathrm{Var}(s)$, $x_{\sigma(1)} \in \mathrm{Var}(s) \setminus \mathrm{Var}(t)$). Let $\chi(\sigma) \in \{1, 2\}$ denote the kind of a permutation σ.

Example 1. Let us consider an algebraic set $Y_0 = \mathrm{V}(x_1x_2 = x_1x_3)$. By the table, Y_0 is the union of the following irreducible components:

$$Y_1 = \mathrm{V}(x_1 \leqslant x_2 \leqslant x_3),\ Y_2 = \mathrm{V}(x_1 \leqslant x_3 \leqslant x_2),\ Y_3 = \mathrm{V}(x_2 = x_3 \leqslant x_1).$$

The irreducible components Y_1, Y_2, Y_3 define the following permutations:

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Moreover, σ_1 and σ_2 are permutations of the first kind, whereas σ_3 is of the second kind.

A permutation σ defines an algebraic set Y_σ as follows:

$$Y_\sigma = \mathrm{V}\left(\bigcup_{i=1}^{n-1} \{x_{\sigma(i)} \leqslant x_{\sigma(i+1)}\}\right) \tag{4}$$

if $\chi(\sigma) = 1$, and

$$Y_\sigma = \mathrm{V}\left(\{x_{\sigma(1)} = x_{\sigma(2)}\} \bigcup_{i=2}^{n-1} \{x_{\sigma(i)} \leqslant x_{\sigma(i+1)}\}\right) \tag{5}$$

if $\chi(\sigma) = 2$.

Example 2. Let $\sigma_1, \sigma_2, \sigma_3$ be permutations from Example 1. Obviously, the sets $Y_{\sigma_1}, Y_{\sigma_2}, Y_{\sigma_3}$ defined by (4) and (5) coincide with the sets Y_1, Y_2, Y_3 respectively.

Lemma 1. Let $\chi(\sigma) \in \{1, 2\}$, then the set Y_σ is irreducible and, moreover,

$$\Gamma(Y_\sigma) \cong \begin{cases} L_n, \text{ if } \chi(\sigma) = 1, \\ L_{n-1}, \text{ if } \chi(\sigma) = 2. \end{cases} \tag{6}$$

Proof. By the definition of a coordinate semilattice, $\Gamma(Y_\sigma)$ is generated by the elements of $\{x_1, x_2, \ldots, x_n\}$ and has the following defined relations:

$$x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant \ldots x_{\sigma(n)} \text{ if } \chi(Y_\sigma) = 1$$

and

$$x_{\sigma(1)} = x_{\sigma(2)} \leqslant \ldots x_{\sigma(n)} \text{ if } \chi(Y_\sigma) = 2.$$

Thus, $\Gamma(Y_\sigma)$ is a linearly ordered semilattice, and (6) holds. By Proposition 2, the set Y_σ is irreducible. ■

The following lemma gives the irreducible decomposition of an algebraic set $Y = \mathrm{V}(t(X) = s(X))$.

Lemma 2. An algebraic set $Y = \mathrm{V}(t(X) = s(X))$ is a union

$$Y = \bigcup_{\chi(\sigma) \in \{1,2\}} Y_\sigma. \tag{7}$$

Proof. Suppose $P = (p_1, p_2, \ldots, p_n) \in Y$. Let us sort p_i in the ascending order

$$p_{\sigma(1)} \leqslant p_{\sigma(2)} \leqslant \ldots \leqslant p_{\sigma(n)},$$

where σ is a permutation of the set $\{1, 2, \ldots, n\}$. We have that σ induces the sorting of the variable set X. Obviously, we may assume that $x_\sigma(1) \in \mathrm{Var}(t)$, otherwise the properties of L_l provide the existence of a variable $x_\sigma(i) \in \mathrm{Var}(t)$ such that $p_{\sigma(i)} = p_{\sigma(1)}$, and we can swap the values $\sigma(1)$ and $\sigma(i)$.

For example, the point $P = (a_2, a_1, a_1) \in \mathrm{V}(x_1x_2 = x_1x_3)$ defines $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$ (the permutation obtained equals σ_3 from Example 1, so the point (a_2, a_1, a_1) belongs to the set Y_3).

Since σ is defined by the inequalities between the coordinates p_i, it follows $P \in Y_\sigma$.

Now we prove that $Y_\sigma \subseteq Y$ for each σ. Suppose $P = (p_1, p_2, \ldots, p_n) \in Y_\sigma$. If $\chi(Y_\sigma) = 1$, then

$$x_{\sigma(1)} \in \mathrm{Var}(t) \cap \mathrm{Var}(s) \Rightarrow t(P) = s(P) = p_{\sigma(1)} \Rightarrow P \in \mathrm{V}(t(X) = s(X)).$$

Otherwise ($\chi(Y_\sigma) = 2$), $t(P) = p_{\sigma(1)}$, $s(P) = p_{\sigma(2)}$, and (5) gives $p_{\sigma(1)} = p_{\sigma(2)}$. Therefore $P \in \mathrm{V}(t(X) = s(X))$. ■

Lemma 3. For distinct permutations σ and σ', we have $Y_\sigma \nsubseteq Y_{\sigma'}$ in (7).

Proof. Let σ be a permutation of the first or second kind, and P_σ denote the following point:

$$p_{\sigma(i)} = a_i \text{ if } \chi(\sigma) = 1,$$

and

$$p_{\sigma(i)} = \begin{cases} a_i, \ 2 \leqslant i \leqslant n, \\ a_2, \ i = 1, \end{cases} \qquad \text{if } \chi(\sigma) = 2.$$

For example, the permutations $\sigma_1, \sigma_2, \sigma_3$ from Example 1 define the points $P_1 = (a_1, a_2, a_3)$, $P_2 = (a_1, a_3, a_2)$, $P_3 = (a_3, a_2, a_2)$, respectively.

Since P_σ preserves the order of variables, we have $P_\sigma \in Y_\sigma$.

Now we can show that $P_\sigma \notin Y_{\sigma'}$ for every $\sigma' \neq \sigma$ (for example, each of the points P_1, P_2, P_3 above belongs to a unique irreducible component from Example 1:

$$P_1 \in Y_1 \setminus (Y_2 \cup Y_3), \ P_2 \in Y_2 \setminus (Y_1 \cup Y_3), \ P_3 \in Y_3 \setminus (Y_1 \cup Y_2)).$$

There exist numbers i and j such that $i < j$, $i = \sigma(\alpha)$, $j = \sigma(\beta)$ with $\alpha < \beta$ and $i = \sigma'(\alpha')$, $j = \sigma'(\beta')$ with $\alpha' > \beta'$. Hence, the inequality $x_i \leqslant x_j$ holds in Y_σ, and the inequality $x_j \leqslant x_i$ holds in $Y_{\sigma'}$. Let us consider the two possible cases:

1) If $\chi(\sigma) = 1$, then $p_i < p_j$ in P_σ, and we immediately obtain $P_\sigma \notin Y_{\sigma'}$.
2) Suppose $\chi(\sigma) = 2$. Assume that $p_i = p_j = a_2$ (if $p_i < p_j$, we immediately obtain $P_\sigma \notin Y_{\sigma'}$). Then $\alpha = 1$, $\beta = 2$ and $i = \sigma(1)$, $j = \sigma(2)$ (one can similarly consider the case $i = \sigma(2)$, $j = \sigma(1)$). Hence, $x_i \in \mathrm{Var}(t) \setminus \mathrm{Var}(s)$, $x_j \in \mathrm{Var}(s) \setminus \mathrm{Var}(t)$. By the definition of a permutation of the second kind, $\sigma'(1) = k \neq j$, and the inequality $x_k \leqslant x_j$ holds in $Y_{\sigma'}$. Let $\sigma(\gamma) = k$. Since $\alpha = 1$, $\beta = 2$, we have $\gamma > 2$. Then $p_k = a_\gamma$ and $p_j < p_k$ for P_σ. Thus, $P \notin Y_{\sigma'}$.

The Lemma 3 is proved. ■

According to Lemmas 1–3, we obtain the following statement.

Theorem 1. The union (7) is the irreducible decomposition of the set $Y = \mathrm{V}(t(X) = s(X))$. The number of irreducible components is equal to the number of permutations of the first and second kind.

4. Average number of irreducible components

One can directly compute that any (k_1, k_2)-equation admits

$$(n - k_1 - k_2)(n-1)!$$

permutations of the first kind and

$$k_1 k_2 (n-2)!$$

permutations of the second kind.

By Theorem 1, for a (k_1, k_2)-equation $t(X) = s(X)$ the number of its irreducible components equals

$$\mathrm{Irr}(k_1, k_2, n) = (n - k_1 - k_2)(n-1)! + k_1 k_2 (n-2)!$$

The average number of irreducible components of algebraic sets defined by equations from $Eq(n)$ is

$$\overline{\mathrm{Irr}}(n) = \frac{\sum\limits_{(k_1,k_2)\in K_n} \#Eq(k_1,k_2,n)\mathrm{Irr}(k_1,k_2,n)}{\#Eq(n)} =$$

$$= \frac{\sum\limits_{k_1=0}^{n-1}\sum\limits_{k_2=0}^{n-k_1} \#Eq(k_1,k_2,n)\mathrm{Irr}(k_1,k_2,n) - \#Eq(0,n,n)\mathrm{Irr}(0,n,n)}{\#Eq(n)}.$$

Since $\mathrm{Irr}(0,n,n) = (n-0-n)(n-1)! + 0n(n-2)! = 0$, we obtain

$$\overline{\mathrm{Irr}}(n) = \frac{\sum\limits_{k_1=0}^{n-1}\sum\limits_{k_2=0}^{n-k_1} \#Eq(k_1,k_2,n)\mathrm{Irr}(k_1,k_2,n)}{\#Eq(n)}.$$

Below we compute $\overline{\mathrm{Irr}}$ using the following notation:

1) $A \overset{(1)}{=} B$: an expression B is obtained from A by the binomial identity

$$a\binom{n}{a} = n\binom{n-1}{a-1};$$

2) $A \overset{(2)}{=} B$: an expression B is obtained from A by the following identity of binomial coefficients

$$\sum_{t=0}^{n}\binom{n}{t}t2^t = 2n3^{n-1}. \tag{8}$$

Here is a proof of (8):

$$\sum_{t=0}^{n}\binom{n}{t}t2^t \overset{(1)}{=} n\sum_{t=0}^{n}\binom{n-1}{t-1}2^t = 2n\sum_{t=0}^{n}\binom{n-1}{t-1}2^{t-1} = 2n\sum_{u=0}^{n-1}\binom{n-1}{u}2^u = 2n3^{n-1}.$$

Let us compute $\overline{\mathrm{Irr}}(n)$. We have that

$$\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1} \#Eq(k_1,k_2,n)\mathrm{Irr}(k_1,k_2,n) =$$

$$= \sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}(n-k_1-k_2)(n-1)! + k_1k_2(n-2)!) =$$

$$= n!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2} - (n-1)!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}k_1 -$$

$$-(n-1)!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}k_2 + (n-2)!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}k_1k_2 = S_1 - S_2 - S_3 + S_4,$$

where

$$S_1 = n!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2} = n!\sum_{k_1=0}^{n-1}\binom{n}{k_1}2^{n-k_1} = n!(3^n-1),$$

$$S_2 = (n-1)!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}k_1 = (n-1)!\sum_{k_1=0}^{n-1}\binom{n}{k_1}k_12^{n-k_1} \overset{(1)}{=}$$

$$\overset{(1)}{=} n!\sum_{k_1=0}^{n-1}\binom{n-1}{k_1-1}2^{n-k_1} = n!\sum_{t=0}^{n-2}\binom{n-1}{t}2^{n-1-t} = n!\left(\sum_{t=0}^{n-1}\binom{n-1}{t}2^{n-1-t}-1\right) = n!(3^{n-1}-1),$$

$$S_3 = (n-1)!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}k_2 \overset{(1)}{=} (n-1)!\sum_{k_1=0}^{n-1}\binom{n}{k_1}(n-k_1)\sum_{k_2=0}^{n-k_1}\binom{n-k_1-1}{k_2-1} =$$
$$= (n-1)!\sum_{k_1=0}^{n-1}\binom{n}{k_1}(n-k_1)2^{n-k_1-1} = (n-1)!\sum_{t=0}^{n}\binom{n}{t}t2^{t-1} = \frac{(n-1)!}{2}\sum_{t=0}^{n}\binom{n}{t}t2^t \overset{(2)}{=} n!3^{n-1},$$

$$S_4 = (n-2)!\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-k_1}\binom{n}{k_1}\binom{n-k_1}{k_2}k_1k_2 \overset{(1)}{=} (n-2)!\sum_{k_1=0}^{n-1}\binom{n}{k_1}k_1(n-k_1)\sum_{k_2=0}^{n-k_1}\binom{n-k_1-1}{k_2-1} =$$
$$= (n-2)!\sum_{k_1=0}^{n-1}\binom{n}{k_1}k_1(n-k_1)2^{n-k_1-1} =$$
$$= \frac{(n-2)!}{2}\sum_{k_1=0}^{n}\binom{n}{k_1}k_1(n-k_1)2^{n-k_1} = \frac{(n-2)!}{2}\sum_{t=0}^{n}\binom{n}{t}t(n-t)2^t =$$
$$= \frac{(n-2)!}{2}\left(n\sum_{t=0}^{n}\binom{n}{k_1}t2^t - \sum_{t=0}^{n}\binom{n}{t}t^22^t\right) \overset{(2)}{=} \frac{(n-2)!}{2}\left(2n^23^{n-1}-S_5\right),$$

$$S_5 = \sum_{t=0}^{n}\binom{n}{k_1}t^22^t \overset{(1)}{=} n\sum_{t=0}^{n}\binom{n-1}{t-1}t2^t = n\left(\sum_{t=0}^{n}\binom{n-1}{t-1}(t-1)2^t + \sum_{t=0}^{n}\binom{n-1}{t-1}2^t\right) =$$
$$= n\left(2\sum_{t=0}^{n}\binom{n-1}{t-1}(t-1)2^{t-1} + \sum_{t=0}^{n}\binom{n-1}{t-1}2^t\right) \overset{(2)}{=} n\left(4(n-1)3^{n-2}+2\cdot 3^{n-1}\right).$$

Finally, we obtain that

$$S_1 - S_2 - S_3 + S_4 = n!(3^n-1) - n!(3^{n-1}-1) - n!3^{n-1} +$$
$$+\frac{(n-2)!}{2}\left(2n^23^{n-1} - n(4(n-1)3^{n-2}+2\cdot 3^{n-1})\right) =$$
$$= n!3^{n-1} + (n-2)!3^{n-2}n\left(3n-2(n-1)-3\right) = n!3^{n-1}+n!3^{n-2} = 4n!3^{n-2},$$

and

$$\overline{\mathrm{Irr}}(n) = \frac{4n!3^{n-2}}{3^n-2} \sim \frac{4}{9}n! \tag{9}$$

Notice that the final answer does not depend on l if $l \leqslant n$. In particular, (9) gives

$$\overline{\mathrm{Irr}}(3) = \frac{72}{25} = 2.88 \tag{10}$$

for $n = 3$, and (10) coincides with (3).

REFERENCES

1. *Shevlyakov A. N.* On irreducible algebraic sets over linearly ordered semilattices. Groups, Complexity, Cryptology, 2016, vol. 8, no. 2, pp. 187–196.
2. *Daniyarova E. Yu., Myasnikov A. G., and Remeslennikov V. N.* Algebraicheskaya geometriya nad algebraicheskimi sistemami [Algebraic Geometry over Algebraic Systems]. Novosibirsk, SB RAS Publ., 2016. 243 p. (in Russian)
3. *Ben-Or M.* Lower bounds for algebraic computation trees. 15th Ann. Symp. Theory Computing, 1983, pp. 80–86.

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

UDC 519.7 DOI 10.17223/20710410/38/4

SUBSTITUTION BLOCK CIPHERS WITH FUNCTIONAL KEYS[1]

G. P. Agibalov

National Research Tomsk State University, Tomsk, Russia

We define a substitution block cipher $\mathcal{C}$ with the plaintext and ciphertext blocks in $\mathbb{F}_2^n$ and with the keyspace $K_{s_0,n}(g)$ that is the set $\{f(x) : f(x) = \pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$; $\sigma_1, \sigma_2 \in \mathbb{F}_2^n; \pi_1, \pi_2 \in S_n\}$, where s_0 is an integer, $1 \leqslant s_0 \leqslant n$; $g : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a bijective vector function $g(x) = g_1(x)g_2(x)\dots g_n(x)$ such that every its coordinate function $g_i(x)$ essentially depends on some $s_i \leqslant s_0$ variables in the string $x = x_1x_2\dots x_n$; S_n is the set of all permutations of the row $(1\ 2\ \dots\ n)$; π_i and σ_i are the permutation and negation operations, that is, $(\pi = (i_1i_2\dots i_n)) \Rightarrow (\pi(a_1a_2\dots a_n) = a_{i_1}a_{i_2}\dots a_{i_n})$, $(\sigma = b_1b_2\dots b_n) \Rightarrow ((a_1a_2\dots a_n)^\sigma = a_1^{b_1}a_2^{b_2}\dots a_n^{b_n})$ and, for a and b in $\mathbb{F}_2$, $a^b = a$ if $b = 1$ and $a^b = \neg a$ if $b = 0$. Like g, any key f in $K_{s_0,n}(g)$ is a bijection on $\mathbb{F}_2^n$, $f(x) = f_1(x)f_2(x)\dots f_n(x)$, and every its coordinate function $f_i(x)$ essentially depends on not more than s_0 variables in x. The encryption of a plaintext block x and the decryption of a ciphertext block y on the key f are defined in $\mathcal{C}$ as follows: $y = f(x)$ and $x = f^{-1}(y)$. Here, we suggest a known plaintext attack on $\mathcal{C}$ with the threat of discovering the key f that was used. Let $P_1, P_2, \dots, P_m$ be some blocks of a plaintext, $C_1, C_2, \dots, C_m$ be the corresponding blocks of a ciphertext, i.e., $C_l = f(P_l)$ for $l = 1, 2, \dots, m$, and $P_l = P_{l1}P_{l2}\dots P_{ln}$, $C_l = C_{l1}C_{l2}\dots C_{ln}$. The object is to determine the coordinate function $f_i(x)$ of f for each $i \in \{1, 2, \dots, n\}$. The suggested attack consists of two steps, namely we first determine the essential variables $x_{i_1}, \dots, x_{i_s}$ of $f_i(x)$ and then compute a Boolean function $h(x_{i_1}, \dots, x_{i_s})$ such that $h(a_{i_1}, \dots, a_{i_s}) = f_i(a_1, \dots, a_n)$ for all n-tuples $(a_1a_2\dots a_n) \in \mathbb{F}_2^n$. For determining the essential variables of f_i, we construct a Boolean matrix $||\inf D(f_i)||$ with the set of rows $\inf D(f_i)$, where $D(f_i) = \{P_l \oplus P_j : C_{li} \neq C_{ji};\ l, j = 1, 2, \dots, m\}$, $C_{li} = f_i(P_l)$, $l = 1, \dots, m$, $i = 1, \dots, n$, and $\inf D(f_i)$ is the subset of all the minimal vectors in $D(f_i)$. Then the numbers of essential variables for f_i are the numbers of columns in the intersection of all covers of $||\inf D(f_i)||$ with the cardinalities not more than s_0, where a cover of a Boolean matrix M is defined as a subset C of its columns such that each row in M has '1' in a column in C. For computing $h(x_{i_1}, \dots, x_{i_s})$, we first set $h(P_{li_1}, \dots, P_{li_s}) = C_{li}$ for $l = 1, \dots, m$ and then, if h_i is not yet completely determined on $\mathbb{F}_2^s$, we increase the number m of known blocks (P_i, C_i) of plain- and ciphertexts or extend h_i on $\mathbb{F}_2^s$ in such a way that the vector function $h = h_1h_2\dots h_n$ with the completely defined coordinate functions is a bijection on $\mathbb{F}_2^n$. We also describe some special known plaintext attacks on substitution block ciphers with keyspaces being subsets of $K_{s_0,n}(g)$.

Keywords: *substitution ciphers, block ciphers, functional keys, cryptanalysis, known plaintext attack, Boolean functions, essential variables, bijective functions.*

[1]The author was supported by the RFBR-grant no. 17-01-00354.

1. Introduction

In cryptography, the cryptosystems with the functional keys are widely used as cryptographic primitives including key-stream generators, s-boxes, cryptofilters, cryptocombiners, key hash functions as well as the symmetric and public-key ciphers, digital signature schemes. For the author, the research, including the definition, characterisation and cryptanalysis of such cryptosystems had beginnings at the 1960-th years. First object of this research was the key-stream generator based on a finite autonomous automaton (state machine) with the output function depending on a bounded number of coordinates of the automaton state and being the key of the generator and of the corresponding stream cipher [1, 2]. Later, two sets of symmetric iterative block ciphers with the functional keys were proposed [3]. They were constructed according to the known cryptographic schemes originally suggested by H. Feistel and implemented in the ciphers LUCIFER and DES and, therefore, were named after Lucifer and Feistel respectively. At the last years, our research in this area was related to definitions and cryptanalysis and synthesis methods for some other kinds of cryptalgorithms with functional keys including watermarking ciphers [4], finite automata cryptographic generators with two-valued controlled steps [5], and cryptautomata [6] where a cryptautomaton is described by a set C of automata networks and a set K of keys such that the choosing a key in K determines a network in C as a specific cryptographic algorithm. In the case, when the key contains transition and (or) output functions of some components in networks in C, we have a cryptosystem with the functional keys. In this paper, we describe another class of cryptosystems with functional keys, namely that named in the title.

2. Definition

Here is a general formal mathematical definition of the ciphers under consideration. Let $\mathcal{C}$ be a symmetric cipher and $\mathcal{C} = (X, Y, K, E, D)$, where X, Y, and K are the sets of plaintexts, ciphertexts and keys respectively and E and D are, respectively, the encryption and decryption algorithms, $E : X \times K \to Y$, $D : Y \times K \to X$ and $E(x, k) = y \Rightarrow D(y, k) = x$ for any $x \in X$, $y \in Y$, and $k \in K$. Suppose $X = Y = \mathbb{F}_2^n$, $K \subseteq B^n$, B is a class of Boolean functions having some bounded both computational and capacity complexities and depending on not more than n variables such that the mapping $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$, defined for $x \in X$ and $f_1 f_2 \dots f_n \in K$ as $f(x) = f_1(x) f_2(x) \dots f_n(x)$, is a bijection. In this case, the cipher $\mathcal{C}$ is said to be a *substitution block cipher with functional keys*, or, shortly, a *funkeysubcipher*. For each block $x = x_1 x_2 \dots x_n \in X$, for each key $k = f_1 f_2 \dots f_n \in K$, and for each ciphertext $y \in Y$ in it, we have $E(x, k) = f(x)$ and $D(y, k) = f^{-1}(y)$. Further, these equalities are called the *invertibility condition* of $\mathcal{C}$.

Note that in this definition, the bounded complexity of a function means the existence of its practical specification and computation.

As usually, there are two general problems in the funkeysubcipher theory — synthesis and analysis. The second problem is very typical of block ciphers and its solving ways significantly depend on the way the first problem is solved. According to the definition above, the first problem consists in generating a proper key space K, namely which is over a set B of Boolean functions of a bounded complexity, satisfies the invertibility conditions, and is great enough to withstand exhaustive search attacks. A method for solving this problem is described in the following section.

3. Synthesis method

Let IS_n denote the set of all invertible systems each consisting of n functions in B. Further, we also consider the systems in IS_n as Boolean bijective vector functions, that is, as substitutions on $\mathbb{F}_2^n$. To synthesize a funkeysubcipher $\mathcal{C} = (\mathbb{F}_2^n, K, \mathbb{F}_2^n, E, D)$, where $K \subseteq IS_n$, we need generating the vector functions in IS_n as keys in K. Without knowing how to generate all of them, we propose here to generate keys in K as some members of IS_n which can be obtained by inverse and permutation operations over bits on inputs and outputs of a chosen or given function in IS_n. For this purpose, we, first, introduce some auxiliary notations related to the permutation and inverse operations and, then, define some subsets of functions in B^n.

Let S_n be the set of all the permutations of numbers 1, 2, ..., n, namely $S_n = \{i_1 i_2 \dots i_n : i_j \in \{1, 2, \dots, n\}; j \neq k \Rightarrow i_j \neq i_k; j, k \in \{1, 2, \dots, n\}\}$. For any permutation $\pi = i_1 i_2 \dots i_n \in S_n$ and any vector $v = v_1 v_2 \dots v_n$, let $\pi(v_j) = v_{i_j}$, $j = 1, 2, \dots, n$, and $\pi(v) = \pi(v_1)\pi(v_2)\dots\pi(v_n) = v_{i_1} v_{i_2} \dots v_{i_n}$. Also, if $v_1, v_2, \dots, v_n$ are Boolean values (variables or constants) and $\sigma = b_1 b_2 \dots b_n \in \mathbb{F}_2^n$, then let $v^\sigma = v_1^{b_1} v_2^{b_2} \dots v_n^{b_n}$, where, for any Boolean values a and b, $a^b = \neg a$ if $b = 0$ and $a^b = a$ if $b = 1$. We say that $\pi(v)$ and v^σ are obtained by, respectively, *permutation* and *inverse* operations π and σ over v. In cases when $\pi = 12 \dots n$ or $\sigma = 11 \dots 1$, that is, the operations π or σ are identity ones, we write $\pi = 1$ or $\sigma = 1$ respectively.

Taking any $g(x_1, x_2, \dots, x_n)$ in IS_n, σ_1, σ_2 in $\mathbb{F}_2^n$, and π_1, π_2 in S_n, we then can define a vector function $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ as $f(x) = \pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$, $x = x_1 x_2 \dots x_n$. Particularly, $g(x)$ can be the identical function, that is, for each i in $\{1, 2, \dots, n\}$ its coordinate function $g_i(x)$ can be equal to x_i. In any case, the table of the function $f(x)$ is obtained from the table of the function $g(x)$ by

- substituting columns corresponding to some variables for inverses (in σ_1),
- transposing (according to π_1) columns corresponding to some variables,
- substituting columns corresponding to some coordinate functions of $g(x)$ for inverses (in σ_2), and
- transposing (according to π_2) columns corresponding to some coordinate functions of the function $g(x)$.

In other words, $f(x)$ is computed from the function $g(x)$ by the inversion and transposition of some its inputs and outputs and, like g, is of a bounded complexity and satisfies the invertibility condition. Therefore, $f(x) \in IS_n$.

Define $K_n(g) = \{\pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1}))) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_1, \pi_2 \in S_n\}$. Thus, we get that $K_n(g) \subseteq IS_n$ and $|K_n(g)| \leqslant (2^n n!)^2$. Any subset $K \subseteq K_n(g)$ of an exponential cardinality can be taken as a synthesis result — the key space of a funkeysubcipher $\mathcal{C}$. The following subsets of $K_n(g)$ are possible candidates for playing this role:

$K_n(g, 1) = \{g(x^{\sigma_1}) : \sigma_1 \in \mathbb{F}_2^n\}$, $|K_n(g, 1)| \leqslant 2^n$;
$K_n(g, 2) = \{g(\pi_1(x)) : \pi_1 \in S_n\}$, $|K_n(g, 2)| \leqslant n!$;
$K_n(g, 3) = \{g(\pi_1(x^{\sigma_1})) : \sigma_1 \in \mathbb{F}_2^n, \pi_1 \in S_n\}$, $|K_n(g, 3)| \leqslant 2^n n!$;
$K_n(g, 4) = \{g^{\sigma_2}(x) : \sigma_2 \in \mathbb{F}_2^n\}$, $|K_n(g, 4)| \leqslant 2^n$;
$K_n(g, 5) = \{g^{\sigma_2}(x^{\sigma_1}) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n\}$, $|K_n(g, 5)| \leqslant 2^{2n}$;
$K_n(g, 6) = \{g^{\sigma_2}(\pi_1(x)) : \sigma_2 \in \mathbb{F}_2^n, \pi_1 \in S_n\}$, $|K_n(g, 6)| \leqslant 2^n n!$;
$K_n(g, 7) = \{g^{\sigma_2}(\pi_1(x^{\sigma_1})) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_1 \in S_n\}$, $|K_n(g, 7)| \leqslant 2^{2n} n!$;
$K_n(g, 8) = \{\pi_2(g(x)) : \pi_2 \in S_n\}$, $|K_n(g, 8)| \leqslant n!$;
$K_n(g, 9) = \{\pi_2(g(x^{\sigma_1})) : \sigma_1 \in \mathbb{F}_2^n, \pi_2 \in S_n\}$, $|K_n(g, 9)| \leqslant 2^n n!$;
$K_n(g, 10) = \{\pi_2(g(\pi_1(x))) : \pi_2, \pi_1 \in S_n\}$, $|K_n(g, 10)| \leqslant (n!)^2$;

$K_n(g,11) = \{\pi_2(g(\pi_1(x^{\sigma_1}))) : \sigma_1 \in \mathbb{F}_2^n, \pi_1, \pi_2 \in S_n\}$, $|K_n(g,11)| \leqslant 2^n(n!)^2$;
$K_n(g,12) = \{\pi_2(g^{\sigma_2}(x)) : \sigma_2 \in \mathbb{F}_2^n, \pi_2 \in S_n\}$, $|K_n(g,12)| \leqslant 2^n n!$;
$K_n(g,13) = \{\pi_2(g^{\sigma_2}(x^{\sigma_1})) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_2 \in S_n\}$, $|K_n(g,13)| \leqslant 2^{2n} n!$;
$K_n(g,14) = \{\pi_2(g^{\sigma_2}(\pi_1(x))) : \sigma_2 \in \mathbb{F}_2^n, \pi_1, \pi_2 \in S_n\}$, $|K_n(g,14)| \leqslant 2^n(n!)^2$;
$K_n(g,15) = K_n(g)$.

4. Funkeysubciphers with key functions in a bounded number of essential variables

Let s_0 and n be some integers, $1 \leqslant s_0 \leqslant n$, and $B_{s_0,n}$ be the set of all Boolean functions $f(x_1,\dots,x_n)$ essentially depending on not more than s_0 variables $x_1,\dots,x_n$, that is, for any $f : \mathbb{F}_2^n \to \mathbb{F}_2$,

$$f(x_1,\dots,x_n) \in B_{s_0,n} \Leftrightarrow$$
$$\Leftrightarrow \exists s \leqslant s_0\ \exists i_1,\dots,i_s \in \{1,\dots,n\} \exists g : \mathbb{F}_2^s \to \mathbb{F}_2 (f(x_1,\dots,x_n) = g(x_{i_1},\dots,x_{i_s})).$$

The set of variables $x_{i_1},\dots,x_{i_s}$ satisfying this equation is said to be a *sufficient subset* of arguments for the function f. If U is a sufficient subset for f and, for any $V \subset U$, V isn't a sufficient for f, then the variables in U are said to be *essential* arguments for f. For natural $s \leqslant s_0$, let $B^*_{s,n}$ be the set of all functions in $B_{s_0,n}$ essentially depending on exactly s variables. It is clear that $B_{s_0,n} = \bigcup\limits_{s=1}^{s_0} B^*_{s,n}$. We suppose that the number s_0 is small enough for accepting functions in $B_{s_0,n}$ to be of a bounded complexity.

Let $IS_{s_0,n}$ denote the set of all bijective Boolean vector functions each consisting of n coordinate functions in $B_{s_0,n}$. Balancedness of each coordinate function of a Boolean vector function f is the necessary condition for bijectivity of f. So the cardinality of $IS_{s_0,n}$ doesn't exceed the number $N_{s_0,n} = \left(\binom{n}{s_0}\binom{2^{s_0}}{2^{s_0-1}}\right)^n$, that is the number of all n-dimensional vectors with coordinates being balanced Boolean functions in s_0 variables taken in all possible ways from the set $\{x_1,\dots,x_n\}$.

A *funkeysubcipher with key functions in a bounded number of essential variables* is a funkeysubcipher $\mathcal{C} = (\mathbb{F}_2^n, K, \mathbb{F}_2^n, E, D)$, where $K \subseteq IS_{s_0,n}$. To synthesize these ciphers, we need generating the key spaces K for them from vector functions in $IS_{s_0,n}$. Here, we propose to do this just in the same way as we have done above in the set IS_n using the inverse and permutation operations.

Namely, take a vector function $g(x_1,x_2,\dots,x_n)$ in $IS_{s_0,n}$. Let $g = (g_1,\dots,g_n)$. By the definition of $IS_{s_0,n}$, for every $i \in \{1,\dots,n\}$, there exists a natural $s_i \leqslant s_0$ such that $g_i \in$ $\in B^*_{s_i,n}$, that is, g_i essentially depends on s_i variables. Define $K_{s_0,n}(g) = \{\pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1}))) :$ $\sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_1, \pi_2 \in S_n\}$, where $x = x_1x_2\dots x_n$. Thus, we get that $K_{s_0,n}(g) \subseteq IS_{s_0,n}$ and $|K_{s_0,n}(g)| \leqslant (2^n n!)^2$. Moreover, for any function $f = (f_1,\dots,f_n) \in K_{s_0,n}(g)$ and any $i \in \{1,\dots,n\}$, the number of essential variables of f_i equals s_j, the number of essential variables of g_j where $j = \pi_2^{-1}(i)$.

Any subset $K \subseteq K_{s_0,n}(g)$ of an exponential cardinality can be taken as the key space of the funkeysubcipher $\mathcal{C}$ with key functions in a bounded number of essential variables. In particular, this role can be successfully played by the subsets $K_{s_0,n}(g,j)$ that are formally defined, just as $K_n(g,j)$, $j = 1,2,\dots,15$, have been done. For example, $K_{s_0,n}(g,7) =$ $= \{g^{\sigma_2}(\pi_1(x^{\sigma_1})) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_1 \in S_n\}$, $|K_{s_0,n}(g,7)| \leqslant 2^{2n} n!$, and $K_{s_0,n}(g,15) = K_{s_0,n}(g)$. The only difference is in the class of the function g that, for $K_n(g,j)$, belongs to IS_n and, for $K_{s_0,n}(g,j)$, belongs to $IS_{s_0,n}$.

To produce subsets $K \subseteq K_{s_0,n}(g)$ as key spaces for funkeysubciphers with key functions in a bounded number of essential variables, we need to have a capability to generate vector Boolean functions $g = (g_1 \dots g_n)$ in $I_{s_0,n}$ with various values of their parameters n, s_0, s_1, …, s_n. Unfortunately, we have no any exhaustive solution of this problem and can only present now a pair of some restricted relevant methods.

Let $IS^*_{s,n}$ denote the set of all bijective Boolean vector functions each consisting of n coordinate functions in $B^*_{s,n}$. The methods just mentioned construct functions from $IS^*_{s,n}$.

The first method is used in the case when $s \geqslant 3$ and $s|n$, i.e. $n = st$ for some $t \in \mathbb{N}$. It is proved in [7] that $IS^*_{s,s} \neq \varnothing$ for all $s \geqslant 3$. So, we can construct t functions $g^{(i)} = g_1^{(i)} \dots g_s^{(i)} \in IS^*_{s,s}$, $i = 1, \dots, t$. Then the function $g(x_1, \dots, x_n) = g_1^{(1)}(x_1, \dots, x_s) \dots g_s^{(1)}(x_1, \dots, x_s) g_1^{(2)}(x_{s+1}, \dots, x_{2s}) \dots g_s^{(2)}(x_{s+1}, \dots, x_{2s}) \dots g_1^{(t)}(x_{(t-1)s+1}, \dots, x_n) \dots g_s^{(t)}(x_{(t-1)s+1}, \dots, x_n)$ belongs to $IS^*_{s,n}$.

The second method starts from $g^{(1)}(x_1, \dots, x_s) = g_1 \dots g_s \in IS^*_{s,s}$ too. Then we construct the function $g^{(2)}(x_1, \dots, x_s, x_{s+1}) = g_1 \dots g_s h$ where $h = x_{s+1} \oplus q(x_1, \dots, x_s)$ and $q \in B^*_{s-1,s}$. It is proved in [8] that $g^{(2)} \in IS^*_{s,s+1}$. Repeating this step, we successively obtain the functions $g^{(3)} \in IS^*_{s,s+2}$ (using the functions $h = x_{s+2} \oplus q(x_1, \dots, x_s, x_{s+1})$ and $q \in B^*_{s-1,s+1}$), …, $g^{(n+s-1)} \in IS^*_{s,n}$.

5. Cryptanalysis

5.1. Cryptanalysis problem

In this section, we consider the cryptanalysis problem for funkeysubciphers giving our attention to ciphers with key functions in bounded numbers of essential variables. Moreover, we confine the consideration to ciphers with key spaces $K = K_{s_0,n}(g, j)$, where g is an arbitrary function in $(B_{s_0,n})^n$ and j can be assigned any value from $\{1, \dots, 15\}$. However, for some parameter j values, the cryptanalysis methods proposed here actually hold for ciphers with the wider key spaces, particularly with $K = K_n(g, j)$.

We assume that the cryptanalyst exploits a known plaintext attack with the threat of total break (secret key recovery). This means that he possesses some blocks $P_1, \dots, P_m$ of a plaintext and corresponding blocks $C_1, \dots, C_m$ of a ciphertext and tries to determine the key that was used, that is, a function $f(x) \in K = K_{s_0,n}(g, j)$ such that $C_l = f(P_l)$ for all $l \in \{1, \dots, m\}$. According to Kerckhoff's principle, it is supposed that the cryptanalyst knows the cipher $\mathcal{C} = (\mathbb{F}_2^n, K, \mathbb{F}_2^n, E, D)$ being used. Particularly, he knows the key space $K = K_{s_0,n}(g, j)$ and its parameters $g \in (B_{s_0,n})^n$, $n \in \mathbb{N}$, $s_0 \leqslant n$, and $j \in \{1, \dots, 15\}$. The knowledge of the function $g(x_1, \dots, x_n)$ yields the knowledge of its inverse g^{-1}, coordinate functions $g_1, \dots, g_n$ in $B_{s_0,n}$ and the sets $X_1, \dots, X_n$ of their essential variables respectively, $X_i \subseteq X = \{x_1, \dots, x_n\}$, $i = 1, \dots, n$. On the base of this information, the cryptanalyst has to determine the coordinate functions $f_1, \dots, f_n$ of a key function $f(x_1, \dots, x_n)$ in $K_{s_0,n}(g, j)$ which satisfies the equalities $f(P_l) = C_l$ for all $l \in \{1, \dots, m\}$. Here, for each $i \in \{1, \dots, n\}$, the function f_i belongs to $B_{s_0,n}$, its essential variables form a subset $U_i \subseteq X$, and $|U_i| = |X_i| = s_i$ if the permutation π_2 in the expression for $K_{s_0,n}(g, j)$ is the identity one. For the cryptanalyst, to determine the function f_i means to determine the set U_i of its essential variables and the value of f_i for each combination of values of variables in U_i.

Below we first give a general solution of the problem comprising the all fifteen partial cases of it and then present specific solutions for some of these cases.

5.2. General cryptanalysis method

The method concerns the funkeysubcipher $\mathcal{C}$ with the general key space $K = K_{s_0,n}(g) = \{\pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1}))) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_1, \pi_2 \in S_n\}$ which includes the partial key spaces $K_{s_0,n}(g, j)$ for all $j \in \{1, \ldots, 15\}$.

Recall that we have a string of Boolean variables $x = x_1 x_2 \ldots x_n$, a vector Boolean function $g(x) = g_1(x)g_2(x)\ldots g_n(x)$ with coordinate functions $g_1, \ldots, g_n$, where $g_i \in B^*_{s_i,n}$ for $1 \leqslant s_i \leqslant s_0$ and $i = 1, \ldots, n$, the blocks of a plain text $P_1, \ldots, P_m$ and the corresponding blocks of a ciphertext $C_1, \ldots, C_m$.

Let $k \in K$ and $C_l = C_{l1}C_{l2}\ldots C_{ln}$, $l = 1, \ldots, m$. Denote $f(x) = f_1(x)f_2(x)\ldots f_n(x) = \pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$. Then $f_i \in B_{s_0,n}$, $k = f(x)$, $C_l = f(P_l)$, and $C_{li} = f_i(P_l)$, $l = 1, \ldots, m$ and $i = 1, \ldots, n$.

Thus, the cryptanalysis problem is as follows: for every $i \in \{1, \ldots, n\}$ and given equalities $C_{li} = f_i(P_l)$, $l = 1, \ldots, m$, determine the function $f_i(x)$. The problem is divided into two subproblems: find out essential variables of the function f_i and compute its values for all possible values of these variables. In connection with the first subproblem, we need to note that the number of essential variables of f_i depends on whether the permutation π_2 in $f(x) = \pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$ is the identity one or not. If the answer is "yes", then f_i has the same number s_i of essential variables as g_i. Otherwise we can only say that this number is less or equal to $\max\{s_1, \ldots, s_n\}$ and doesn't exceed s_0.

To solve the first subproblem, we now present some auxiliary results. Let $f'(x_1, \ldots, x_n)$ be a, possibly, partial Boolean function given by two subsets $M^0_{f'} \subseteq \mathbb{F}_2^n$ and $M^1_{f'} \subseteq \mathbb{F}_2^n$ so that $\alpha \in M^b_{f'} \Leftrightarrow f'(\alpha) = b$, $b \in \mathbb{F}_2$.

We first define the following sets:

$$D(f') = \{\alpha \oplus \beta : \alpha \in M^0_{f'}, \beta \in M^1_{f'}\},$$
$$\inf D(f') = \{\delta : \delta \in D(f'),\ \neg\exists\delta' \in D(f')(\delta' < \delta)\},$$

where for $\delta = d_1 \ldots d_n$ and $\delta' = d'_1 \ldots d'_n$ in $\mathbb{F}_2^n$, $\delta' < \delta \Leftrightarrow \delta' \neq \delta\ \&\ \forall t \in \{1, \ldots, n\}\ (d'_t \leqslant d_t)$. Particularly, in our case,

$$D(f_i) = \{P_l \oplus P_j : C_{li} \neq C_{ji},\ l, j = 1, 2, \ldots, m\}.$$

We next construct the Boolean matrix $M' = ||\inf D(f')||$ with the set of rows that is equal to $\inf D(f')$. The columns in M' with the numbers 1, 2, ..., n are assigned to variables $x_1, x_2, \ldots, x_n$ respectively. A subset J of them is said to be a *cover* of M' if for each row in M', there is a column in J with the value 1 in this row. The cover J is *minimal* if it doesn't contain as a subset another cover of M'.

At last, we note that in [9] we have proved that a subset of variables $U = \{x_{j_1}, \ldots, x_{j_s}\}$ is sufficient for f' iff the subset J of columns in M' with the numbers $j_1, \ldots, j_s$ is a cover of M', and U is essential for f' iff J is a minimal cover of M'. Moreover, U is a unique subset of essential variables for f' iff J is a unique cover of M'; in this case, each row in M' is a unit vector e_j (with a 1 in the j-th coordinate and 0's elsewhere) and $U = \{x_{j_1}, \ldots, x_{j_s}\}$ if all the rows in M' are $e_{j_1}, \ldots, e_{j_s}$.

Also, in [10], we have proved that a subset of variables $\{x_{j_1}, \ldots, x_{j_s}\}$ is a unique subset of essential variables for a function f' in $B_{s_0,n}$ iff all the covers of the matrix $||\inf D(f')||$, the cardinalities of which don't exceed s_0, have a non-empty intersection consisting of columns with the numbers $j_1, \ldots, j_s$.

So, finding a unique subset of essential variables (if it exists) for the function f_i in $B_{s_0,n}$ and thus solving the first cryptanalysis subproblem is reduced to computing, for the

matrix $||\inf D(f_i)||$, the intersection of all covers whose cardinalities are not more than s_0. The computational complexity of this work is $O(2^{s_0})$.

Under the known essential variables $x_{i_1}, \dots, x_{i_{s_i}}$ of f_i, any solution of the second cryptanalysis subproblem for $i \in \{1, \dots, n\}$ can be obtained as $f_i(x) = h_i(x_{i_1}, \dots, x_{i_{s_i}})$, where $h_i : \mathbb{F}_2^{s_i} \to \mathbb{F}_2$, the vector function $h_1(x_{1_1}, \dots, x_{1_{s_1}})h_2(x_{2_1}, \dots, x_{2_{s_2}}) \dots h_n(x_{n_1}, \dots, x_{n_{s_n}})$ is a bijection on $\mathbb{F}_2^n$, and, for all $\alpha = a_1 a_2 \dots a_n \in \mathbb{F}_2^n$, if $\alpha = P_l$ and $l \in \{1, \dots, m\}$, then $h_i(a_{i_1}, \dots, a_{i_{s_i}}) = C_{li}$.

In particular, if for each $i \in \{1, \dots, n\}$, the set $\{x_{i_1}, \dots, x_{i_{s_i}}\}$ is a unique subset of essential variables for f_i and $P = \{P_{li_1} P_{li_2} \dots P_{li_{s_i}} : l = 1, \dots, m\} = \mathbb{F}_2^{s_i}$, then the solution $f(x)$ of the cryptanalysis problem for the cipher $\mathcal{C}$ is unique and, for $i \in \{1, \dots, n\}$, it has $f_i(x) = h_i(x_{i_1}, \dots, x_{i_{s_i}})$, where $h_i(P_{li_1} P_{li_2} \dots P_{li_{s_i}}) = C_{li}$, $l \in \{1, \dots, m\}$.

In the case of $P \neq \mathbb{F}_2^{s_i}$, the following problem arises: given a partially defined Boolean function $f'(x_1, \dots, x_n)$ and a subset $\{i_1, \dots, i_s\} \subset \{1, \dots, n\}$, find (if exists) a completely defined Boolean function $h(x_{i_1}, \dots, x_{i_s})$ such that $h(a_{i_1} a_{i_2} \dots a_{i_s}) = f'(a_1 a_2 \dots a_n)$ for each n-tuple $(a_1 a_2 \dots a_n)$ from the domain of f'. This problem is a special case of the known problem of completing a partial function in a functional class and isn't a subject of this research.

For making references to the general cryptanalysis method described here, we name it GCM. The core of GCM is the algorithm for finding, for a given partially defined Boolean function $f'(x_1, \dots, x_n)$ from $B_{s_0,n}$, such a function $h(x_{i_1}, \dots, x_{i_s}) \in B^*_{s,n}$ that $h(x_{i_1}, \dots, x_{i_s}) = f'(x_1, \dots, x_n)$ on the domain of f'. We denote this algorithm by $\mathcal{B}$.

As for the parameters of the cryptanalysis problem, namely $g, \sigma_1, \sigma_2, \pi_1, \pi_2$, GCM doesn't depend directly on them both in the contents and in a result. This is not an accidental fact, but it is because these parameters are not really the key k of the cipher $\mathcal{C}$, they only form the expression $\pi_2(g^{\sigma_2}(\pi_1(x^{\sigma_1})))$ to specify a bijective function $f : \mathbb{F}^n \to \mathbb{F}^n$ which is in fact the key k of $\mathcal{C}$ and the result of GCM execution over the given pairs (P_i, C_i), $i = 1, \dots, n$.

5.3. Some particular cryptanalysis methods

Some particular cryptanalysis methods for a cipher $\mathcal{C}$ under consideration can be obtained by applying GCM to ciphers $\mathcal{C}_j$ with key spaces $K_{s_0,n}(g, j)$ for $j \in \{1, \dots, 14\}$. We think of these methods as key space limitations of the general method and denote by GCM_j. For example, GCM_9 and GCM_{14} are GCM for ciphers with key spaces $K = K_{s_0,n}(g, 9) = \{\pi_2(g(x^{\sigma_1})) : \sigma_1 \in \mathbb{F}_2^n, \pi_2 \in S_n\}$ and $K = K_{s_0,n}(g, 14) = \{\pi_2(g^{\sigma_2}(\pi_1(x))) : \sigma_2 \in \mathbb{F}_2^n, \pi_1, \pi_2 \in S_n\}$ respectively.

Now, we consider some other particular cryptanalysis methods that are not exactly key space limitations of GCM, but give special solutions to some ciphers $\mathcal{C}_j$ with limited key spaces.

Cases $j = 1, 4, 5$

Describing cryptanalysis methods in these cases, we limit our exposition to determination of the inverse operations for obtaining f from g.

Let $g^{-1} = (g_1^{-\prime}, g_2^{-\prime}, \dots, g_n^{-\prime})$, $f^{-1} = (f_1^{-\prime}, f_2^{-\prime}, \dots, f_n^{-\prime})$, $\sigma_1 = \sigma_{11}\sigma_{12} \dots \sigma_{1n}$, $\sigma_2 = \sigma_{21}\sigma_{22} \dots \sigma_{2n}$, $P_l = P_{l1}P_{l2} \dots P_{ln}$, and $C_l = C_{l1}C_{l2} \dots C_{ln}$, $l = 1, 2, \dots, m$.

In the *case* $j = 1$, where $K = K_{s_0,n}(g, 1) = \{g(x^{\sigma_1}) : \sigma_1 \in \mathbb{F}_2^n\}$, the cryptanalysis problem is trivial because, for every $l \in \{1, \dots, m\}$, $C_l = g(P_l^{\sigma_1})$, $P_l^{\sigma_1} = g^{-1}(C_l)$, $P_{li}^{\sigma_{1i}} = g_i^{-\prime}(C_l)$, $i \in \{1, \dots, n\}$, and σ_{1i} is computed by using the Boolean implication

$$(a^b = c) \Rightarrow (b = 1 \Leftrightarrow c = a),$$

namely $\sigma_{1i} = 1 \Leftrightarrow g_i^{-\prime}(C_l) = P_{li}$ for all $i \in \{1, 2, \ldots, n\}$ and some (any) $l \in \{1, \ldots, m\}$, particularly for $l = 1$.

By the same reason, the problem is trivial in the *case* $j = 4$, where $K = K_{s_0,n}(g, 4) = \{g^{\sigma_2}(x) : \sigma_2 \in \mathbb{F}_2^n\}$, because $C_l = g^{\sigma_2}(P_l)$, $C_l^{\sigma_2} = g(P_l)$ and σ_2 is computed by using the same implication: $\sigma_{2i} = 1 \Leftrightarrow g_i(P_l) = C_{li}$, $i = 1, 2, \ldots, n$.

In the *case* $j = 5$, where $K = K_{s_0,n}(g, 5) = \{g^{\sigma_2}(x^{\sigma_1}) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n\}$, we have $C_l = g^{\sigma_2}(P_l^{\sigma_1})$, $C_l^{\sigma_2} = g(P_l^{\sigma_1})$, and $P_l^{\sigma_1} = g^{-1}(C_l^{\sigma_2})$. For every pair (σ_2, l), where $\sigma_2 \in \mathbb{F}_2^n$ and $l = 1, 2, \ldots, m$, compute the value $\sigma_1^{(\sigma_2,l)} = \sigma_{11}^{(\sigma_2,l)}\sigma_{12}^{(\sigma_2,l)} \ldots \sigma_{1n}^{(\sigma_2,l)}$ of the vector σ_1 in $\mathbb{F}_{2^n}$ by using the algorithm of the case $j = 1$, namely

$$\sigma_{1i}^{(\sigma_2,l)} = 1 \Leftrightarrow g_i^{-\prime}(C_l^{\sigma_2}) = P_{li}, i = 1, \ldots, m.$$

The result of the cryptanalysis is a pair (σ_1, σ_2) satisfying the equality $\sigma_1 = \sigma_1^{(\sigma_2,l)}$ for all $l \in \{1, \ldots, m\}$. Note that this answer is not sure to be unique. The computational complexity of the algorithm is $O(2^n)$.

Note that the attacks described in these cases successfully work on ciphers with $K = K_n(g, j)$ for $j = 1, 4, 5$ respectively and $g \in IS_n$.

Case $j = 7$

In this case, $K = K_{s_0,n}(g, 7) = \{g^{\sigma_2}(\pi_1(x^{\sigma_1})) : \sigma_1, \sigma_2 \in \mathbb{F}_2^n, \pi_1 \in S_n\}$ and the cipher under consideration is $\mathcal{C}_7$ that is the partial case of $\mathcal{C}$, where $\pi_2 = 1$. Besides, the ciphers $\mathcal{C}_j$ for all $j \in \{1, \ldots, 6\}$ are partial cases of $\mathcal{C}_7$, and the cryptanalysis problem for them can be solved by any method solving this problem for $\mathcal{C}_7$. The method presented here is an amplification of GCM, namely, instead of method $\mathcal{B}$, a method $\mathcal{A}$ is used, which takes into attention the condition $\pi_2 = 1$, yielding the fact that a function $f_i(x)$ to be found has the same number s_i of essential variables as the known function g_i. So, finding essential variables for f_i is reduced in $\mathcal{A}$ to finding, for the matrix $||\inf D(f_i)||$, a minimal cover of the given cardinality — s_i. The computational complexity of the last problem doesn't exceed $\binom{n}{s_i}$. In other details, the cryptanalysis method for $\mathcal{C}_7$ coincides with GCM.

Some program implementations of algorithms $\mathcal{A}$ and $\mathcal{B}$ and the results of their thorough testing on computers have been presented in [2].

REFERENCES

1. *Agibalov G. P. and Levashnikov A. A.* Statisticheskoe issledovanie zadachi opoznaniya bulevykh funktsiy odnogo klassa [Statistical study of the identifying problem for a class of Boolean functions]. Proc. ASDA Conf., Novosibirsk, 1966, pp. 40–45. (in Russian)
2. *Agibalov G. P. and Sungurova O. G.* Kriptoanaliz konechno-avtomatnogo generatora klyuchevogo potoka s funktsiey vykhodov v kachestve klyucha [Cryptanalysis of a finite-state keystream generator with an output function as a key]. Vestnik TSU. Prilozhenie, 2006, no. 17, pp. 104–108. (in Russian)
3. *Agibalov G. P.* SIBCiphers — simmetrichnye iterativnye blochnye shifry iz bulevykh funktsiy s klyuchevymi argumentami [SIBCiphers — symmetric iterative block ciphers composed of Boolean functions depending on small number of variables]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2014, no. 7, pp. 43–48. (in Russian)
4. *Agibalov G. P.* Watermarking ciphers. Prikladnaya Diskretnaya Matematika, 2016, no. 1(31), pp. 62–66.
5. *Agibalov G. P. and Pankratova I. A.* O dvukhkaskadnykh konechno-avtomatnykh kriptograficheskikh generatorakh i metodakh ikh kriptoanaliza [About 2-cascade finite

automata cryptographic generators and their cryptanalysis]. Prikladnaya Diskretnaya Matematika, 2017, no. 35, pp. 38–47. (in Russian)

6. *Agibalov G. P.* Kriptoavtomaty s funktsional'nymi klyuchami [Cryptautomata with functional keys]. Prikladnaya Diskretnaya Matematika, 2017, no. 36, pp. 59–72. (in Russian)
7. *Pankratova I. A.* Construction of invertible vectorial Boolean functions with coordinates depending on given number of variables. Proc. CSIST'2016, Minsk, BSU Publ., 2016, pp. 519–521.
8. *Pankratova I. A.* Ob obratimosti vektornykh bulevykh funktsiy [On the invertibility of vector Boolean functions]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2015, no. 8, pp. 35–37. (in Russian)
9. *Agibalov G. P.* Minimizatsiya chisla argumentov bulevykh funktsiy [Number minimization for variables a partial Boolean function depends on]. Problemy Sinteza Tsifrovykh Avtomatov, Moscow, Nauka Publ., 1967, pp. 96–100. (in Russian)
10. *Agibalov G. P.* O nekotorykh doopredeleniyakh chastichnoy bulevoy funktsii [Some completions of partial Boolean function]. Trudy SPhTI, 1970, iss. 49, pp. 12–19. (in Russian)

МАТЕМАТИЧЕСКИЕ ОСНОВЫ НАДЁЖНОСТИ ВЫЧИСЛИТЕЛЬНЫХ И УПРАВЛЯЮЩИХ СИСТЕМ

УДК 519.718.7

ЕДИНИЧНЫЕ ПРОВЕРЯЮЩИЕ ТЕСТЫ ДЛЯ СХЕМ ИЗ ФУНКЦИОНАЛЬНЫХ ЭЛЕМЕНТОВ В БАЗИСЕ «КОНЪЮНКЦИЯ-ОТРИЦАНИЕ»[1]

К. А. Попков

Институт прикладной математики им. М. В. Келдыша РАН, г. Москва, Россия

Рассматривается задача синтеза неизбыточных схем из функциональных элементов, реализующих булевы функции от n переменных и допускающих короткие единичные проверяющие тесты относительно однотипных константных неисправностей на выходах элементов, в базисе $\{\&, \neg\}$ и схожих базисах. Для каждой булевой функции, допускающей реализацию неизбыточной схемой, найдено минимально возможное значение длины такого теста. В частности, доказано, что оно не превосходит трёх.

Ключевые слова: *схема из функциональных элементов, константная неисправность, единичный проверяющий тест.*

DOI 10.17223/20710410/38/5

SINGLE FAULT DETECTION TESTS FOR LOGIC NETWORKS OF AND, NOT GATES

K. A. Popkov

Keldysh Institute of Applied Mathematics, Moscow, Russia

E-mail: kirill-formulist@mail.ru

Let $D_1(f)$ ($D_0(f)$) be the least length of a fault detection test for irredundant logic networks consisting of logic gates in the basis $\{\&, \neg\}$, implementing a given Boolean function $f(x_1, \ldots, x_n)$, and having at most one stuck-at-1 (stuck-at-0 respectively) fault on outputs of the logic gates. Let $f_1 = x_i$; $f_2 = 0$, $\overline{x_i}$, or $x_{i_1}^{\sigma_1}\&x_{i_2}\ldots\&x_{i_k}$; $f_3 = \overline{x_{i_1}}\&\overline{x_{i_2}}\&x_{i_3}\&\ldots\&x_{i_k}$ or $\underbrace{(\ldots((}_{k-1}x_{i_1}^{\sigma_1}\&x_{i_2})^{\sigma_2}\&x_{i_3})^{\sigma_3}\&\ldots\&x_{i_k})^{\sigma_k}$; $f_4 = x_{i_1}^{\sigma_1}\&\ldots\&x_{i_k}^{\sigma_k}$; $f_5 = \underbrace{(\ldots((}_{k-1}x_{i_1}^{\sigma_1}\&x_{i_2}^{\sigma_2})^{\delta_1}\&x_{i_3}^{\sigma_3})^{\delta_2}\&\ldots\&x_{i_k}^{\sigma_k})^{\delta_{k-1}}$, where $2 \leqslant k \leqslant n$ for f_2, f_3, and f_5; $1 \leqslant k \leqslant n$ for f_4; $\sigma_1, \ldots, \sigma_k, \delta_1, \ldots, \delta_{k-1} \in \{0, 1\}$; $i, i_1, \ldots, i_k \in \{1, \ldots, n\}$; indices $i_1, \ldots, i_k$ are pairwise different; for f_3, at least one of numbers $\sigma_2, \ldots, \sigma_k$ equals 0 and if $k = 2$, then assume $x_{i_3}\&\ldots\&x_{i_k} \equiv 1$; for f_5, at least one of numbers $\delta_1, \ldots, \delta_{k-1}$

[1]Работа выполнена при поддержке гранта РНФ, проект № 14-21-00025 П.

equals 0. It is proved that, for each Boolean function $f(x_1, \ldots, x_n) \not\equiv 1$,

$$D_1(f) = \begin{cases} 0, & \text{iff the function } f \text{ is representable in the form of } f_1, \\ 1, & \text{iff the function } f \text{ is representable in the form of } f_2, \\ 2, & \text{iff the function } f \text{ is representable in the form of } f_3, \\ 3 & \text{otherwise.} \end{cases}$$

If $f \equiv 1$ then the value $D_1(f)$ is undefined. Also, it is proved that, for each Boolean function $f(x_1, \ldots, x_n)$ which is different from constants,

$$D_0(f) = \begin{cases} 0, & \text{iff the function } f \text{ is representable in the form of } f_1, \\ 1, & \text{iff the function } f \text{ is representable in the form of } f_4 \text{ but not of } f_1, \\ 2, & \text{iff the function } f \text{ is representable in the form of } f_5, \\ 3 & \text{otherwise.} \end{cases}$$

If $f \equiv 1$ or $f \equiv 0$ then the value $D_0(f)$ is undefined.

Keywords: *logic network, stuck-at fault, single fault detection test.*

Введение

В работе рассматривается задача синтеза легкотестируемых схем, реализующих заданные булевы функции. Логический подход к тестированию электрических схем предложен С. В. Яблонским и И. А. Чегис в [1]; этот подход также применим к тестированию схем из функциональных элементов [2–4]. Пусть имеется схема из функциональных элементов S с одним выходом (без обратных связей), реализующая булеву функцию $f(\tilde{x}^n)$, где $\tilde{x}^n = (x_1, \ldots, x_n)$. Под воздействием некоторого источника неисправностей один или несколько элементов схемы S могут перейти в неисправное состояние. В результате схема S вместо исходной функции $f(\tilde{x}^n)$ будет реализовывать некоторую булеву функцию $g(\tilde{x}^n)$, вообще говоря, отличную от f. Все такие функции $g(\tilde{x}^n)$, получающиеся при всевозможных допустимых для рассматриваемой задачи неисправностях элементов схемы S, называются *функциями неисправности* данной схемы.

Введём следующие определения [2–4]. *Проверяющим тестом* для схемы S называется такое множество T наборов значений переменных $x_1, \ldots, x_n$, что для любой отличной от $f(\tilde{x}^n)$ функции неисправности схемы S в T найдётся набор $\tilde{\sigma}$, на котором $f(\tilde{\sigma}) \neq g(\tilde{\sigma})$. *Диагностическим тестом* для схемы S называется такое множество T наборов значений переменных $x_1, \ldots, x_n$, что T является проверяющим тестом и, кроме того, для любых двух различных функций неисправности $g_1(\tilde{x}^n)$ и $g_2(\tilde{x}^n)$ схемы S в T найдётся набор $\tilde{\sigma}$, на котором $g_1(\tilde{\sigma}) \neq g_2(\tilde{\sigma})$. Число наборов в T называется *длиной* теста. В качестве тривиального диагностического (и проверяющего) теста длины 2^n для схемы S всегда можно взять множество, состоящее из всех двоичных наборов длины n. Тест называется *полным*, если в схеме могут быть неисправны сколько угодно элементов, и *единичным*, если в схеме может быть неисправен только один элемент. Единичные тесты обычно рассматривают для неизбыточных схем [4], т. е. для таких схем, в которых любая допустимая неисправность любого одного элемента приводит к функции неисправности, отличной от исходной функции, реализуемой схемой.

Любое множество булевых функций будем называть *базисом*.

Пусть зафиксирован вид неисправностей элементов, B — произвольный функционально полный базис и T — единичный проверяющий тест (ЕПТ) для некоторой

схемы S в базисе B. Введём следующие обозначения: $D^B_{\mathrm{s,detect}}(T)$ — длина теста T; $D^B_{\mathrm{s,detect}}(S) = \min D^B_{\mathrm{s,detect}}(T)$, где минимум берётся по всем ЕПТ T для схемы S; $D^B_{\mathrm{s,detect}}(f) = \min D^B_{\mathrm{s,detect}}(S)$, где минимум берётся по всем неизбыточным схемам S в базисе B, реализующим функцию f; $D^B_{\mathrm{s,detect}}(n) = \max D^B_{\mathrm{s,detect}}(f)$, где максимум берётся по всем булевым функциям f от n переменных, для которых определено значение $D^B_{\mathrm{s,detect}}(f)$. Функция $D^B_{\mathrm{s,detect}}(n)$ называется *функцией Шеннона* длины ЕПТ. По аналогии с функциями $D^B_{\mathrm{s,detect}}$ можно ввести функции $D^B_{\mathrm{s,diagn}}$, $D^B_{\mathrm{c,detect}}$ и $D^B_{\mathrm{c,diagn}}$ для соответственно единичного диагностического теста, полного проверяющего и полного диагностического тестов, зависящие от T, от S, от f и от n (в определениях функций $D^B_{\mathrm{c,detect}}(f)$ и $D^B_{\mathrm{c,diagn}}(f)$ не требуется предполагать неизбыточность схем). Так, например, $D^B_{\mathrm{c,diagn}}(n)$ — функция Шеннона длины полного диагностического теста.

Перечислим основные результаты, касающиеся тестирования схем из функциональных элементов. Класс допустимых неисправностей функциональных элементов ограничим константными неисправностями на выходах элементов, при которых значение на выходе любого неисправного элемента становится равно некоторой булевой константе. Неисправности на выходах элементов называются однотипными константными типа p, если эта константа одна и та же для каждого неисправного элемента и равна p, и произвольными константными, если эта константа может быть равна как 0, так и 1 для каждого неисправного элемента независимо от неисправностей других элементов. Для удобства над буквой D после символов, обозначающих базис, через точку с запятой будем ставить символы «0, 1», «0» или «1» в случаях, когда в схемах допускаются соответственно произвольные константные неисправности, однотипные константные неисправности типа 0 или типа 1 на выходах элементов. Вполне разумно предполагать, что если в базисе содержится булева константа α, то у элемента, её реализующего, не может быть неисправности типа α.

В работе С. М. Редди [5] для базиса Жегалкина $B_1 = \{\&, \oplus, 1, 0\}$ получена оценка $D^{B_1;\,0,1}_{\mathrm{s,detect}}(n) \leqslant n + 3$. В дальнейшем этот результат был обобщён С. С. Колядой в [6] на случай произвольного функционально полного конечного базиса. Последний результат, в свою очередь, был усилен Д. С. Романовым, который в [7] для любого функционально полного базиса B получил оценку $D^{B;\,0,1}_{\mathrm{s,detect}}(n) \leqslant 4$ (правда, в [7] используется несколько другое определение неизбыточных схем). Для полных проверяющих тестов Н. П. Редькин в [8, 9] для любого полного конечного базиса B_2 получил оценку $D^{B_2;\,0,1}_{\mathrm{c,detect}}(n) \leqslant 2\left(2^{\lfloor n/2\rfloor} + 2^{\lceil n/2\rceil} + n\right)$. Д. С. Романов в [10] доказал, что существует базис B_3, содержащий функциональные элементы с числом входов от одного до семи, в котором $2 \leqslant D^{B_3;\,0,1}_{\mathrm{c,detect}}(n) \leqslant 4$. В [4, с. 113, теорема 9] с использованием идей С. В. Яблонского установлено, что для любого полного базиса B функция $D^{B;\,0,1}_{\mathrm{s,diagn}}(n)$ асимптотически не превосходит $2^{n+1}/n$; аналогично можно показать, что $D^{B;\,p}_{\mathrm{s,diagn}}(n) \lesssim 2^n/n$, $p = 0, 1$. Для базиса $B_4 = \{\&, \vee, \neg\}$ Н. П. Редькин в [11, 12] получил оценки $D^{B_4;\,p}_{\mathrm{c,detect}}(n) \leqslant n$ и $D^{B_4;\,p}_{\mathrm{s,diagn}}(n) \leqslant 2n + 1$ для $p = 0, 1$. Первая из этих оценок впоследствии была улучшена Ю. В. Бородиной, которая в [13] установила, что $D^{B_4;\,p}_{\mathrm{c,detect}}(n) = 2$; вторая оценка улучшена в [14], где, в частности, доказано, что $D^{B_4;\,p}_{\mathrm{s,diagn}}(n) = 2$. Ю. В. Бородина в базисе $B_5 = \{|\}$ (штрих Шеффера) получила оценку $D^{B_5;\,1}_{\mathrm{c,detect}}(n) \leqslant n + 1$ [15]. Ей же в базисе Жегалкина B_1 удалось найти точное значение функций Шеннона $D^{B_1;\,1}_{\mathrm{s,detect}}(n) = 1$ [16] и $D^{B_1;\,0}_{\mathrm{c,detect}}(n) = 1$ [17] (совместно с П. А. Бородиным). В работе [18], в частности, установлено равенство $D^{B_1;\,0}_{\mathrm{s,diagn}}(n) = 2$.

Определим множество булевых функций $\hat{B}_0 = \{\overline{x_1}, x_1\& \ldots \&x_m : m \geqslant 2\}$. В данной работе рассматриваются только единичные проверяющие тесты, в качестве неисправ-

ностей функциональных элементов — однотипные константные неисправности типа p ($p \in \{0,1\}$) на выходах элементов, а в качестве базиса — произвольное функционально полное подмножество B_0 множества $\hat{B}_0$, например множество $\{\overline{x_1}, x_1\&x_2\}$. Для краткости вместо $D^{B_0;\,p}_{\mathrm{s,detect}}(f)$ и $D^{B_0;\,p}_{\mathrm{s,detect}}(n)$ будем писать соответственно $D_p(f)$ и $D_p(n)$.

Любой элемент, реализующий функцию вида $x_1\&\ldots\&x_m$, где $m \geqslant 2$ (функцию $\overline{x_1}$), будем называть *конъюнктором* (соответственно *инвертором*). Введём обозначения $\tilde{0}^r = \underbrace{0,\ldots,0}_{r}$, $\tilde{1}^r = \underbrace{1,\ldots,1}_{r}$, где $r \in \mathbb{Z}^+$. Будем говорить, что функциональный элемент E' расположен в схеме S *выше* (*ниже*) функционального элемента E, если в ней существует ориентированный путь от E' к E (соответственно от E к E').

Ограничимся рассмотрением только таких схем из функциональных элементов, в которых выход каждого элемента, не являющегося выходным, соединён хотя бы с одним входом хотя бы одного элемента, т. е. нет «висячих» элементов. В противном случае все «висячие» элементы схемы можно удалить, и это никак не скажется на функции, реализуемой схемой, множестве функций неисправности данной схемы и её неизбыточности, если исходная схема неизбыточна. Кроме того, будем считать, что каждый функциональный элемент, реализующий некоторую булеву функцию из базиса B_0, имеет столько входов, сколько существенных переменных у этой функции. Это предположение также не ограничивает общности.

1. Вспомогательные утверждения

Пусть S — произвольная схема из функциональных элементов. Произвольный элемент E этой схемы будем называть *разделяющим*, если любая цепочка, соединяющая любой элемент, расположенный в этой схеме выше элемента E, с выходом схемы S, проходит через элемент E.

Лемма 1. Пусть S — произвольная схема из функциональных элементов, неизбыточная относительно однотипных константных неисправностей типа 0 или 1 на выходах элементов; T — произвольный ЕПТ для этой схемы; E — произвольный разделяющий элемент схемы S; S' — подсхема схемы S, получающаяся из неё удалением всех элементов, расположенных в схеме S не выше элемента E, кроме него самого, и переносом выхода схемы на выход элемента E. Тогда схема S' неизбыточна относительно неисправностей такого же типа и множество T является для неё ЕПТ.

Доказательство. Пусть E' — произвольный элемент, содержащийся в схеме S', тогда он содержится и в схеме S. Так как T — ЕПТ для неизбыточной схемы S, то в T найдётся набор $\tilde{\sigma}$, на котором значение на выходе схемы S при неисправности элемента E' изменится. Тогда на этом наборе при указанной неисправности изменится и значение на выходе элемента E. Действительно, в противном случае переход элемента E' в неисправное состояние никак не отразился бы на значении, выдаваемом схемой S на наборе $\tilde{\sigma}$, что невозможно. Таким образом, при указанной неисправности изменится значение на выходе схемы S', откуда следует, что эта схема неизбыточна и множество T является для неё ЕПТ. ■

Лемма 2. Пусть S — произвольная схема из функциональных элементов в базисе B_0, неизбыточная относительно однотипных константных неисправностей типа 1 на выходах элементов; в S содержится хотя бы один элемент и вход любого инвертора соединён с одним из входов схемы; T — произвольный ЕПТ для схемы S; $x_{i_1},\ldots,x_{i_d}$ — все попарно различные переменные, подаваемые в S на входы инверторов (если таких переменных нет, полагаем $d = 0$); $x_{i_{d+1}},\ldots,x_{i_k}$ — все попарно различ-

ные переменные, каждая из которых подаётся на вход хотя бы одного конъюнктора (если таких переменных нет, полагаем $k = d$; множества индексов $\{i_1, \ldots, i_d\}$ и $\{i_{d+1}, \ldots, i_k\}$ могут пересекаться). Тогда $k \geqslant 1$, на выходе схемы S реализуется функция $\overline{x_{i_1}}\& \ldots \&\overline{x_{i_d}}\&x_{i_{d+1}}\& \ldots \&x_{i_k}$ (в случае $d = 0$ полагаем $\overline{x_{i_1}}\& \ldots \&\overline{x_{i_d}} \equiv 1$; в случае $k = d$ полагаем $x_{i_{d+1}}\& \ldots \&x_{i_k} \equiv 1$) и $|T| \geqslant d$.

Доказательство. Неравенство $k \geqslant 1$ очевидно, так как хотя бы одна переменная обязана подаваться на входы произвольного «верхнего» элемента схемы S. Все инверторы схемы S и все её входы, отвечающие переменным $x_{i_{d+1}}, \ldots, x_{i_k}$, соединяются со входами некоторой подсхемы, состоящей из одних конъюнкторов (эта подсхема содержит все конъюнкторы схемы S). Поэтому на выходе схемы S реализуется функция $f = \overline{x_{i_1}}\& \ldots \&\overline{x_{i_d}}\&x_{i_{d+1}}\& \ldots \&x_{i_k}$.

Докажем неравенство $|T| \geqslant d$. При $d = 0$ оно очевидно. Пусть $d \geqslant 1$. Если какая-то из переменных $x_{i_1}, \ldots, x_{i_d}$ подаётся в схеме S на входы хотя бы двух инверторов, то при неисправности любого из них функция неисправности этой схемы равна $1\&\overline{x_{i_1}}\& \ldots \&\overline{x_{i_d}}\&x_{i_{d+1}}\& \ldots \&x_{i_k} \equiv f$, т. е. схема S избыточна, что невозможно. Поэтому каждая из переменных $x_{i_1}, \ldots, x_{i_d}$ подаётся в схеме S на вход только одного инвертора (и других инверторов в этой схеме нет). Тогда при неисправности инвертора, на вход которого подаётся переменная x_{i_j}, $j \in \{1, \ldots, d\}$, функция неисправности схемы S равна $g_j = \overline{x_{i_1}}\& \ldots \&\overline{x_{i_{j-1}}}\&\overline{x_{i_{j+1}}}\& \ldots \&\overline{x_{i_d}}\&x_{i_{d+1}}\& \ldots \&x_{i_k}$. Заметим, что $g_j \not\equiv f$, так как схема S неизбыточна, и функцию g_j можно отличить от функции f только на тех наборах, у которых компоненты с номерами $i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_d$ равны нулю, а с номерами $i_j, i_{d+1}, i_{d+2}, \ldots, i_k$ — единице. Очевидно, что при $j = 1, \ldots, d$ множества этих наборов попарно не пересекаются. В тест T должно входить хотя бы по одному набору из каждого из этих d множеств, т. е. хотя бы d наборов, откуда $|T| \geqslant d$. ■

2. Единичные проверяющие тесты при однотипных константных неисправностях типа 1 на выходах элементов

Выделим три возможных представления функции $f(\tilde{x}^n)$:

$$f(\tilde{x}^n) = x_i; \tag{1}$$

$$f(\tilde{x}^n) = 0,\ \overline{x_i} \text{ или } x_{i_1}^{\sigma_1}\&x_{i_2}\& \ldots \&x_{i_k}; \tag{2}$$

$$f(\tilde{x}^n) = \overline{x_{i_1}}\&\overline{x_{i_2}}\&x_{i_3}\& \ldots \&x_{i_k} \text{ или } \underbrace{(\ldots((}_{k-1}x_{i_1}^{\sigma_1}\&x_{i_2})^{\sigma_2}\&x_{i_3})^{\sigma_3}\& \ldots \&x_{i_k})^{\sigma_k}, \tag{3}$$

где $2 \leqslant k \leqslant n$; $i, i_1, \ldots, i_k \in \{1, \ldots, n\}$, индексы $i_1, \ldots, i_k$ попарно различны и $\sigma_1, \ldots, \sigma_k \in \{0, 1\}$, причём в представлении (3) хотя бы одно из чисел $\sigma_2, \ldots, \sigma_k$ равно 0 и если $k = 2$, то полагаем $x_{i_3}\& \ldots \&x_{i_k} \equiv 1$.

Теорема 1. Для любой булевой функции $f(\tilde{x}^n)$, отличной от тождественной единицы, справедливо равенство

$$D_1(f) = \begin{cases} 0, & \text{если функция } f \text{ представима в виде (1)}, \\ 1, & \text{если функция } f \text{ представима в виде (2)}, \\ 2, & \text{если функция } f \text{ представима в виде (3)}, \\ 3, & \text{если функция } f \text{ непредставима в видах (1)–(3)}. \end{cases}$$

Если $f \equiv 1$, то значение $D_1(f)$ не определено.

Следствие 1. Для любого $n \geqslant 2$ справедливо равенство $D_1(n) = 3$.

Для доказательства следствия 1 достаточно заметить, что функция $x_1 \oplus \ldots \oplus x_n$ при $n \geqslant 2$ непредставима в видах (1)–(3).

Доказательство теоремы 1. Вместо $D_1(f)$ для краткости будем писать $D(f)$. Будем считать, что в базисе B_0 содержится функция $x_1 \& x_2$; в противном случае можно отождествить все входы, кроме одного, у произвольного конъюнктора из этого базиса и получить двухвходовой конъюнктор, допускающий те же самые неисправности (а именно неисправность типа 1 на выходе), что и исходный конъюнктор. Пусть сначала $f \equiv 1$; S — произвольная схема в базисе B_0, реализующая функцию f. Выход схемы S, очевидно, не может совпадать ни с одним из её входов, поэтому он является выходом некоторого функционального элемента. Тогда при неисправности этого элемента получающаяся схема по-прежнему будет реализовывать тождественную единицу, т. е. схема S избыточна. Получаем, что неизбыточных схем, реализующих функцию f, не существует и, следовательно, значение $D(f)$ не определено. Далее будем считать, что $f \not\equiv 1$. Рассмотрим четыре случая:

1. Функция f представима в виде (1). Тогда её можно реализовать схемой, не содержащей функциональных элементов. У такой схемы нет ни одной функции неисправности, поэтому пустое множество является для неё ЕПТ, откуда следует равенство $D(f) = 0$.

2. Функция f представима в виде (2). Поскольку она непредставима в виде (1), выход любой схемы, реализующей функцию f, не может совпадать ни с одним из её входов, поэтому он является выходом некоторого функционального элемента. Тогда при неисправности этого элемента схема реализует тождественную единицу, которую надо отличить от функции f хотя бы на одном наборе, откуда следует, что $D(f) \geqslant 1$. Докажем неравенство $D(f) \leqslant 1$. Рассмотрим три подслучая:

2.1. Пусть $f \equiv 0$. Реализуем функцию f схемой S в базисе B_0, содержащей один инвертор и один двухвходовой конъюнктор. На вход инвертора подадим переменную x_1, его выход соединим с левым входом конъюнктора, а на правый вход подадим переменную x_1. Очевидно, что схема S реализует тождественный нуль и имеет две функции неисправности — тождественную единицу и x_1. Каждую из них можно отличить от функции f на наборе $(\tilde{1}^n)$, поэтому схема S неизбыточна и множество $\{(\tilde{1}^n)\}$ является для неё ЕПТ длины 1, откуда $D(f) \leqslant 1$.

2.2. Пусть $f = \overline{x_i}$, где $i \in \{1, \ldots, n\}$. Реализуем функцию f схемой в базисе B_0, содержащей один инвертор. Эта схема имеет только одну функцию неисправности — тождественную единицу, которую можно отличить от функции f на наборе $(\tilde{1}^n)$. Поэтому схема неизбыточна и $D(f) \leqslant 1$.

2.3. Пусть $f = x_{i_1}^{\sigma_1} \& x_{i_2} \& \ldots \& x_{i_k}$, где $2 \leqslant k \leqslant n$; $i_1, \ldots, i_k$ — попарно различные индексы из множества $\{1, \ldots, n\}$ и $\sigma_1 \in \{0, 1\}$. Без ограничения общности можно считать, что $i_1 = 1$, $\ldots$, $i_k = k$, т. е. $f(\tilde{x}^n) = x_1^{\sigma_1} \& x_2 \& \ldots \& x_k$. Реализуем функцию f схемой S в базисе B_0, представляющей собой цепочку из одного инвертора (в случае $\sigma_1 = 0$) и $k - 1$ двухвходовых конъюнкторов, на правые входы которых подаются переменные $x_2, \ldots, x_k$. Если $\sigma_1 = 0$, то верхний элемент этой цепочки — инвертор, на вход которого подаётся переменная x_1, а его выход соединяется с левым входом верхнего конъюнктора; если $\sigma_1 = 1$, то верхний элемент этой цепочки — конъюнктор, на левый вход которого подаётся переменная x_1.

Очевидно, что схема S реализует функцию f. Докажем, что она неизбыточна и множество $\{\tilde{\pi}\}$ является для неё ЕПТ, где $\tilde{\pi} = (\overline{\sigma_1}, \tilde{1}^{n-1})$. Заметим, что $f(\tilde{\pi}) = 0$, причём на правые входы всех конъюнкторов в схеме S на наборе $\tilde{\pi}$ подаются единицы.

Если неисправен некоторый элемент в схеме, то на указанном наборе на выходе этого и всех следующих за ним элементов в схеме S, а значит, и на выходе схемы появятся единицы и неисправность будет обнаружена. Поэтому схема S неизбыточна и $D(f) \leqslant 1$.

В итоге для любой функции f вида (2) получаем равенство $D(f) = 1$.

3. Функция f представима в виде (3). Докажем сначала неравенство $D(f) \leqslant 2$. Рассмотрим два подслучая:

3.1. Пусть $f = \overline{x_{i_1}}\&\overline{x_{i_2}}\&x_{i_3}\&\ldots\&x_{i_k}$, где $2 \leqslant k \leqslant n$ и $i_1, \ldots, i_k$ — попарно различные индексы из множества $\{1, \ldots, n\}$. Без ограничения общности можно считать, что $i_1 = 1, \ldots, i_k = k$, т. е. $f(\tilde{x}^n) = \overline{x_1}\&\overline{x_2}\&x_3\&\ldots\&x_k$. Реализуем функцию f схемой S в базисе B_0, содержащей два инвертора и $k-1$ двухвходовых конъюнкторов. На входы инверторов подадим переменные x_1 и x_2; затем выходы этих инверторов и входы схемы, отвечающие переменным $x_3, \ldots, x_k$, последовательно соединим со входами цепочки из $k-1$ конъюнкторов: входы верхнего конъюнктора из цепочки соединим с выходами инверторов; на правые входы остальных $k-2$ конъюнкторов подадим переменные $x_3, \ldots, x_k$.

Очевидно, что схема S реализует функцию f. Докажем, что она неизбыточна и множество $\{\tilde{\pi}_1, \tilde{\pi}_2\}$ является для неё ЕПТ, где $\tilde{\pi}_1 = (0, \tilde{1}^{n-1})$, $\tilde{\pi}_2 = (1, 0, \tilde{1}^{n-2})$. Заметим, что $f(\tilde{\pi}_1) = f(\tilde{\pi}_2) = 0$, причём на один из входов верхнего конъюнктора и на правые входы всех остальных конъюнкторов в схеме S на каждом из наборов $\tilde{\pi}_1, \tilde{\pi}_2$ подаются единицы. Если неисправен некоторый элемент, то по крайней мере на одном из указанных наборов на выходе этого и всех следующих за ним элементов, а значит, и на выходе схемы появятся единицы и неисправность будет обнаружена. Поэтому схема S неизбыточна и $D(f) \leqslant 2$.

3.2. Пусть $f = \underbrace{(\ldots((}_{k-1} x_{i_1}^{\sigma_1}\&x_{i_2})^{\sigma_2}\&x_{i_3})^{\sigma_3}\&\ldots\&x_{i_k})^{\sigma_k}$, где $2 \leqslant k \leqslant n$; $i_1, \ldots, i_k$ — попарно различные индексы из множества $\{1, \ldots, n\}$ и $\sigma_1, \ldots, \sigma_k \in \{0, 1\}$, причём хотя бы одно из чисел $\sigma_2, \ldots, \sigma_k$ равно 0. Без ограничения общности можно считать, что $i_1 = 1, \ldots, i_k = k$, т. е. $f(\tilde{x}^n) = \underbrace{(\ldots((}_{k-1} x_1^{\sigma_1}\&x_2)^{\sigma_2}\&x_3)^{\sigma_3}\&\ldots\&x_k)^{\sigma_k}$. Реализуем функцию f схемой S в базисе B_0, представляющей собой цепочку из инверторов и конъюнкторов, в соответствии с последним равенством. В этой схеме содержатся $k-1$ двухвходовых конъюнкторов и столько инверторов, сколько чисел из $\sigma_1, \ldots, \sigma_k$ равны 0. Если $\sigma_1 = 1$, то верхний элемент схемы S — двухвходовой конъюнктор, на входы которого подаются переменные x_1 и x_2. Если $\sigma_1 = 0$, то верхний элемент схемы S — инвертор, на вход которого подаётся переменная x_1; выход инвертора соединяется с левым входом конъюнктора, на правый вход которого подаётся переменная x_2. Если $\sigma_2 = 0$, то выход указанного конъюнктора соединяется со входом инвертора. Далее, если $k \geqslant 3$, то выход последнего построенного элемента соединяется с левым входом конъюнктора, на правый вход которого подаётся переменная x_3, и т. д.

Очевидно, что схема S реализует функцию f. Докажем, что она неизбыточна и множество $T = \{\tilde{\pi}_1, \tilde{\pi}_2\}$ является для неё ЕПТ, где $\tilde{\pi}_1 = (0, \tilde{1}^{n-1})$, $\tilde{\pi}_2 = (\tilde{1}^n)$. Заметим, что $f(x_1, \tilde{1}^{n-1}) = \underbrace{(\ldots((}_{k-1} x_1^{\sigma_1})^{\sigma_2})^{\sigma_3}\ldots)^{\sigma_k}$, поэтому

$$f(\tilde{\pi}_1) \neq f(\tilde{\pi}_2). \tag{4}$$

Единственная цепь, связывающая вход x_1 схемы S с её выходом, проходит через каждый элемент. Поэтому если в схеме какой-то элемент неисправен, то при подаче на её

входы вместо переменных $x_2, \ldots, x_n$ константы 1 изменение значения переменной x_1 с 0 на 1, т. е. переход от входного набора $\tilde{\pi}_1$ к набору $\tilde{\pi}_2$, никак не отразится на значении, выдаваемом схемой S. Следовательно, функция неисправности g схемы S удовлетворяет условию $g(\tilde{\pi}_1) = g(\tilde{\pi}_2)$. Отсюда и из (4) вытекает, что $g \not\equiv f$, т. е. схема S неизбыточна и множество T является для неё ЕПТ, поэтому $D(f) \leqslant 2$.

Докажем теперь, что $D(f) \geqslant 2$. Пусть S — произвольная неизбыточная схема, реализующая функцию f вида (3); T — произвольный ЕПТ для этой схемы. Надо доказать, что $|T| \geqslant 2$. Рассмотрим два подслучая:

3.1′. В схеме S найдётся хотя бы один инвертор I, вход которого соединён с выходом некоторого функционального элемента E. Пусть на выходе элемента E в схеме S реализуется булева функция φ, тогда на выходе элемента I реализуется функция $\overline{\varphi}$. Чтобы обнаружить неисправность элемента E (I), в тесте T должен содержаться набор $\tilde{\sigma}_1$ ($\tilde{\sigma}_2$), для которого $\varphi(\tilde{\sigma}_1) = 0$ (соответственно $\overline{\varphi}(\tilde{\sigma}_2) = 0$). Из двух последних равенств следует, что $\tilde{\sigma}_1 \neq \tilde{\sigma}_2$, поэтому $|T| \geqslant 2$, что и требовалось доказать.

3.2′. Вход любого инвертора схемы S соединён с одним из входов схемы. Пусть $x_{i_1}, \ldots, x_{i_d}$ — все попарно различные переменные, подаваемые на входы инверторов (если таких переменных нет, полагаем $d = 0$), а $x_{i_{d+1}}, \ldots, x_{i_k}$ — все попарно различные переменные, каждая из которых подаётся на вход хотя бы одного конъюнктора (если таких переменных нет, полагаем $k = d$). Тогда в силу леммы 2 имеем $k \geqslant 1$, $f = \overline{x_{i_1}} \& \ldots \& \overline{x_{i_d}} \& x_{i_{d+1}} \& \ldots \& x_{i_k}$ (в случае $d = 0$ полагаем $\overline{x_{i_1}} \& \ldots \& \overline{x_{i_d}} \equiv 1$; в случае $k = d$ полагаем $x_{i_{d+1}} \& \ldots \& x_{i_k} \equiv 1$) и $|T| \geqslant d$. Так как функция f непредставима в видах (1), (2), то $d \geqslant 2$, следовательно, $|T| \geqslant 2$, что и требовалось доказать.

В итоге для любой функции f вида (3) получаем равенство $D(f) = 2$.

4. Функция f отлична от тождественной единицы и непредставима в видах (1)–(3). Докажем сначала неравенство $D(f) \leqslant 3$. Идеи доказательства сходны с идеями, использованными в [6, 7] при получении верхних оценок длин ЕПТ в базисе $\{\&, \neg\}$. Представим функцию f полиномом Жегалкина

$$f(\tilde{x}^n) = K_1 \oplus \ldots \oplus K_m \oplus c, \tag{5}$$

где $c \in \{0, 1\}$ и K_1 — самая короткая конъюнкция в полиноме. Без ограничения общности $K_1 = x_1 \& \ldots \& x_k$. Отметим, что $k < n$, так как в противном случае функция f была бы равна $x_1 \& \ldots \& x_n \oplus c = (x_1 \& \ldots \& x_n)^{\overline{c}}$, т. е. имела бы вид (2) или (3). По аналогичной причине $m \geqslant 2$. Если в полиноме Жегалкина для функции f присутствует слагаемое $x_1 \& \ldots \& x_k \& x_{k+1}$, то будем без ограничения общности считать, что это слагаемое K_m. Представим K_1 в виде $K_1' \oplus K_{m+1}$, где $K_1' = x_1 \& \ldots \& x_k \& \overline{x_{k+1}}$, $K_{m+1} = x_1 \& \ldots \& x_k \& x_{k+1}$. Тогда

$$f = K_1' \oplus K_2 \oplus \ldots \oplus K_m \oplus K_{m+1} \oplus c = K_1' \oplus K_2 \oplus \ldots \oplus K_q \oplus c, \tag{6}$$

где $q = m - 1$ при $K_m = K_{m+1}$ и $q = m + 1$ при $K_m \neq K_{m+1}$. Отметим, что случай $m = 2$, $K_m = K_{m+1}$ невозможен, так как тогда функция f была бы равна $K_1' \oplus c = (\overline{x_{k+1}} \& x_1 \ldots \& x_k)^{\overline{c}}$, т. е. была бы представима в виде (2) или (3). Поэтому $q \geqslant 2$.

Так как K_1 — самая короткая конъюнкция в полиноме Жегалкина для функции f, в каждую конъюнкцию $K_i = x_{j_1(i)} \& \ldots \& x_{j_{t_i}(i)}$, где $i = 2, \ldots, q$, входит хотя бы одна переменная, отличная от переменных $x_1, \ldots, x_k$. Без ограничения общности это переменная $x_{j_1(i)}$. В случае $t_i \geqslant 3$ представим конъюнкцию K_i, $i = 2, \ldots, q$, в виде $(K_i^{(2)} \oplus \ldots \oplus K_i^{(t_i-1)} \oplus K_i^{(t_i)}) \oplus (K_i^{(2)} \oplus \ldots \oplus K_i^{(t_i-1)})$, где $K_i^{(s)} = x_{j_1(i)} \& \ldots \& x_{j_s(i)}$, $s = 2, \ldots, t_i$

(конъюнкция первых s переменных из K_i). Тогда представление (6) примет вид

$$f = K_1' \oplus \bigoplus_{i=2}^{q} \left(\bigoplus_{s=2}^{t_i} K_i^{(s)} \oplus \bigoplus_{s=2}^{t_i-1} K_i^{(s)} \right) \oplus c \tag{7}$$

(в случае $t_i \in \{1, 2\}$ полагаем $\bigoplus\limits_{s=2}^{t_i} K_i^{(s)} \oplus \bigoplus\limits_{s=2}^{t_i-1} K_i^{(s)} = K_i$).

Наконец, заменим в представлении (7) все операции $\oplus$ на $\oplus'$, где по определению полагаем $x \oplus' y = x \oplus y \oplus 1$. Эта замена приведёт к прибавлению к правой части представления (7) некоторого числа единиц (по одному на каждую операцию $\oplus$). Сложив все эти единицы, а также константу c по модулю 2, получим некоторую булеву константу c', такую, что

$$f = K_1' \oplus' \bigoplus_{i=2}^{q}{}' \left(\bigoplus_{s=2}^{t_i}{}' K_i^{(s)} \oplus' \bigoplus_{s=2}^{t_i-1}{}' K_i^{(s)} \right) \oplus c'. \tag{8}$$

Пусть $S_{\oplus'}$ — схема в базисе B_0 с двумя входами и одним выходом, изображённая на рис. 1 (все конъюнкторы двухвходовые). Нетрудно проверить, что на выходе этой схемы реализуется функция $x \oplus' y$, её всевозможными функциями неисправности являются функции $1, 0, x \vee \overline{y}, \overline{x} \vee y$, $\overline{x}, \overline{y}, xy$ (и, следовательно, схема $S_{\oplus'}$ неизбыточна), а в качестве ЕПТ для неё можно взять множество $T_{\oplus'} = \{(0, 0), (0, 1), (1, 0)\}$ (действительно, единственной булевой функцией от двух переменных, которую нельзя отличить от функции $x \oplus' y$ на наборах из множества $T_{\oplus'}$, кроме неё самой, является функция $\overline{x}\,\overline{y}$, но она не входит в число функций неисправности схемы $S_{\oplus'}$).

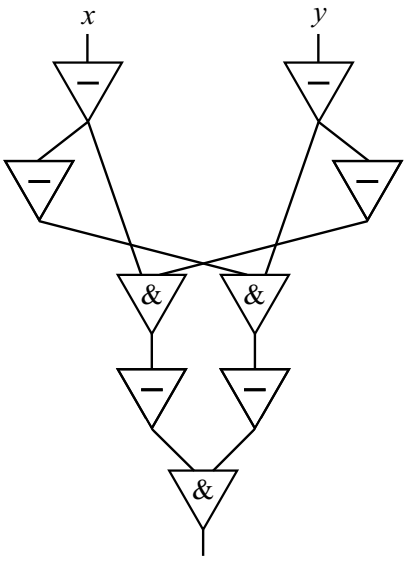

Рис. 1. Схема $S_{\oplus'}$

Реализуем функцию $f(\tilde{x}^n)$ схемой S в базисе B_0 в соответствии с представлением (8) (рис. 2). Конъюнкцию K_1' реализуем цепочкой Z_1 из одного инвертора и k двухвходовых конъюнкторов, верхним элементом в которой является инвертор; на вход инвертора подадим переменную x_{k+1}, а на правые входы конъюнкторов — последовательно переменные $x_1, \ldots, x_k$. Каждую конъюнкцию K_i, $i = 2, \ldots, q$, реализуем цепочкой Z_i из $t_i - 1$ двухвходовых конъюнкторов, на входы которой последовательно подадим переменные $x_{j_1(i)}, x_{j_2(i)} \ldots, x_{j_{t_i}(i)}$ (в случае $t_i = 1$ в этой цепочке не содержится элементов, а

её выход совпадает с входом $x_{j_1(i)}$ схемы S). Далее, для каждого $i \in \{2, \ldots, q\}$, такого, что $t_i \geqslant 3$, реализуем конъюнкцию $K_i^{(t_i-1)}$ цепочкой $\hat{Z}_i$ из t_i-2 двухвходовых конъюнкторов, на входы которой последовательно подадим переменные $x_{j_1(i)}, x_{j_2(i)} \ldots, x_{j_{t_i-1}(i)}$. Затем выход цепочки Z_1, выходы всех цепочек Z_i, в которых не содержится элементов, и выходы всех конъюнкторов, содержащихся в цепочках $Z_2, \ldots, Z_q$, $\hat{Z}_2, \ldots, \hat{Z}_q$ (точнее, в тех из них, которые были определены), соединим со входами цепочки $Z_{\oplus'}$, состоящей из блоков $S_{\oplus'}$ и (в случае $c' = 1$) инвертора, вход которого соединён с выходом нижнего из этих блоков, причём левый верхний вход этой цепочки соединим с выходом цепочки Z_1. Выход нижнего элемента цепочки $Z_{\oplus'}$ объявим выходом схемы S.

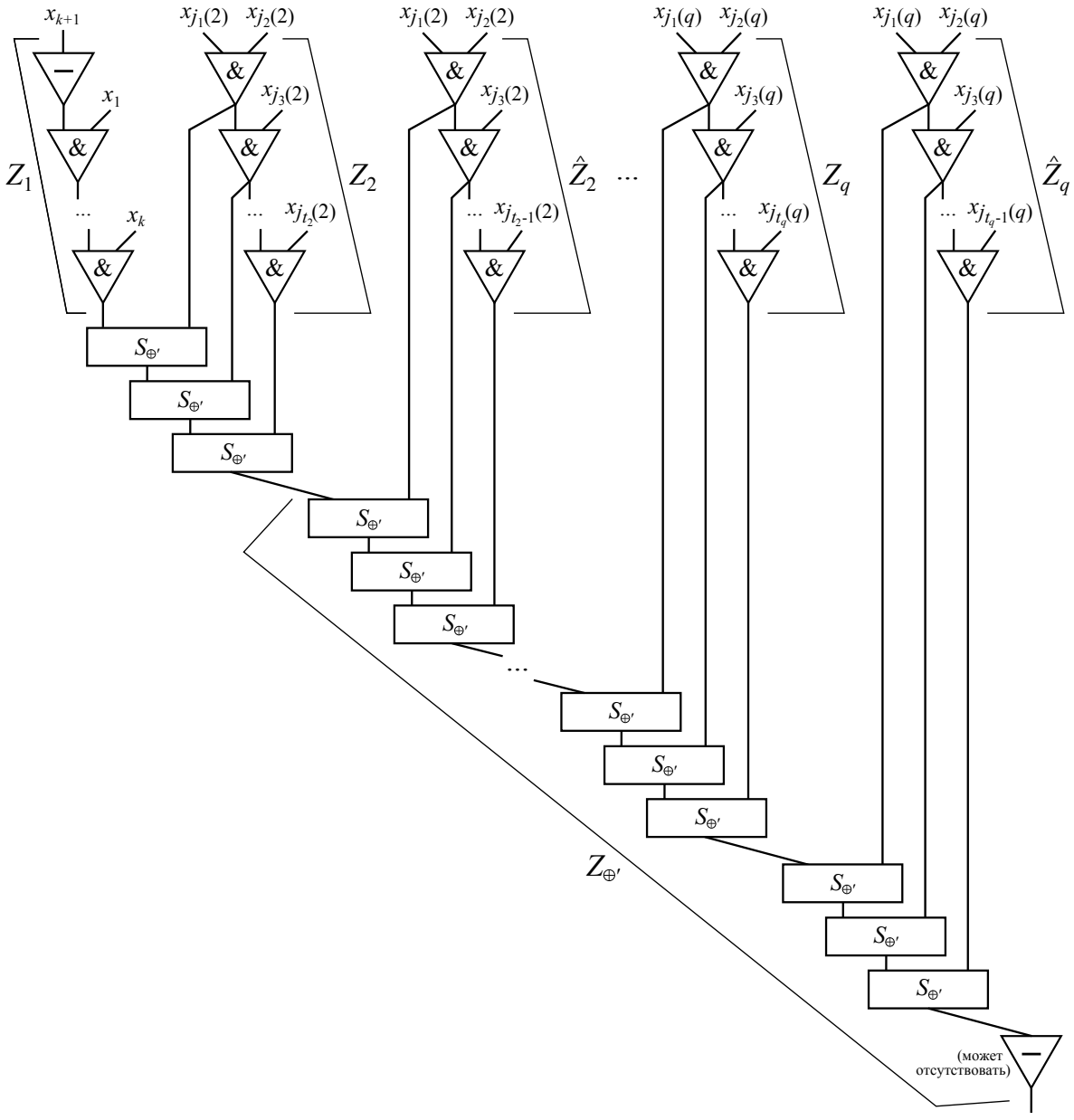

Рис. 2. Схема S

Нетрудно убедиться, что построенная схема S реализует функцию $f(\tilde{x}^n)$. Для этого достаточно заметить, что на выходах конъюнкторов, содержащихся в цепочке Z_i, $i = 2, \ldots, q$, при движении по этой цепочке «сверху вниз» реализуются функции

$K_i^{(2)}, \ldots, K_i^{(t_i)}$ (при $t_i \geqslant 2$), а на выходах конъюнкторов, содержащихся в цепочке $\hat{Z}_i$, $i = 2, \ldots, q$, — функции $K_i^{(2)}, \ldots, K_i^{(t_i-1)}$ (при $t_i \geqslant 3$).

Докажем, что схема S неизбыточна и множество $T = \{\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3\}$ является для неё ЕПТ, где $\tilde{\sigma}_1 = (\tilde{0}^n)$, $\tilde{\sigma}_2 = (\tilde{1}^k, \tilde{0}^{n-k})$, $\tilde{\sigma}_3 = (\tilde{1}^n)$. В случае исправности всех элементов схемы S на наборе $\tilde{\sigma}_1$ на выходах всех конъюнкторов, содержащихся в цепочках $Z_i, \hat{Z}_i$, где $i = 2, \ldots, q$, появятся нули. Если неисправен некоторый конъюнктор в одной из цепочек, то на наборе $\tilde{\sigma}_1$ на выходе этого конъюнктора будет единица, а на выходах всех остальных конъюнкторов — по-прежнему нули. Учитывая, что выход неисправного конъюнктора соединяется ровно с одним входом цепочки $Z_{\oplus'}$, реализующей линейную функцию от своих входов, значение на выходе данной цепочки, т. е. на выходе всей схемы S, изменится, поэтому рассматриваемая неисправность будет обнаружена на наборе $\tilde{\sigma}_1$.

Далее, в случае исправности всех элементов схемы S на наборе $\tilde{\sigma}_3$ на выходах всех элементов, содержащихся в цепочке Z_1, будут нули, причём на правые входы всех конъюнкторов в цепочке подаются единицы. Если неисправен некоторый элемент в цепочке, то на наборе $\tilde{\sigma}_3$ на выходе этого и всех следующих за ним элементов в цепочке Z_1, т. е. на выходе этой цепочки, очевидно, возникнут единицы. Учитывая, что выход цепочки Z_1 соединяется ровно с одним входом цепочки $Z_{\oplus'}$, реализующей линейную функцию от своих входов, значение на выходе данной цепочки, т. е. на выходе всей схемы S, изменится, поэтому рассматриваемая неисправность будет обнаружена на наборе $\tilde{\sigma}_3$.

Осталось рассмотреть неисправность некоторого элемента в цепочке $Z_{\oplus'}$. В случае исправности всех элементов схемы S на наборах $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ на выходе цепочки Z_1 возникают значения соответственно $0, 1, 0$, а на выходах всех цепочек Z_i, в которых не содержится элементов, и выходах всех конъюнкторов, содержащихся в цепочках $Z_i, \hat{Z}_i$, где $i = 2, \ldots, q$, — соответственно $0, 0, 1$ (здесь используется то свойство, что каждая переменная $x_{j_1(i)}$ отлична от переменных $x_1, \ldots, x_k$). В таком случае на входы верхнего блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ подаются наборы $(0,0), (1,0), (0,1)$, а на его выходе реализуются значения $0 \oplus' 0, 1 \oplus' 0, 0 \oplus' 1$, т. е. $1, 0, 0$. Тогда на входы второго сверху блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ (если он существует) подаются наборы $(1,0), (0,0), (0,1)$, а на его выходе реализуются значения $1 \oplus' 0, 0 \oplus' 0, 0 \oplus' 1$, т. е. $0, 1, 0$. Далее, на входы третьего сверху блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ (если он существует) подаются наборы $(0,0), (1,0), (0,1)$, а на его выходе реализуются значения $0 \oplus' 0, 1 \oplus' 0, 0 \oplus' 1$, т. е. $1, 0, 0$, и т.д. Таким образом, на входах каждого блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ при подаче на входы схемы S наборов $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ возникают все наборы из множества $T_{\oplus'}$. Так как это множество является ЕПТ для неизбыточной схемы $S_{\oplus'}$, то при неисправности любого элемента в любом блоке $S_{\oplus'}$ цепочки $Z_{\oplus'}$ хотя бы на одном из наборов $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ значение на выходе этого блока изменится. Учитывая, что выход указанного блока либо совпадает с выходом схемы S, либо соединяется ровно с одним входом некоторой нижней части цепочки $Z_{\oplus'}$, реализующей линейную функцию от своих входов, значение на выходе схемы S изменится, поэтому рассматриваемая неисправность будет обнаружена на одном из наборов $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$.

Наконец, если неисправен выходной инвертор цепочки $Z_{\oplus'}$ (в случае $c' = 1$), то функция неисправности схемы S равна тождественной единице, которую можно отличить от функции f на одном из наборов $\tilde{\sigma}_1, \tilde{\sigma}_2$ (поскольку $f(\tilde{\sigma}_1) = c$, $f(\tilde{\sigma}_2) = \bar{c}$ — это следует из представления (5)).

В итоге получаем, что любую функцию неисправности схемы S можно отличить от функции $f(\tilde{x}^n)$ хотя бы на одном наборе из множества T. Это означает, что схема S

неизбыточна и множество T является для неё ЕПТ. Его длина равна 3, откуда следует неравенство $D(f) \leqslant 3$.

Докажем теперь, что $D(f) \geqslant 3$. Пусть S — произвольная неизбыточная схема, реализующая функцию f, отличную от тождественной единицы и непредставимую в видах (1)–(3); T — произвольный ЕПТ для этой схемы. Надо доказать, что $|T| \geqslant 3$. Рассмотрим два подслучая:

4.1. В схеме S найдётся элемент, входы которого соединены с выходами по крайней мере двух различных функциональных элементов. Среди всех элементов с таким свойством выберем произвольный «нижний» элемент E, ниже которого в схеме S не существует элемента с указанным свойством (это можно сделать, так как схема конечна и не содержит ориентированных циклов). Очевидно, что элемент E — конъюнктор, а входы каждого элемента, расположенного в схеме S ниже элемента E, соединены с выходом ровно одного функционального элемента (напомним, что мы рассматриваем только схемы без «висячих» элементов) и, как следствие, E — разделяющий элемент.

Среди всех элементов, выходы которых соединены в схеме S непосредственно со входами элемента E (а таких элементов не меньше двух), выберем произвольный «нижний» элемент E_1, ниже которого в схеме S не существует элемента с указанным свойством, и любой другой элемент E_2. Пусть в случае исправности всех элементов схемы S на выходе элемента E_1 (E_2) этой схемы реализуется булева функция φ_1 (соответственно φ_2). Тогда на выходе элемента E реализуется булева функция $\varphi_1\&\varphi_2\&\varphi_3$, где φ_3 — конъюнкция функций, подаваемых на все те входы элемента E, которые не соединены с выходом ни одного из элементов E_1, E_2, либо тождественная единица, если таких входов не существует. Рассмотрим три подслучая:

4.1.1. В схеме S найдётся хотя бы один инвертор, расположенный ниже элемента E. Пусть I — «верхний» из этих инверторов. Очевидно, что I — разделяющий элемент. По лемме 1 схема S', получающаяся из схемы S удалением всех элементов, расположенных в ней не выше элемента I, кроме него самого, и переносом выхода схемы на выход элемента I, неизбыточна и T — ЕПТ для схемы S'. Все элементы, расположенные в схеме S ниже элемента E, но выше элемента I (если такие есть) — конъюнкторы. Поэтому на вход элемента I подаётся функция $\varphi_1\&\varphi_2\&\varphi_3\&\varphi_4$, где φ_4 — некоторая булева функция. Тогда на выходе элемента I, т. е. на выходе схемы S', реализуется функция $f' = \overline{\varphi_1\&\varphi_2\&\varphi_3\&\varphi_4}$.

При неисправности инвертора I на выходе схемы S' возникнет функция неисправности $g_1 \equiv 1$. При неисправности элемента E_1 на его выходе вместо функции φ_1 возникнет тождественная единица, а функции на выходах всех остальных элементов, соединяющихся со входами элемента E, не изменятся, поскольку все эти элементы расположены в схеме S' не ниже элемента E_1 в силу его выбора. Поэтому на выходе схемы S' появится функция неисправности $g_2 = \overline{1\&\varphi_2\&\varphi_3\&\varphi_4} = \overline{\varphi_2\&\varphi_3\&\varphi_4}$.

Так как схема S' неизбыточна, каждая из функций g_1, g_2 отлична от функции f'. Чтобы отличить функцию f' от функции g_1, в множестве T должен содержаться хотя бы один набор $\tilde{\sigma}_1$, для которого $f'(\tilde{\sigma}_1) = 0$, т. е. $\varphi_1(\tilde{\sigma}_1) = \varphi_2(\tilde{\sigma}_1) = \varphi_3(\tilde{\sigma}_1) = \varphi_4(\tilde{\sigma}_1) = 1$. Чтобы отличить функцию f' от функции g_2, в множестве T должен содержаться хотя бы один набор $\tilde{\sigma}_2$, для которого $\varphi_1(\tilde{\sigma}_2) = 0$, $\varphi_2(\tilde{\sigma}_2) = \varphi_3(\tilde{\sigma}_2) = \varphi_4(\tilde{\sigma}_2) = 1$. Из равенств $\varphi_1(\tilde{\sigma}_1) = 1$, $\varphi_1(\tilde{\sigma}_2) = 0$ следует, что $\tilde{\sigma}_1 \neq \tilde{\sigma}_2$, а из равенств $\varphi_2(\tilde{\sigma}_1) = \varphi_2(\tilde{\sigma}_2) = 1$ — что неисправность элемента E_2, на выходе которого в случае исправности всех элементов схемы S' реализуется функция φ_2, нельзя обнаружить на наборах $\tilde{\sigma}_1$, $\tilde{\sigma}_2$. Поэтому в тесте T должен содержаться ещё какой-то набор, отличный от этих двух наборов. Таким образом, $|T| \geqslant 3$, что и требовалось доказать.

4.1.2. Условие случая 4.1.1 не выполнено, но при этом в схеме S найдётся хотя бы один инвертор, расположенный выше элемента E, вход которого соединён с выходом некоторого функционального элемента. Среди всех инверторов с такими свойствами выберем произвольный «нижний» инвертор I, ниже которого в схеме S не существует инвертора с указанными свойствами. Пусть Z — произвольная цепочка, соединяющая элемент I с элементом E. Будем считать, что сам элемент I в неё не входит. Очевидно тогда, что все элементы в цепочке Z — конъюнкторы. По лемме 1 схема S', получающаяся из схемы S удалением всех элементов, расположенных в ней не выше элемента E, кроме него самого, и переносом выхода схемы на выход элемента E, неизбыточна и T — ЕПТ для схемы S'. На выходе элемента E, т. е. на выходе схемы S', как было показано ранее, реализуется функция $f' = \varphi_1 \& \varphi_2 \& \varphi_3$.

При неисправности того элемента, выход которого соединён в схеме S (и, следовательно, в схеме S') со входом инвертора I, на вход I подаётся тождественная единица, а на его выходе появится тождественный нуль. Он будет подаваться на один из входов цепочки Z из конъюнкторов, поэтому на её выходе, т. е. на выходе элемента E, а значит, и схемы S', возникнет функция неисправности $g_1 \equiv 0$. Далее, при неисправности элемента E_1 на его выходе вместо функции φ_1 будет тождественная единица, а функции на выходах всех остальных элементов, соединяющихся со входами элемента E, не изменятся, поскольку все эти элементы расположены в схеме S' не ниже элемента E_1 в силу его выбора. Поэтому на выходе схемы S' возникнет функция неисправности $g_2 = 1 \& \varphi_2 \& \varphi_3 = \varphi_2 \& \varphi_3$.

Так как схема S' неизбыточна, каждая из функций g_1, g_2 отлична от функции f'. Чтобы отличить функцию f' от функции g_1, в множестве T должен содержаться хотя бы один набор $\tilde{\sigma}_1$, для которого $f'(\tilde{\sigma}_1) = 1$, т. е. $\varphi_1(\tilde{\sigma}_1) = \varphi_2(\tilde{\sigma}_1) = \varphi_3(\tilde{\sigma}_1) = 1$. Чтобы отличить функцию f' от функции g_2, в множестве T должен содержаться хотя бы один набор $\tilde{\sigma}_2$, для которого $\varphi_1(\tilde{\sigma}_2) = 0$, $\varphi_2(\tilde{\sigma}_2) = \varphi_3(\tilde{\sigma}_2) = 1$. Из равенств $\varphi_1(\tilde{\sigma}_1) = 1$, $\varphi_1(\tilde{\sigma}_2) = 0$ следует, что $\tilde{\sigma}_1 \neq \tilde{\sigma}_2$, а из равенств $\varphi_2(\tilde{\sigma}_1) = \varphi_2(\tilde{\sigma}_2) = 1$ — что неисправность элемента E_2, на выходе которого в случае исправности всех элементов схемы S' реализуется функция φ_2, нельзя обнаружить на наборах $\tilde{\sigma}_1$, $\tilde{\sigma}_2$. Поэтому в тесте T должен содержаться ещё какой-то набор. Таким образом, $|T| \geqslant 3$, что и требовалось доказать.

4.1.3. Отрицание объединения случаев 4.1.1 и 4.1.2: в схеме S ниже элемента E не расположено ни одного инвертора, а вход любого инвертора этой схемы, расположенного выше элемента E, соединён с одним из входов схемы. В таком случае вход любого инвертора схемы S соединён с одним из входов этой схемы. Пусть $x_{i_1}, \ldots, x_{i_d}$ — все попарно различные переменные, подаваемые в схеме на входы инверторов (если таких переменных нет, полагаем $d = 0$), а $x_{i_{d+1}}, \ldots, x_{i_k}$ — все попарно различные переменные, каждая из которых подаётся на вход хотя бы одного конъюнктора (если таких переменных нет, полагаем $k = d$). Тогда в силу леммы 2 имеем $k \geqslant 1$, $f = \overline{x_{i_1}} \& \ldots \& \overline{x_{i_d}} \& x_{i_{d+1}} \& \ldots \& x_{i_k}$ (в случае $d = 0$ полагаем $\overline{x_{i_1}} \& \ldots \& \overline{x_{i_d}} \equiv 1$; в случае $k = d$ полагаем $x_{i_{d+1}} \& \ldots \& x_{i_k} \equiv 1$) и $|T| \geqslant d$. Так как функция f непредставима в видах (1)–(3), то $d \geqslant 3$, следовательно, $|T| \geqslant 3$, что и требовалось доказать. Случай 4.1 разобран.

4.2. Входы каждого элемента схемы S соединены с выходом не более чем одного функционального элемента. Очевидно, что в таком случае схема S представляет собой цепочку из элементов, причём в ней содержится хотя бы один элемент, так как функция f непредставима в виде (1). Пусть в этой цепочке при движении сверху вниз расположены элементы $E_1, \ldots, E_r$, где r — общее число элементов в схеме S.

Докажем по индукции, что на выходе каждого из этих элементов реализуется функция вида

$$0, 1 \text{ или } \underbrace{(\ldots((}_{k-1} x_{i_1}^{\sigma_1} \& x_{i_2})^{\sigma_2} \& x_{i_3})^{\sigma_3} \& \ldots \& x_{i_k})^{\sigma_k}, \tag{9}$$

где $1 \leqslant k \leqslant n$; $i_1, \ldots, i_k$ — попарно различные индексы и $\sigma_1, \ldots, \sigma_k \in \{0, 1\}$ (при $k = 1$ этот вид превращается в $x_{i_1}^{\sigma_1}$). Легко видеть, что вне зависимости от того, является элемент E_1 инвертором или конъюнктором, на его выходе будет реализована функция вида (9). В случае $r = 1$ утверждение доказано. Далее будем считать, что $r \geqslant 2$. Пусть утверждение верно для элемента E_j, где $j \in \{1, \ldots, r-1\}$; докажем его для элемента E_{j+1}. На выходе элемента E_j по предположению индукции реализуется функция f_j вида (9); этот выход соединяется в схеме S с одним или несколькими входами элемента E_{j+1}, на все остальные входы которого подаются переменные. Если E_{j+1} — инвертор, то утверждение очевидно, так как отрицание любой функции вида (9) является функцией того же вида. Пусть этот элемент — конъюнктор и все попарно различные переменные, которые подаются на его входы, не соединённые с выходом элемента E_j, — это переменные $x_{i_{k+1}}, \ldots, x_{i_l}$, где $l \geqslant k$ и $i_{k+1}, \ldots, i_l$ — попарно различные индексы (если таких переменных нет, полагаем $l = k$). Тогда на выходе E_{j+1} реализуется функция $f_j \& x_{i_{k+1}} \& \ldots \& x_{i_l}$ (в случае $l = k$ полагаем $x_{i_{k+1}} \& \ldots \& x_{i_l} \equiv 1$). Справедливо тождество

$$f_j \& x_{i_{k+1}} \& \ldots \& x_{i_l} \equiv h_j \& x_{i_{k+1}} \& \ldots \& x_{i_l}, \tag{10}$$

где h_j — булева функция, получающаяся подстановкой в функцию f_j вместо всех переменных из множества $\{x_{i_{k+1}}, \ldots, x_{i_l}\}$ константы 1 (в случае $l = k$ полагаем $\{x_{i_{k+1}}, \ldots, x_{i_l}\} = \varnothing$). Для доказательства тождества (10) достаточно рассмотреть случай, когда хотя бы одна из переменных $x_{i_{k+1}}, \ldots, x_{i_l}$ равна 0, и противоположный случай. Но функция $h_j \& x_{i_{k+1}} \& \ldots \& x_{i_l}$, как нетрудно видеть, представима в виде (9) при указанных в начале этого абзаца условиях на индексы. Индуктивный переход доказан.

Получаем, что на выходе элемента E_r, т. е. на выходе всей схемы S, реализуется функция f вида (9). Но тогда эта функция имеет один из видов (1)–(3) или равна тождественной единице, что невозможно по предположению случая 4. Противоречие.

В итоге для любой булевой функции f, отличной от тождественной единицы и непредставимой в видах (1)–(3), получаем равенство $D(f) = 3$. ■

3. Единичные проверяющие тесты при однотипных константных неисправностях типа 0 на выходах элементов

Выделим ещё два возможных представления функции $f(\tilde{x}^n)$:

$$f(\tilde{x}^n) = x_{i_1}^{\sigma_1} \& \ldots \& x_{i_k}^{\sigma_k}, \tag{11}$$

$$f(\tilde{x}^n) = \underbrace{(\ldots((}_{k-1} x_{i_1}^{\sigma_1} \& x_{i_2}^{\sigma_2})^{\delta_1} \& x_{i_3}^{\sigma_3})^{\delta_2} \& \ldots \& x_{i_k}^{\sigma_k})^{\delta_{k-1}}, \tag{12}$$

где $1 \leqslant k \leqslant n$ в представлении (11) и $2 \leqslant k \leqslant n$ в представлении (12); $i_1, \ldots, i_k$ — попарно различные индексы из множества $\{1, \ldots, n\}$ и $\sigma_1, \ldots, \sigma_k, \delta_1, \ldots, \delta_{k-1} \in \{0, 1\}$, причём хотя бы одно из чисел $\delta_1, \ldots, \delta_{k-1}$ (в представлении (12)) равно 0.

Теорема 2. Для любой булевой функции $f(\tilde{x}^n)$, отличной от констант, справедливо равенство

$$D_0(f) = \begin{cases} 0, & \text{если функция } f \text{ представима в виде (1)},\\ 1, & \text{если функция } f \text{ представима в виде (11), но не в виде (1)},\\ 2, & \text{если функция } f \text{ представима в виде (12)},\\ 3, & \text{если функция } f \text{ непредставима в видах (1), (11), (12)}. \end{cases}$$

Если $f \equiv 0$ или $f \equiv 1$, то значение $D_0(f)$ не определено.

Следствие 2. Для любого $n \geqslant 2$ справедливо равенство $D_0(n) = 3$.

Для доказательства следствия 2 достаточно заметить, что функция $x_1 \oplus \ldots \oplus x_n$ при $n \geqslant 2$ непредставима в видах (1), (11), (12).

Доказательство теоремы 2. Вместо $D_0(f)$ для краткости будем писать $D(f)$. Можно считать, что в базисе B_0 содержится функция $x_1 \& x_2$ (см. начало доказательства теоремы 1). Пусть $f \equiv 0$ или $f \equiv 1$. Рассуждая аналогично доказательству теоремы 5 из [7], получаем, что неизбыточных схем, реализующих функцию f, не существует и, следовательно, значение $D(f)$ не определено. Далее будем считать, что $f \not\equiv 0$ и $f \not\equiv 1$. Рассмотрим четыре случая:

1. Функция f представима в виде (1). Тогда её можно реализовать схемой, не содержащей функциональных элементов. У такой схемы нет ни одной функции неисправности, поэтому пустое множество для неё является ЕПТ, откуда следует равенство $D(f) = 0$.

2. Функция f представима в виде (11), но не в виде (1). Выход любой схемы, реализующей функцию f, не может совпадать ни с одним из её входов, поэтому он является выходом некоторого функционального элемента. Тогда при неисправности этого элемента схема реализует тождественный нуль, которую надо отличить от функции f хотя бы на одном наборе, откуда следует, что $D(f) \geqslant 1$.

Докажем неравенство $D(f) \leqslant 1$. Имеем $f = x_{i_1}^{\sigma_1} \& \ldots \& x_{i_k}^{\sigma_k}$, где $1 \leqslant k \leqslant n$; $i_1, \ldots, i_k$ — попарно различные индексы и $\sigma_1, \ldots, \sigma_k \in \{0, 1\}$. Реализуем функцию f схемой S в базисе B_0 в соответствии с последним равенством. В этой схеме содержатся $k-1$ двухвходовых конъюнкторов и столько инверторов, сколько чисел из $\sigma_1, \ldots, \sigma_k$ равны 0. Каждый множитель вида $\overline{x_{i_j}}$ реализуем с использованием одного инвертора. Затем все построенные инверторы и все входы x_{i_j} соединим цепочкой из конъюнкторов. Очевидно, что полученная схема реализует функцию f, а единственной её функцией неисправности является тождественный нуль. Отсюда следует, что схема S неизбыточна и множество, состоящее из любого одного набора, на котором функция f принимает значение 1, является для этой схемы ЕПТ длины 1. Поэтому $D(f) \leqslant 1$.

В итоге получаем равенство $D(f) = 1$. Случай 2 разобран.

3. Функция f представима в виде (12). Докажем сначала неравенство $D(f) \leqslant 2$. Без ограничения общности можно считать, что $i_1 = 1$, …, $i_k = k$, т. е. $f(\tilde{x}^n) = \underbrace{(\ldots((}_{k-1} x_1^{\sigma_1} \& x_2^{\sigma_2})^{\delta_1} \& x_3^{\sigma_3})^{\delta_2} \& \ldots \& x_k^{\sigma_k})^{\delta_{k-1}}$. Реализуем функцию f схемой S в базисе B_0 в соответствии с последним равенством. В этой схеме содержатся $k-1$ двухвходовых конъюнкторов и столько инверторов, сколько чисел из $\sigma_1, \ldots, \sigma_k, \delta_1, \ldots, \delta_{k-1}$ равны 0. Для каждого $i \in \{1, \ldots, k\}$, такого, что $\sigma_i = 0$, подадим входную переменную x_i на вход инвертора. Затем подадим функции $x_1^{\sigma_1}$ и $x_2^{\sigma_2}$ (каждая из которых берётся либо со входа схемы, либо с выхода одного из инверторов) соответственно на левый и правый

входы конъюнктора. Если $\delta_1 = 0$, то выход конъюнктора соединим со входом инвертора. Далее, если $k \geqslant 3$, то выход последнего построенного элемента соединим с левым входом конъюнктора, на правый вход которого подадим функцию $x_3^{\sigma_3}$ (взятую либо со входа схемы, либо с выхода одного из инверторов). Если $\delta_2 = 0$, то выход конъюнктора соединим со входом инвертора, и т. д.

Очевидно, что схема S реализует функцию f. Докажем, что она неизбыточна и множество $T = \{\tilde{\pi}_1, \tilde{\pi}_2\}$ является для неё ЕПТ, где $\tilde{\pi}_1 = (0, \sigma_2, \ldots, \sigma_k, \tilde{1}^{n-k})$, $\tilde{\pi}_2 = (1, \sigma_2, \ldots, \sigma_k, \tilde{1}^{n-k})$. Заметим, что $f(x_1, \sigma_2, \ldots, \sigma_k, \tilde{1}^{n-k}) = \underbrace{(\ldots((}_{k-1} x_1^{\sigma_1})^{\delta_1})^{\delta_2} \ldots)^{\delta_{k-1}}$, поэтому верно соотношение (4): $f(\tilde{\pi}_1) \neq f(\tilde{\pi}_2)$. Неисправность каждого инвертора в схеме S, кроме выходного (если выходной элемент — инвертор), приводит к той же функции неисправности схемы S, что и неисправность следующего за ним конъюнктора. Поэтому достаточно рассмотреть неисправности конъюнкторов и, быть может, выходного инвертора. Единственная цепь, связывающая вход x_1 схемы S с её выходом, проходит через каждый из этих элементов. Поэтому если какой-то из них неисправен, то при подаче на входы схемы S вместо переменных $x_2, \ldots, x_n$ констант соответственно $\sigma_2, \ldots, \sigma_k, \tilde{1}^{n-k}$ изменение значения переменной x_1 с 0 на 1, т. е. переход от входного набора $\tilde{\pi}_1$ к входному набору $\tilde{\pi}_2$, никак не отразится на значении, выдаваемом схемой. Следовательно, получающаяся функция неисправности g схемы S удовлетворяет условию $g(\tilde{\pi}_1) = g(\tilde{\pi}_2)$. Отсюда и из (4) вытекает, что $g \not\equiv f$, т. е. схема S неизбыточна, и множество T является для неё ЕПТ, поэтому $D(f) \leqslant 2$.

Докажем теперь, что $D(f) \geqslant 2$. Пусть S — произвольная неизбыточная схема, реализующая функцию f вида (12); T — произвольный ЕПТ для этой схемы. Надо доказать, что $|T| \geqslant 2$. Рассмотрим два подслучая:

3.1. В схеме S найдётся хотя бы один инвертор I, вход которого соединён с выходом некоторого функционального элемента E. Пусть на выходе элемента E в схеме S реализуется булева функция φ, тогда на выходе элемента I реализуется функция $\overline{\varphi}$. Чтобы обнаружить неисправность элемента E (I), в тесте T должен содержаться набор $\tilde{\sigma}_1$ ($\tilde{\sigma}_2$), для которого $\varphi(\tilde{\sigma}_1) = 1$ (соответственно $\overline{\varphi}(\tilde{\sigma}_2) = 1$). Из двух последних равенств следует, что $\tilde{\sigma}_1 \neq \tilde{\sigma}_2$, поэтому $|T| \geqslant 2$, что и требовалось доказать.

3.2. Вход любого инвертора схемы S соединён с одним из входов схемы. Рассуждая аналогично первому абзацу из доказательства леммы 2 и учитывая соотношение $f \not\equiv 0$, получаем, что функция f представима в одном из видов (1), (11). Однако это невозможно, так как она представима в виде (12). Противоречие.

В итоге для любой функции f вида (12) получаем равенство $D(f) = 2$. Случай 3 разобран.

4. Функция f отлична от булевых констант и непредставима в видах (1), (11) и (12). Докажем сначала неравенство $D(f) \leqslant 3$. Идеи доказательства сходны с идеями, использованными при получении аналогичного неравенства в случае 4 из доказательства теоремы 1. Представим функцию f полиномом Жегалкина (5), где $c \in \{0, 1\}$, а K_1 — самая короткая конъюнкция в этом полиноме. Пусть $K_i = x_{j_1(i)} \& \ldots \& x_{j_{t_i}(i)}$, $i = 1, \ldots, m$. Без ограничения общности $K_1 = x_1 \& \ldots \& x_k$. Отметим, что $k < n$, так как в противном случае функция f равна $x_1 \& \ldots \& x_n \oplus c = (x_1 \& \ldots \& x_n)^{\overline{c}}$, т. е. имеет вид (11) или (12). По аналогичной причине $m \geqslant 2$.

Так как K_1 — самая короткая конъюнкция в полиноме Жегалкина для функции f, в каждую конъюнкцию K_i, $i = 2, \ldots, m$, входит хотя бы одна переменная, отличная от переменных $x_1, \ldots, x_k$. Без ограничения общности это переменная $x_{j_1(i)}$.

Заменим в представлении (5) все операции $\oplus$ на операции $\oplus'$. Эта замена приведёт к прибавлению к правой части представления (5) некоторого числа единиц (по одному на каждую операцию $\oplus$). Сложив все эти единицы, а также константу c по модулю 2, получим некоторую булеву константу c', такую, что

$$f = K_1 \oplus' \ldots \oplus' K_m \oplus c'. \tag{13}$$

Пусть $S_{\oplus'}$ — схема в базисе B_0 с двумя входами и одним выходом (рис. 3). Нетрудно проверить, что на выходе схемы реализуется функция $x \oplus' y$, её всевозможными функциями неисправности являются функции $0, 1, xy, \overline{x}\,\overline{y}$ (и, следовательно, схема $S_{\oplus'}$ неизбыточна), а в качестве ЕПТ для неё можно взять множество $T_{\oplus'} = \{(0,0), (1,0), (1,1)\}$ (действительно, единственной булевой функцией от двух переменных, которую нельзя отличить от функции $x \oplus' y$ на наборах из множества $T_{\oplus'}$, кроме неё самой, является функция $\overline{x} \vee y$, но она не входит в число функций неисправности схемы $S_{\oplus'}$).

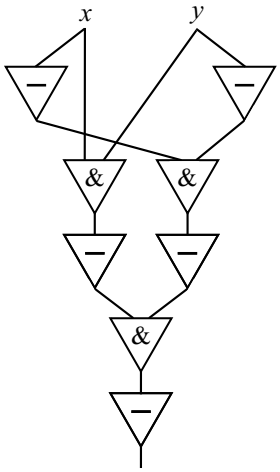

Рис. 3. Схема $S_{\oplus'}$

Реализуем функцию $f(\tilde{x}^n)$ схемой S в базисе B_0 в соответствии с представлением (13) (рис. 4). Каждую конъюнкцию K_i, $i = 1, \ldots, m$, реализуем цепочкой Z_i из конъюнкторов, причём если $i \geqslant 2$ и ранг конъюнкции K_i не менее 2, то на левый вход верхнего элемента цепочки подадим переменную $x_{j_1(i)}$, а если ранг конъюнкции K_i равен 1, то в цепочке Z_i не содержится элементов, а её выход совпадает со входом $x_{j_1(i)}$ схемы S. Затем выходы всех построенных цепочек $Z_1, \ldots, Z_m$ соединим со входами цепочки $Z_{\oplus'}$, состоящей из блоков $S_{\oplus'}$ и (в случае $c' = 1$) инвертора, вход которого соединён с выходом нижнего из этих блоков, причём левый верхний вход этой цепочки соединим с выходом цепочки Z_1. Выход нижнего элемента цепочки $Z_{\oplus'}$ объявим выходом схемы S.

Легко видеть, что схема S реализует функцию $f(\tilde{x}^n)$. Докажем, что она неизбыточна и множество $T = \{\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3\}$ является для неё ЕПТ, где $\tilde{\sigma}_1 = (\tilde{0}^n)$, $\tilde{\sigma}_2 = (\tilde{1}^k, \tilde{0}^{n-k})$, $\tilde{\sigma}_3 = (\tilde{1}^n)$. В случае исправности всех элементов схемы S на наборе $\tilde{\sigma}_3$ на выходах всех конъюнкторов, содержащихся в цепочках $Z_1, \ldots, Z_m$, будут единицы. Если неисправен некоторый конъюнктор в цепочке Z_i, $i \in \{1, \ldots, m\}$, то на наборе $\tilde{\sigma}_3$ на выходе этого конъюнктора и всех следующих за ним конъюнкторов в указанной цепочке, а значит, и на выходе цепочки Z_i, будут нули, а на выходах всех остальных конъюнкторов из этих цепочек по-прежнему будут единицы. Поскольку выход цепочки Z_i соединяется ровно с одним входом цепочки $Z_{\oplus'}$, реализующей линейную функцию, значение на выходе

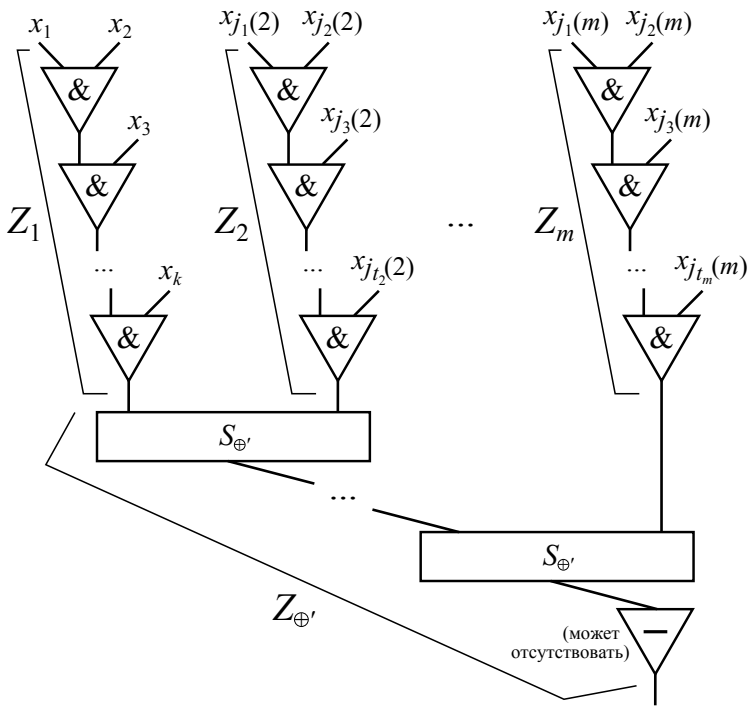

Рис. 4. Схема S

данной цепочки, т. е. на выходе всей схемы S, изменится, поэтому рассматриваемая неисправность будет обнаружена на наборе $\tilde{\sigma}_3$.

Далее, в случае исправности всех элементов схемы S на наборах $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ на выходе цепочки Z_1 возникают значения соответственно $0, 1, 1$, а на выходе каждой из цепочек $Z_2, \ldots, Z_m$ — соответственно $0, 0, 1$ (здесь используется то свойство, что каждая из переменных $x_{j_1(2)}, \ldots, x_{j_1(m)}$ отлична от переменных $x_1, \ldots, x_k$). В таком случае на входы верхнего блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ подаются наборы $(0,0), (1,0), (1,1)$, а на его выходе реализуются значения $0 \oplus' 0, 1 \oplus' 0, 1 \oplus' 1$, т. е. $1, 0, 1$. Тогда на входы второго сверху блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ (если он существует) подаются наборы $(1,0), (0,0), (1,1)$, а на его выходе реализуются значения $1 \oplus' 0, 0 \oplus' 0, 1 \oplus' 1$, т. е. $0, 1, 1$. Далее, на входы третьего сверху блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ (если он существует) подаются наборы $(0,0), (1,0), (1,1)$, а на его выходе реализуются значения $0 \oplus' 0, 1 \oplus' 0, 1 \oplus' 1$, т. е. $1, 0, 1$, и т. д. Таким образом, на входах каждого блока $S_{\oplus'}$ цепочки $Z_{\oplus'}$ при подаче на входы схемы S наборов $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ возникают все наборы из множества $T_{\oplus'}$. Так как это множество является ЕПТ для неизбыточной схемы $S_{\oplus'}$, то при неисправности любого элемента в любом блоке $S_{\oplus'}$ цепочки $Z_{\oplus'}$ хотя бы на одном из наборов $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ значение на выходе этого блока изменится. Поскольку выход блока либо совпадает с выходом схемы S, либо соединяется ровно с одним входом некоторой нижней части цепочки $Z_{\oplus'}$, реализующей линейную функцию, значение на выходе схемы S изменится, поэтому рассматриваемая неисправность будет обнаружена на одном из наборов $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$.

Наконец, если неисправен выходной инвертор цепочки $Z_{\oplus'}$ (в случае $c' = 1$), то функция неисправности схемы S равна тождественному нулю, которую можно отличить от функции f на одном из наборов $\tilde{\sigma}_1, \tilde{\sigma}_2$ (поскольку $f(\tilde{\sigma}_1) = c$, $f(\tilde{\sigma}_2) = \overline{c}$ — это следует из представления (5)).

В итоге получаем, что любую функцию неисправности схемы S можно отличить от функции $f(\tilde{x}^n)$ хотя бы на одном наборе из множества T. Это означает, что схема S

неизбыточна и множество T является для неё ЕПТ. Его длина равна 3, откуда следует неравенство $D(f) \leqslant 3$.

Докажем теперь, что $D(f) \geqslant 3$. Пусть S — произвольная неизбыточная схема, реализующая функцию f, отличную от констант и непредставимую в видах (1), (11) и (12); T — произвольный ЕПТ для этой схемы. Надо доказать, что $|T| \geqslant 3$.

Пусть E — произвольный конъюнктор, содержащийся в схеме S. Через $A_1(E)$ будем обозначать множество всех таких конъюнкторов схемы S, каждый из которых является верхним элементом хотя бы одной цепочки из конъюнкторов, у которой нижний элемент — E. Очевидно, что $E \in A_1(E)$. Через $A_2(E)$ обозначим множество всех таких элементов схемы S, не принадлежащих множеству $A_1(E)$, выход каждого из которых соединён хотя бы с одним входом хотя бы одного элемента из множества $A_1(E)$. Легко видеть, что все элементы в множестве $A_2(E)$ (если такие есть) — инверторы. Через $A_3(E)$ обозначим множество всех таких элементов схемы S, выход каждого из которых соединён со входом хотя бы одного инвертора из множества $A_2(E)$. Рассмотрим два подслучая:

4.1. В схеме S найдётся разделяющий конъюнктор E, для которого $|A_3(E)| \geqslant 2$. Среди всех элементов из множества $A_3(E)$ выберем произвольный «нижний» элемент E_1, ниже которого в схеме S не существует элемента из этого множества, и любой другой элемент E_2. Пусть I_1 (I_2) — произвольный инвертор из множества $A_2(E)$, вход которого соединён с выходом элемента E_1 (соответственно E_2); E_1' (E_2') — произвольный конъюнктор из множества $A_1(E)$, хотя бы один вход которого соединён с выходом элемента I_1 (соответственно I_2; элементы E_1' и E_2' могут совпадать). Пусть в случае исправности всех элементов схемы S на выходе элемента E_1 (E_2) реализуется булева функция φ_1 (соответственно φ_2). Тогда на выходе элемента I_1 (I_2) реализуется функция $\overline{\varphi_1}$ (соответственно $\overline{\varphi_2}$), на выходе элемента E_1' (E_2') — функция, меньшая либо равная $\overline{\varphi_1}$ (соответственно $\overline{\varphi_2}$), а на выходе элемента E — функция, меньшая либо равная как $\overline{\varphi_1}$, так и $\overline{\varphi_2}$, т. е. функция вида $\overline{\varphi_1}\&\overline{\varphi_2}\&\varphi_3$, где φ_3 — некоторая булева функция.

По лемме 1 схема S', получающаяся из схемы S удалением всех элементов, расположенных в ней не выше элемента E, кроме него самого, и переносом выхода схемы на выход элемента E, неизбыточна и T — ЕПТ для схемы S'. На выходе элемента E, т. е. на выходе схемы S', как показано ранее, реализуется функция $f' = \overline{\varphi_1}\&\overline{\varphi_2}\&\varphi_3$. При неисправности элемента E на выходе схемы S' возникнет функция неисправности $g_1 \equiv 0$. При неисправности элемента E_1 в силу его выбора на выходе элемента E_2 в схеме S (а значит, и в схеме S') по-прежнему будет реализована функция φ_2, а на выходе элемента I_2 — функция $\overline{\varphi_2}$. Тогда на выходе элемента E_2' и, как следствие, элемента E будет реализована функция, меньшая либо равная $\overline{\varphi_2}$. Поэтому получающаяся функция неисправности g_2 схемы S' представима в виде $\overline{\varphi_2}\&\varphi_3'$, где φ_3' — некоторая булева функция.

Так как схема S' неизбыточна, то каждая из функций g_1, g_2 отлична от функции f'. Чтобы отличить функцию f' от функции g_1, в множестве T должен содержаться хотя бы один набор $\tilde{\sigma}_1$, для которого $f'(\tilde{\sigma}_1) = 1$, т. е. $\varphi_1(\tilde{\sigma}_1) = \varphi_2(\tilde{\sigma}_1) = 0$, $\varphi_3(\tilde{\sigma}_1) = 1$. Чтобы отличить функцию f' от функции g_2, в множестве T должен содержаться хотя бы один набор $\tilde{\sigma}_2$, для которого $\varphi_2(\tilde{\sigma}_2) = 0$ (это следует из выражений для функций f', g_2) и $\varphi_1(\tilde{\sigma}_2) = 1$ (чтобы можно было обнаружить появление на выходе элемента E_1 вместо функции φ_1 константы 0). Из равенств $\varphi_1(\tilde{\sigma}_1) = 0$, $\varphi_1(\tilde{\sigma}_2) = 1$ следует, что $\tilde{\sigma}_1 \neq \tilde{\sigma}_2$, а из равенств $\varphi_2(\tilde{\sigma}_1) = \varphi_2(\tilde{\sigma}_2) = 0$ — что неисправность элемента E_2, на выходе которого в случае исправности всех элементов схемы S' реализуется функция φ_2, нельзя обна-

ружить на наборах $\tilde{\sigma}_1$, $\tilde{\sigma}_2$. Поэтому в тесте T должен содержаться ещё какой-то набор, отличный от указанных двух. Таким образом, $|T| \geqslant 3$, что и требовалось доказать.

4.2. Для любого разделяющего конъюнктора E в схеме S выполняется неравенство $|A_3(E)| \leqslant 1$. Так как функция f непредставима в видах (1), (11), то в схеме S содержится хотя бы один конъюнктор. «Нижний» конъюнктор схемы S, очевидно, является разделяющим (ниже него в схеме S может располагаться только цепочка из инверторов). Обозначим этот конъюнктор через E_1. Если $|A_3(E_1)| = 0$, то вход каждого инвертора, расположенного в схеме S выше элемента E_1, соединён с одним из входов этой схемы. Тогда, рассуждая аналогично первому абзацу из доказательства леммы 2, получаем, что на выходе элемента E_1 и, как следствие, на выходе схемы S реализуется функция одного из видов (1), (11), (12) или булева константа, что невозможно. Поэтому $|A_3(E_1)| = 1$. Пусть E'_2 — единственный элемент из множества $A_3(E_1)$; I_1 — произвольный инвертор из множества $A_2(E_1)$, вход которого соединён с выходом элемента E'_2. Входы всех остальных инверторов из множества $A_2(E_1)$ соединены либо с выходом элемента E'_2, либо со входами схемы. Входы всех «верхних» конъюнкторов из множества $A_1(E_1)$, очевидно, соединены либо с выходами инверторов из множества $A_2(E_1)$, либо со входами схемы.

Пусть в случае исправности всех элементов схемы S на выходе элемента E'_2 реализуется булева функция φ_1. Тогда на выходе элемента I_1 реализуется функция $\overline{\varphi_1}$. Заметим, что $E'_2 \notin A_1(E_1) \cup A_2(E_1)$, так как в противном случае на выходе элемента E_1 реализовывалась бы функция $\overline{\varphi_1}\&\varphi_1\& \ldots \equiv 0$, что невозможно. В таком случае легко видеть, что функция f, реализуемая на выходе всей схемы, имеет вид $(\overline{\varphi_1}\&K_1)^{\delta_1}$, где $\delta_1 \in \{0, 1\}$, а K_1 — либо тождественная единица, либо выражение вида $x_{i_1}^{\sigma_1}\& \ldots \&x_{i_k}^{\sigma_k}$, в котором $k \geqslant 1$; $i_1, \ldots, i_k$ — попарно различные индексы из множества $\{1, \ldots, n\}$ и $\sigma_1, \ldots, \sigma_k \in \{0, 1\}$).

Справедливо тождество $\overline{\varphi_1}\&K_1 = h_1\&K_1$, где h_1 — булева функция, получающаяся подстановкой в функцию $\overline{\varphi_1}$ вместо тех переменных, которые входят в конъюнкцию K_1, булевых констант, обращающих эту конъюнкцию в единицу (доказательство аналогично доказательству тождества (10); в случае $K_1 \equiv 1$ полагаем, что таких переменных нет). Тогда $f = (h_1\&K_1)^{\delta_1}$. Функция φ_1 не может иметь вид (1), (11), (12) или равняться константе, так как в противном случае такой же вид имела бы функция $h_1\&K_1$, а значит, и функция $f = (h_1\&K_1)^{\delta_1}$, что невозможно.

Легко видеть, что элемент E'_2 в схеме S является разделяющим. Действительно, любая цепочка, соединяющая любой элемент, расположенный в схеме S выше элемента E'_2, с выходным элементом, обязана проходить через E'_2, так как у всех остальных элементов из множества $A_1(E_1) \cup A_2(E_1) \cup A_3(E_1)$ все входы уже «заняты» либо выходами элементов из этого же множества, либо входами схемы (рис. 5).

Далее, если элемент E'_2 — конъюнктор, то положим $E_2 = E'_2$; если E'_2 — инвертор, то пусть E_2 — «нижний» конъюнктор, расположенный в схеме S выше элемента E'_2 (если такого конъюнктора нет, то выше элемента E'_2 располагается цепочка из инверторов, поэтому функция φ_1, реализуемая на выходе элемента E'_2, имеет вид x_i^{σ}, где $i \in \{1, \ldots, n\}$, $\sigma \in \{0, 1\}$, т. е. вид (1) или (11), что невозможно); во втором случае очевидно, что от элемента E_2 к элементу E'_2 ведёт цепочка из инверторов. В любом случае получаем, что на выходе конъюнктора E_2 реализуется булева функция $\varphi_1^{\delta_2}$, где $\delta_2 \in \{0, 1\}$.

Из того, что элемент E'_2 разделяющий, а между элементами E_2 и E'_2 в схеме S может располагаться только цепочка из инверторов, вход каждого из которых уже «занят» выходом предыдущего элемента в этой цепочке или выходом элемента E_2, лег-

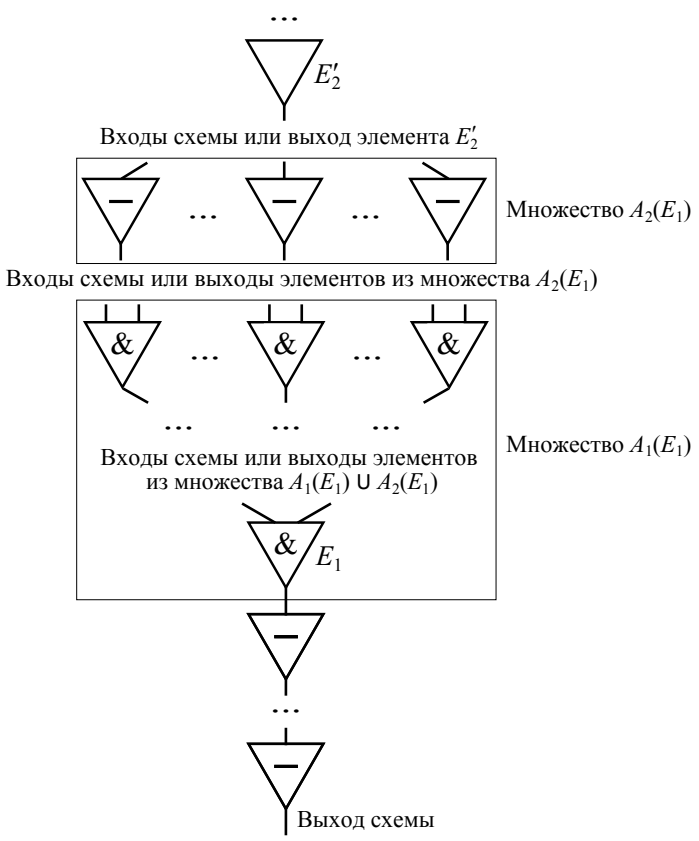

Рис. 5. Нижняя часть схемы S

ко следует, что E_2 — разделяющий конъюнктор. По предположению случая 4.2 имеем $|A_3(E_2)| \leqslant 1$. Далее отдельно разбираем случаи $|A_3(E_2)| = 0$ и $|A_3(E_2)| = 1$ по аналогии с разбором случаев $|A_3(E_1)| = 0$ и $|A_3(E_1)| = 1$ (с использованием того факта, что функция $\varphi_1^{\delta_2}$ отлична от констант и непредставима в видах (1), (11), (12)). Получаем, что в схеме S выше элемента E_2 должен существовать некоторый разделяющий конъюнктор E_3, на выходе которого реализуется функция, отличная от булевых констант и непредставимая в видах (1), (11), (12), и т. д. Поскольку число элементов в схеме S конечно, придём к противоречию. Поэтому случай 4.2 невозможен.

В итоге для любой булевой функции f, отличной от констант и непредставимой в видах (1), (11), (12), получаем равенство $D(f) = 3$. ■

Используя теоремы 1, 2, следствия 1, 2 и принцип двойственности (см., например, [19, с. 24]), а именно рассматривая схемы, получающиеся заменой всех элементов в схемах из доказательства теорем 1, 2 на двойственные, нетрудно получить двойственные им результаты для базиса B_0^*, являющегося произвольным функционально полным подмножеством множества $\hat{B}_0^* = \{\overline{x_1}, x_1 \vee \ldots \vee x_m : m \geqslant 2\}$ (например, для базиса $\{\overline{x_1}, x_1 \vee x_2\}$). В частности, при $n \geqslant 2$ справедливы равенства $D^{B_0^*;\,0}_{\mathrm{s,detect}}(n) = D^{B_0^*;\,1}_{\mathrm{s,detect}}(n) = 3$.

ЛИТЕРАТУРА

1. *Чегис И. А., Яблонский С. В.* Логические способы контроля работы электрических схем // Труды МИАН. 1958. Т. 51. С. 270–360.

2. *Яблонский С. В.* Надежность и контроль управляющих систем // Материалы Всесоюзного семинара по дискретной математике и ее приложениям (Москва, 31 января–2 февраля 1984 г.). М.: Изд-во МГУ, 1986. С. 7–12.
3. *Яблонский С. В.* Некоторые вопросы надежности и контроля управляющих систем // Математические вопросы кибернетики. Вып. 1. М.: Наука, 1988. С. 5–25.
4. *Редькин Н. П.* Надежность и диагностика схем. М.: Изд-во МГУ, 1992. 192 с.
5. *Reddy S. M.* Easily testable realization for logic functions // IEEE Trans. Comput. 1972. V. C-21. No. 11. P. 1183–1188.
6. *Коляда С. С.* Верхние оценки длины проверяющих тестов для схем из функциональных элементов: дис. ... канд. физ.-мат. наук. М., 2013. 77 с.
7. *Романов Д. С.* Метод синтеза легкотестируемых схем, допускающих единичные проверяющие тесты константной длины // Дискретная математика. 2014. Т. 26. № 2. С. 100–130.
8. *Редькин Н. П.* О полных проверяющих тестах для схем из функциональных элементов // Вестник Московского университета. Серия 1. Математика. Механика. 1986. № 1. С. 72–74.
9. *Редькин Н. П.* О полных проверяющих тестах для схем из функциональных элементов // Математические вопросы кибернетики. Вып. 2. М.: Наука, 1989. С. 198–222.
10. *Романов Д. С.* О синтезе схем, допускающих полные проверяющие тесты константной длины относительно произвольных константных неисправностей на выходах элементов // Дискретная математика. 2013. Т. 25. № 2. С. 104–120.
11. *Редькин Н. П.* О схемах, допускающих короткие тесты // Вестник Московского университета. Серия 1. Математика. Механика. 1988. № 2. С. 17–21.
12. *Редькин Н. П.* О единичных диагностических тестах для однотипных константных неисправностей на выходах функциональных элементов // Вестник Московского университета. Серия 1. Математика. Механика. 1992. № 5. С. 43–46.
13. *Бородина Ю. В.* О синтезе легкотестируемых схем в случае однотипных константных неисправностей на выходах элементов // Вестник Московского университета. Серия 15. Вычислительная математика и кибернетика. 2008. № 1. С. 40–44.
14. *Попков К. А.* О точном значении длины минимального единичного диагностического теста для одного класса схем // Препринты ИПМ им. М.В.Келдыша. 2015. № 74. 21 с.
15. *Бородина Ю. В.* Нижняя оценка длины полного проверяющего теста в базисе $\{x \mid y\}$ // Вестник Московского университета. Серия 1. Математика. Механика. 2015. № 4. С. 49–51.
16. *Бородина Ю. В.* О схемах, допускающих единичные тесты длины 1 при константных неисправностях на выходах элементов // Вестник Московского университета. Серия 1. Математика. Механика. 2008. № 5. С. 49–52.
17. *Бородина Ю. В., Бородин П. А.* Синтез легкотестируемых схем в базисе Жегалкина при константных неисправностях типа 0 на выходах элементов // Дискретная математика. 2010. Т. 22. № 3. С. 127–133.
18. *Попков К. А.* О единичных диагностических тестах для схем из функциональных элементов в базисе Жегалкина // Известия высших учебных заведений. Поволжский регион. Физико-математические науки. 2016. № 3. С. 3–18.
19. *Яблонский С. В.* Введение в дискретную математику. М.: Наука, 1986. 384 с.

REFERENCES

1. *Chegis I. A. and Yablonskiy S. V.* Logicheskie sposoby kontrolya raboty elektricheskikh skhem [Logical methods of control of work of electric circuits]. Trudy Mat. Inst. Steklov, 1958, vol. 51, pp. 270–360. (in Russian)
2. *Yablonskiy S. V.* Nadezhnost' i kontrol' upravlyayushchikh sistem [Reliability and verification of control systems]. Materialy Vsesoyuznogo seminara po diskretnoy matematike i ee

prilozheniyam (Moskva, 31 yanvarya–2 fevralya 1984 g.) Moscow., MSU Publ., 1986, pp. 7–12. (In Russian)

3. *Yablonskiy S. V.* Nekotorye voprosy nadezhnosti i kontrolya upravlyayushchikh sistem [Some questions of reliability and verification of control systems]. Matematicheskie Voprosy Kibernetiki, iss. 1, Moscow, Nauka Publ., 1988, pp. 5–25. (in Russian)
4. *Red'kin N. P.* Nadezhnost' i diagnostika skhem [Circuits Reliability and Diagnostics]. Moscow, MSU Publ., 1992. 192 p. (in Russian)
5. *Reddy S. M.* Easily testable realization for logic functions. IEEE Trans. Comput., 1972, vol. C-21, no. 11, pp. 1183–1188.
6. *Kolyada S. S.* Verkhnie otsenki dlin proveryayushchikh testov dlya skhem iz funktsional'nykh elementov [Upper bounds on the lengths of fault detection tests for logic networks]. Cand. Sci. Dissertation, MSU, Moscow, 2013. 77 p. (in Russian)
7. *Romanov D. S.* Method of synthesis of easily testable circuits admitting single fault detection tests of constant length. Discrete Math. Appl., 2014, vol. 24, no. 4, pp. 227–251.
8. *Red'kin N. P.* O polnykh proveryayushchikh testakh dlya skhem iz funktsional'nyhkh elementov [On complete fault detection tests for logic networks]. Vestnik MSU, ser. 1, 1986, no. 1, pp. 72–74. (in Russian)
9. *Red'kin N. P.* O polnykh proveryayushchikh testakh dlya skhem iz funktsional'nyhkh elementov [On complete fault detection tests for logic networks]. Matematicheskie Voprosy Kibernetiki, iss. 2, Moscow, Nauka Publ., 1989, pp. 198–222. (in Russian)
10. *Romanov D. S.* On the synthesis of circuits admitting complete fault detection test sets of constant length under arbitrary constant faults at the outputs of the gates. Discrete Math. Appl., 2013, vol. 23, no. 3–4, pp. 343–362.
11. *Red'kin N. P.* O skhemakh, dopuskayshchikh korotkiye testy [On circuits admitting short tests]. Vestnik MSU, ser. 1, 1988, no. 2, pp. 17–21. (in Russian)
12. *Red'kin N. P.* O edinichnykh diagnosticheskikh testakh dlya odnotipnykh konstantnykh neispravnostey na vykhodakh funktsional'nykh elementov [On single diagnostic tests for one-type stuck-at faults at outputs of logic gates]. Vestnik MSU, ser. 1, 1992, no. 5, pp. 43–46. (in Russian)
13. *Borodina Yu. V.* Synthesis of easily-tested circuits in the case of single-type constant malfunctions at the element outputs. Mosc. Univ. Comput. Math. Cybern., 2008, vol. 32, no. 1, pp. 42–46.
14. *Popkov K. A.* O tochnom znachenii dliny minimal'nogo edinichnogo diagnosticheskogo testa dlya odnogo klassa skhem [On an exact length value of a minimal single diagnostic test for a particular class of circuits]. Preprinty IPM im. M.V.Keldysha, 2015, no. 74. 21 p. (in Russian)
15. *Borodina Yu. V.* Lower estimate of the length of the complete test in the basis $\{x \mid y\}$. Mosc. Univ. Math. Bull., 2015, vol. 70, no. 4, pp. 185–186.
16. *Borodina Yu. V.* Circuits admitting single-fault tests of length 1 under constant faults at outputs of elements. Mosc. Univ. Math. Bull., 2008, vol. 63, no. 5, pp. 202–204.
17. *Borodina Yu. V., Borodin P. A.* Synthesis of easily testable circuits over the Zhegalkin basis in the case of constant faults of type 0 at outputs of elements. Discrete. Math. Appl., 2010, vol. 20, no. 4, pp. 441–449.
18. *Popkov K. A.* O edinichnykh diagnosticheskikh testakh dlya skhem iz funktsional'nykh elementov v bazise Zhegalkina [On single diagnostic tests for logic networks in the Zhegalkin basis]. Izvestiya Vysshikh Uchebnykh Zavedeniy. Povolzhskiy Region. Fiziko-Matematicheskie Nauki, 2016, no. 3, pp. 3–18. (In Russian)
19. *Yablonskiy S. V.* Vvedenie v diskretnuyu matematiku [Introduction to Discrete Mathematics]. Moscow, Nauka, 1986. 384 p. (in Russian)

ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ

УДК 519.17

О МИНИМАЛЬНЫХ ВЕРШИННЫХ 1-РАСШИРЕНИЯХ ОРИЕНТАЦИЙ ЦЕПЕЙ

М. Б. Абросимов*, О. В. Моденова**

*Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, г. Саратов, Россия

**Научно-образовательный центр «Эрудит», г. Саратов, Россия

Исследуются верхняя и нижняя оценки числа дополнительных дуг $ec(P_n)$ минимального вершинного 1-расширения ориентации цепи P_n. Если ориентация цепи $\overrightarrow{P}_n$ имеет концы разного типа и отлична от гамильтоновой, то $\lceil (n+1)/6 \rceil + 2 \leqslant \leqslant ec(P_n) \leqslant n+3$. Если ориентация цепи $\overrightarrow{P}_n$ имеет концы одинакового типа, то $\lceil (n+1)/4 \rceil + 2 \leqslant ec(P_n) \leqslant n+3$.

Ключевые слова: *минимальное вершинное 1-расширение, вершинная отказоустойчивая реализация, ориентация цепи.*

DOI 10.17223/20710410/38/6

ON MINIMAL VERTEX 1-EXTENSIONS OF PATH ORIENTATION

M. B. Abrosimov*, O. V. Modenova**

*Saratov State University, Saratov, Russia

**SEC "Erudit", Saratov, Russia

E-mail: oginiel@rambler.ru

In 1976, J. Hayes proposed a graph theoretic model for the study of system fault tolerance by considering faults of nodes. In 1993, the model was expanded to the case of failures of links between nodes. A graph G^* is a k-vertex extension of a graph G if every graph obtained by removing k vertex from G^* contains G. A k-vertex extension G^* of graph G is said to be minimal if it contains $n+k$ vertices, where n is the number of vertices in G, and G^* has the minimum number of edges among all k-vertex extensions of graph G with $n+k$ vertices. In the paper, the upper and lower bounds for the number of additional arcs $ec(\overrightarrow{P}_n)$ of a minimal vertex 1-extension of an oriented path $\overrightarrow{P}_n$ are obtained. For the oriented path $\overrightarrow{P}_n$ with ends of different types which is not isomorphic to Hamiltonian path, we have $\lceil (n+1)/6 \rceil + 2 \leqslant ec(\overrightarrow{P}_n) \leqslant n+3$. For the oriented path $\overrightarrow{P}_n$ with ends of equal types, we have $\lceil (n+1)/4 \rceil + 2 \leqslant ec(\overrightarrow{P}_n) \leqslant \leqslant n+3$.

Keywords: *minimal vertex extension, node fault tolerance, path orientation.*

Введение

Для изучения отказов элементов технической системы предложено понятие вершинной отказоустойчивой реализации (вершинного расширения) [1, 2], а для изучения отказов связей между элементами — понятие рёберной отказоустойчивой реализации (рёберного расширения) [3]. В [4] доказано, что задача проверки вершинного или рёберного расширения графа является NP-полной. В общем виде задачу описания вершинных расширений произвольного графа решить не удаётся. Основное направление работ в этой области продолжает следовать подходу [1 – 3], при котором описывается частное решение для графов определённого вида: цепей, циклов, деревьев. Основное внимание уделяется неориентированным графам.

В некоторых работах получены результаты для частных случаев ориентированных графов: функциональных графов [5] и контуров [6]. В данной работе исследуются ориентации цепей. Под цепью понимается дерево, степени вершин которого не больше 2. Две вершины степени 1 называются концами цепи. Под ориентацией цепи понимается орграф, получающийся из цепи заданием ориентации каждого ребра. n-Вершинную цепь будем обозначать P_n, а ориентацию цепи — $\overrightarrow{P}_n$. Очевидно, что цепь и её ориентация содержат $n-1$ рёбер и дуг соответственно. Среди важных частных случаев ориентаций цепи выделим гамильтонову цепь (все рёбра ориентируются в одну сторону) и цепь, состоящую из чередующихся источников и стоков (все рёбра ориентируются поочередно в разные стороны).

Через $d^+(v)$ и $d^-(v)$ будем обозначать степени (полустепени) исхода и захода вершины v соответственно. Для вершины v будем указывать пару степеней исхода и захода в порядке $(d^+(v), d^-(v))$. Степенью вершины в орграфе называется число дуг, имеющих эту вершину своим началом или концом: $d(v) = d^+(v) + d^-(v)$. Если в ориентации цепи концы имеют одинаковые полустепени исхода и захода, то будем говорить о цепи с концами одинакового типа, в противном случае — о цепи с концами разного типа. Будем использовать основные понятия теории графов, опираясь преимущественно на работу [7].

Дадим далее основные определения по работе [8].

Граф $G^* = (V^*, \alpha^*)$ называется *минимальным вершинным k-расширением* (МВ-kР, k — натуральное) n-вершинного графа $G = (V, \alpha)$, если выполняются следующие условия:

1) граф G^* является вершинным k-расширением графа G, то есть граф G вкладывается в каждый подграф графа G^*, получающийся удалением любых его k вершин;
2) граф G^* содержит $n + k$ вершин, то есть $|V^*| = |V| + k$;
3) α^* имеет минимальную мощность при выполнении условий 1 и 2.

Через $ec(G)$ обозначается количество дополнительных рёбер минимального вершинного расширения по сравнению с числом рёбер графа G.

Некоторые результаты могут быть полезны при переходе от неориентированных графов к ориентированным.

Лемма 1 [9]. Пусть $\overrightarrow{G}^*$ — минимальное вершинное k-расширение орграфа $\overrightarrow{G}$. Тогда симметризация $\overrightarrow{G}^*$ является вершинным k-расширением симметризации $\overrightarrow{G}$.

Напомним, что симметризацией орграфа $\overrightarrow{G} = (V, \alpha)$ называется граф

$$G = (V, (\alpha \cup \alpha^{-1}) \backslash \Delta),$$

то есть симметризация орграфа получается заменой дуг рёбрами и удалением петель.

Одним из наиболее простых результатов для неориентированных графов является утверждение о том, что минимальным вершинным 1-расширением цепи P_n является цикл C_{n+1}, который имеет два дополнительных ребра [1]. Однако перенос этой задачи на случай ориентированных графов, за исключением случая, когда все рёбра ориентируются в одну сторону, оказывается нетривиальной задачей (рис. 1). В данной работе исследуются оценки числа дополнительных дуг минимального вершинного 1-расширения ориентации цепи. В [10] доказан следующий результат.

Рис. 1. Гамильтонова ориентация цепи и её МВ-1Р

Теорема 1 [10]. Среди всех ориентаций произвольной цепи P_n только гамильтонова ориентация имеет МВ-1Р с двумя дополнительными дугами.

В [11] результат удалось улучшить.

Теорема 2 [11]. Не существует ориентаций цепей с числом вершин больше четырёх, таких, что МВ-1Р имеет три дополнительных дуги.

1. Основные результаты

Заметим достаточно очевидный факт, который можно использовать для получения верхней оценки числа дополнительных дуг.

Теорема 3. Неориентированный цикл C_{n+1} при $n > 1$ является вершинным 1-расширением для любой ориентации цепи P_n.

Из теоремы получается оценка

$$ec(P_n) \leqslant n + 3.$$

Лемма 2. Если у ориентации цепи $\overrightarrow{P}_n$ концы имеют одинаковый тип, то в её МВ-1Р не может быть двух смежных вершин степени 2.

Доказательство. Заметим, что ориентация цепи из условия теоремы не может быть гамильтоновой, так как у гамильтоновой ориентации цепи концы имеют разный тип.

Предположим, что утверждение теоремы неверно. Пусть в МВ-1Р ориентации цепи $\overrightarrow{P}_n$ есть две смежные вершины u_1 и u_2 степени 2. Не теряя общности, ориентируем ребро между u_1 и u_2 таким образом, чтобы получилась дуга из u_1 в u_2 (рис. 2).

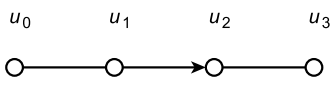

Рис. 2. Ориентация ребра между двумя смежными вершинами степени 2

Если удалим вершину u_0, то u_1 будет иметь полустепени (1,0). Если удалим вершину u_3, то u_2 будет иметь полустепени (0,1). Получили противоречие, так как по условию концы ориентации цепи должны иметь одинаковый тип. ■

Лемма 3. Если у негамильтоновой ориентации цепи $\overrightarrow{P}_n$ концы имеют разный тип, то в её МВ-1Р не может быть вершины степени 2, смежной с двумя вершинами степени 2.

Доказательство. От противного. Пусть в МВ-1Р ориентации цепи $\overrightarrow{P}_n$ есть вершина u_2, смежная с вершинами u_1 и u_3, причём $d(u_1) = d(u_2) = d(u_3) = 2$. Рассмотрим различные способы ориентации рёбер между этими вершинами.

1. Пусть дуга из u_1 идёт в u_2, из u_2 — в u_3 (рис. 3).

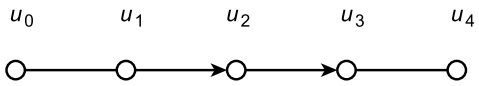

Рис. 3. Случай 1

а) Удалим вершину u_0. Тогда в цепи будет участок, состоящий из двух дуг, направленных в одну сторону (рис. 4), причём он начинается с вершины, имеющей полустепени (1,0).

б) Удалим вершину u_1. Тогда вершина u_2 имеет полустепени (1,0). По п. а получается, что дуга из u_3 должна идти в u_4 (рис. 5).

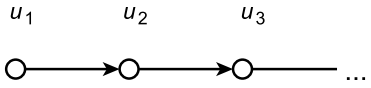

Рис. 4. Случай 1, *а*

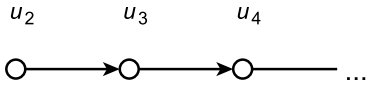

Рис. 5. Случай 1, *б*

Продолжая эти рассуждения, получаем, что все дуги ориентированы в одну сторону, то есть ориентация цепи является гамильтоновой, что противоречит условию теоремы.

2. Пусть в ориентации дуги из u_1 и u_3 идут в u_2 (рис. 6).

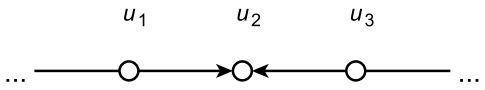

Рис. 6. Случай 2

а) Удалим вершину u_2. По условию концы цепи должны иметь разный тип. Тогда есть либо дуги (u_0, u_1) и (u_3, u_4) (рис. 7), либо дуги (u_1, u_0) и (u_4, u_3) (рис. 8).

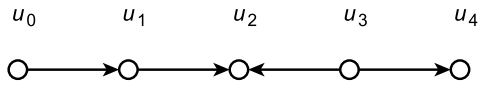

Рис. 7. Случай 2, *а*: дуги (u_0, u_1) и (u_3, u_4)

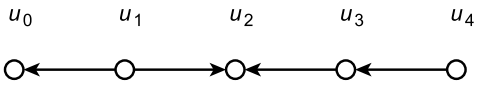

Рис. 8. Случай 2, *а*: дуги (u_1, u_0) и (u_4, u_3)

б) Пусть есть дуги (u_0, u_1) и (u_3, u_4). Если удалить u_1, то u_2 — конечная вершина ориентации цепи, т. е. после вершины со степенью (0,1) должна идти вершина степени (2,0) (рис. 9).

в) Если удалить вершину u_3, то получим противоречие с п. б — после конца ориентации цепи u_2 степени (0,1) идёт вершина степени (1,1) (рис. 10).

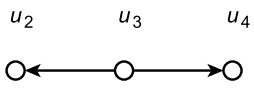

Рис. 9. Случай 2, *б*

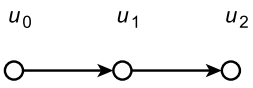

Рис. 10. Случай 2, *в*

Для второго случая (дуги (u_1, u_0) и (u_4, u_3)) доказательство аналогичное.

Для двух других способов ориентации рёбер между вершинами u_1, u_2 и u_3 доказательство проводится аналогично. ■

Теорема 4. Число дополнительных дуг MB-1P негамильтоновой ориентации цепи $\overrightarrow{P}_n$, имеющей концы разного типа, удовлетворяет неравенствам

$$\left\lceil \frac{n+1}{6} \right\rceil + 2 \leqslant ec(P_n) \leqslant n+3.$$

Доказательство. Докажем нижнюю оценку.

Учитывая лемму 3, получаем, что в MB-1P ориентации цепи $\overrightarrow{P}_n$ из условия теоремы вершина степени 2 не может быть смежна с двумя вершинами степени 2. Таким образом, чтобы построить MB-1P негамильтоновой ориентации цепи $\overrightarrow{P}_n$, требуется:

1. Две дуги, которые соединят добавленную вершину с концами цепи.

2. На каждые три вершины должно приходиться минимум по одной дуге.

Вершин в MB-1P $n+1$. Отсюда получается нижняя оценка. Верхняя оценка следует из теоремы 3. ■

Теорема 5. Число дополнительных дуг MB-1P ориентации цепи $\overrightarrow{P}_n$, имеющей концы одинакового типа, удовлетворяет неравенствам

$$\left\lceil \frac{n+1}{4} \right\rceil + 2 \leqslant ec(P_n) \leqslant n+3.$$

Доказательство. Докажем нижнюю оценку.

Учитывая лемму 2, получаем, что в MB-1P ориентации цепи из условия теоремы не может быть двух смежных вершин степени 2. Таким образом, чтобы построить MB-1P негамильтоновой ориентации цепи $\overrightarrow{P}_n$, требуется:

1. Две дуги, которые соединят добавленную вершину с концами цепи.

2. На каждые две вершины должно приходиться минимум по одной дуге.

Вершин в MB-1P $n+1$. Отсюда получается нижняя оценка. Верхняя оценка следует из теоремы 3. ■

ЛИТЕРАТУРА

1. *Hayes J. P.* A graph model for fault-tolerant computing system // IEEE Trans. Comput. 1976. V. C.25. No. 9. P. 875–884.
2. *Harary F. and Hayes J. P.* Node fault tolerance in graphs // Networks. 1996. V. 27. P. 19–23.
3. *Harary F. and Hayes J. P.* Edge fault tolerance in graphs // Networks. 1993. V. 23. P. 135–142.
4. *Абросимов М. Б.* О сложности некоторых задач, связанных с расширениями графов // Матем. заметки. 2010. № 5(88). С. 643–650.

5. *Киреева А. В.* Отказоустойчивость в функциональных графах // Упорядоченные множества и решетки. Саратов : Изд-во Сарат. ун-та, 1995. Вып. 11. С. 32–38.
6. *Sung T. Y., Lin C. Y., Chuang Y. C., and Hsu L. H.* Fault tolerant token ring embedding in double loop networks // Inform. Process. Lett. 1998. V. 66. P. 201–207.
7. *Богомолов А. М., Салий В. Н.* Алгебраические основы теории дискретных систем. М.: Наука, 1997. 368 с.
8. *Абросимов М. Б.* Графовые модели отказоустойчивости. Саратов: Изд-во Сарат. ун-та, 2012. 192 с.
9. *Абросимов М. Б.* Минимальные вершинные расширения направленных звезд // Дискретная математика. 2011. Т. 23. № 2. С. 93–102.
10. *Абросимов М. Б., Моденова О. В.* Характеризация орграфов с малым числом дополнительных дуг минимального вершинного 1-расширения // Изв. Сарат. ун-та. Нов. сер. 2013. Т. 13. Сер. Математика. Механика. Информатика. Вып. 2. Ч. 2. С. 3–9.
11. *Абросимов М. Б., Моденова О. В.* Характеризация орграфов с тремя дополнительными дугами в минимальном вершинном 1-расширении // Прикладная дискретная математика. 2013. № 3. С. 68–75.

REFERENCES

1. *Hayes J. P.* A graph model for fault-tolerant computing system. IEEE Trans. Comput., 1976, vol. C.25, no. 9, pp. 875–884.
2. *Harary F. and Hayes J. P.* Node fault tolerance in graphs. Networks, 1996, vol. 27, pp. 19–23.
3. *Harary F. and Hayes J. P.* Edge fault tolerance in graphs. Networks, 1993, vol. 23, pp. 135–142.
4. *Abrosimov M. B.* On the complexity of some problems related to graph extensions. Math. Notes, 2010, no. 5–6(88), pp. 619–625.
5. *Kireeva A. V.* Otkazoustojchivost' v funkcional'nyh grafah [Fault tolerance of functional graphs]. Uporyadochennye mnozhestva i reshetki. Saratov, SSU Publ., 1995, iss. 11, pp. 32–38. (in Russian)
6. *Sung T. Y., Lin C. Y., Chuang Y. C., and Hsu L. H.* Fault tolerant token ring embedding in double loop networks. Inform. Process. Lett., 1998, vol. 66, pp. 201–207.
7. *Bogomolov A. M. and Salii B. H.* Algebraicheskie osnovy teorii diskretnyh sistem [Algebraic Foundations of the Theory of Discrete Systems]. Moscow, Nauka Publ., 1997. 368 p. (in Russian)
8. *Abrosimov M. B.* Grafovye modeli otkazoustojchivosti [Graph Models for Fault Tolerance]. Saratov, SSU Publ., 2012. 192 p. (in Russian)
9. *Abrosimov M. B.* Minimal'nye vershinnye rasshireniya napravlennyh zvezd [Minimal vertex extensions of directed stars]. Diskr. Math., 2011, vol. 23, no. 2, pp. 93–102. (in Russian)
10. *Abrosimov M. B. and Modenova O. V.* Harakterizaciya orgrafov s malym chislom dopolnitel'nyh dug minimal'nogo vershinnogo 1-rasshireniya [Characterization of graphs with a small number of additional arcs in a minimal 1-vertex extension]. Izv. Saratov Univ. (N.S.), Ser. Math. Mech. Inform., 2013, vol. 13, iss. 2, part. 2, pp. 3–9. (in Russian)
11. *Abrosimov M. B. and Modenova O. V.* Harakterizaciya orgrafov s tremya dopolnitel'nymi dugami v minimal'nom vershinnom 1-rasshirenii [Characterization of graphs with three additional edges in a minimal 1-vertex extension]. Prikladnaya Diskretnaya Matematika, 2013, no. 3, pp. 68–75. (in Russian)

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

УДК 510.52

О ГЕНЕРИЧЕСКОЙ СЛОЖНОСТИ ПРОБЛЕМЫ ИЗВЛЕЧЕНИЯ КОРНЯ В ГРУППАХ ВЫЧЕТОВ[1]

А. Н. Рыбалов

Омский государственный университет им. Ф. М. Достоевского, г. Омск, Россия,
Институт математики им. С. Л. Соболева СО РАН, г. Новосибирск, Россия

Генерический подход к алгоритмическим проблемам предложен А. Мясниковым, И. Каповичем, П. Шуппом и В. Шпильрайном в 2003 г. В рамках этого подхода рассматривается поведение алгоритмов на множествах почти всех входов. В данной работе изучается генерическая сложность классической проблемы извлечения корня в группах вычетов $\mathbb{Z}/(m)$, где $m = pq$ — произведение двух простых чисел. Доказывается, что её естественная подпроблема генерически трудноразрешима (то есть трудна для почти всех входов) при условии, что проблема извлечения корня трудноразрешима в классическом смысле.

Ключевые слова: *генерическая сложность, проблема извлечения корня в группах вычетов, вероятностный алгоритм.*

DOI 10.17223/20710410/38/7

ON GENERIC COMPLEXITY OF THE PROBLEM OF FINDING ROOTS IN GROUPS OF RESIDUES

A. N. Rybalov

Omsk State University, Omsk, Russia,
Sobolev Institute of Mathematics, Novosibirsk, Russia

E-mail: alexander.rybalov@gmail.com

Every algorithmic problem used in modern cryptography must satisfy important conditions. First of all, the problem should be easily decidable for legal users and hard for cryptanalysis. In addition, there should be an effective algorithm for generation of hard inputs. In practically used cryptosystems with open key, such inputs are randomly generated on a sufficiently large set of inputs. A problem may be hard only for a small part of the inputs (for example, only for polynomial number of words among exponential number of all binary words). So, the problem is easy for almost all inputs. This observation leads to the concept of generic complexity and computability. In the framework of this approach, the algorithmic problem is considered on some subset of "almost all" inputs. Such inputs form the so-called generic set. The concept of "almost all" can be formalized by the introduction of a natural measure on the set of inputs. The problem can be hard (moreover, algorithmically undecidable) on

[1]Работа поддержана грантом РФФИ № 15-41-04312.

the whole set of inputs, but decidable (moreover, effectively decidable) for the "almost all" inputs. But cryptographic problems must remain hard in the generic case. In this paper, we study the generic complexity of the classical algorithmic problem of cryptography — the problem of extracting a root in the residue groups $\mathbb{Z}/(m)$, where $m = pq$ is the product of two different prime numbers. It is still unknown whether there exists a polynomial algorithm, deciding this problem for all inputs. Moreover, the famous cryptosystem RSA is based on the assumption of its hardness. We prove that this problem is generically undecidable in polynomial time, provided there is no polynomial probabilistic algorithm for its solution in the worst case. There is a plausible hypothesis (P = BPP) that any polynomial probabilistic algorithm can be efficiently derandomized, i.e. that a polynomial deterministic algorithm can be build to solve the same problem.

Keywords: *generic complexity, problem of finding roots in groups of residues, probabilistic algorithm.*

Введение

Алгоритмические проблемы, используемые в криптографии с открытым ключом, обладают рядом жёстких свойств. Прежде всего это свойства, связанные с вычислительной сложностью (проблема должна быть легкоразрешимой для легальных пользователей и трудноразрешимой для криптоанализа). Кроме того, важное значение имеет задача легкой генерации входов, на которых проблема является вычислительно сложной. В практически используемых криптосистемах с открытым ключом такие входы генерируются случайным образом на достаточно большом множестве входов проблемы. Если рассматриваемая проблема окажется труднорешаемой лишь для небольшой части входов (например, только для $O(n^3)$ слов среди всех 2^n бинарных слов), то почти на каждом сгенерированном входе проблема будет легко решаться. Это наблюдение привело к понятию генерической сложности и вычислимости [1]. В рамках этого подхода алгоритмическая проблема рассматривается не на всём множестве входов, а на некотором подмножестве «почти всех» входов. Такие входы образуют так называемое генерическое множество. Понятие «почти все» формализуется введением естественной меры на множестве входных данных. Проблема может быть труднорешаемой (более того, алгоритмически неразрешимой) на всём множестве входов, но разрешимой (более того, легкоразрешимой) на «почти всём» множестве входных данных. В [1, 2] доказано, что таким поведением обладают многие алгоритмические проблемы алгебры, а в [3] построено генерическое множество, на котором разрешима классическая проблема остановки для машин Тьюринга с лентой, бесконечной в одном направлении. Для многих классических NP-полных проблем существуют полиномиальные генерические алгоритмы [4].

С точки зрения современной криптографии интересны такие алгоритмические проблемы, которые, являясь (гипотетически) трудными в классическом смысле, остаются трудными и в генерическом смысле, т. е. для почти всех входов. Это объясняется тем, что при случайной генерации ключей в криптографическом алгоритме происходит генерация входа некоторой трудной алгоритмической проблемы, лежащей в основе алгоритма. Если проблема будет генерически легкоразрешимой, то для почти всех таких входов её можно будет быстро решить и ключи почти всегда будут нестойкими. В [5, 6] проведён генерический анализ двух классических алгоритмических проблем криптографии: проблемы распознавания квадратичных вычетов и проблемы дискретного логарифма.

В данной работе изучается генерическая сложность ещё одной классической алгоритмической проблемы криптографии — извлечения корня в группах вычетов $\mathbb{Z}/(m)$, где $m = pq$ — произведение двух различных простых чисел. До сих пор не известно полиномиальных алгоритмов её решения. Более того, на предположении об её трудноразрешимости основана знаменитая криптосистема RSA с открытым ключом [7, 8]. В данной работе доказывается, что эта проблема генерически неразрешима за полиномиальное время при условии отсутствия полиномиального вероятностного алгоритма для её решения в худшем случае. Существует правдоподобная гипотеза о том, что любой полиномиальный вероятностный алгоритм можно эффективно дерандомизировать, т. е. построить полиномиальный детерминированный алгоритм, решающий ту же задачу. Хотя это пока не доказано, имеются серьёзные результаты в пользу этой гипотезы [9]. В работе использованы методы, развитые в [5, 6, 10].

1. Генерические алгоритмы

Пусть I есть множество всех входов некоторой алгоритмической проблемы и I_n — множество всех входов размера n. Для подмножества $S \subseteq I$ определим последовательность

$$\rho_n(S) = \frac{|S \cap I_n|}{|I_n|},\ n = 1, 2, 3, \ldots$$

Заметим, что $\rho_n(S)$ — это вероятность попасть в S при случайной и равновероятной генерации входов из I_n. *Асимптотической плотностью* S назовём предел (если он существует)

$$\rho(S) = \lim_{n\to\infty} \rho_n(S).$$

Множество S называется *генерическим*, если $\rho(S) = 1$, и *пренебрежимым*, если $\rho(S) = 0$. Очевидно, что S генерическое тогда и только тогда, когда его дополнение $I \setminus S$ пренебрежимо.

Алгоритм $\mathcal{A}$ с множеством входов I и множеством выходов $J \cup \{?\}$ $(? \notin J)$ называется *генерическим*, если

1) $\mathcal{A}$ останавливается на всех входах из I;
2) множество $\{x \in I : \mathcal{A}(x) =?\}$ пренебрежимо.

Генерический алгоритм $\mathcal{A}$ вычисляет функцию $f : I \to J$, если для всех $x \in I$ $\mathcal{A}(x) = y \in J \Rightarrow f(x) = y$. Ситуация $\mathcal{A}(x) =?$ означает, что $\mathcal{A}$ не может вычислить функцию f на аргументе x, но условие 2 гарантирует, что $\mathcal{A}$ корректно вычисляет f на почти всех входах (входах из генерического множества).

2. Проблема извлечения корня в группах вычетов

Пусть $\mathbb{Z}/(m)$ — мультипликативная группа вычетов по модулю $m \in \mathbb{N}$. Напомним, что проблема извлечения корня в группах вычетов состоит в вычислении функции $\mathrm{root} : I \to \mathbb{N}$, где I — это множество троек (a, e, m), таких, что $m = pq$ (p, q — различные простые числа), $e < m$, $(\varphi(m), e) = 1$ и $a \in \mathbb{Z}/(m)$. Сама функция root определяется следующим образом:

$$\mathrm{root}(a, e, m) = x \Leftrightarrow x^e = a \in \mathbb{Z}/(m).$$

Под размером входа понимается число разрядов в двоичной записи числа m. Заметим, что при условии $(\varphi(m), e) = 1$ для любого $a \in \mathbb{Z}/(m)$ существует единственный корень степени e [11]. Этим корнем является элемент a^d, где $d = e^{-1} \mod \varphi(m)$. В настоящее время неизвестно полиномиальных алгоритмов (даже вероятностных), решающих

проблему извлечения корня в группах вычетов. На этом факте, в частности, основана криптостойкость известной криптосистемы RSA [7, 8].

Для изучения генерической сложности этой проблемы необходимо провести стратификацию на множестве входов. Рассмотрим любую бесконечную последовательность пар натуральных чисел $\mu = \{(e_1, m_1), (e_2, m_2), \ldots\}$, удовлетворяющую следующим условиям:

1) $2^n < m_n < 2^{n+1}$ для любого n;
2) m_n — произведение двух различных простых чисел для любого $n > 1$;
3) $e_n < m_n$ и $(\varphi(m_n), e_n) = 1$ для любого n.

Будем называть такую последовательность *экспоненциальной*. Из знаменитого постулата Бертрана, доказанного П. Л. Чебышевым, следует, что экспоненциальные последовательности существуют.

Теперь определим функцию root_μ как ограничение функции root на следующее множество входных данных:

$$I = \{(a, e, m) : (e, m) \in \mu, a \in \mathbb{Z}/(m)\}.$$

Под размером входа (a, e, m) понимается количество бит в двоичной записи числа m. Заметим, что множество I_n входов функции root_μ размера n состоит из всех пар (a, e, m), где (e, m) — единственная пара из μ, удовлетворяющая условию $2^n < m < 2^{n+1}$; a — любой элемент из $\mathbb{Z}/(m)$. Таким образом, проблема вычисления функции root_μ является подпроблемой проблемы вычисления функции root. Следующая лемма доказывает, что среди функций root_μ существуют функции, такие же сложные, как и оригинальная функция root.

Лемма 1. Если не существует полиномиального вероятностного алгоритма для вычисления функции root, то найдется такая экспоненциальная последовательность μ, что и для вычисления root_μ нет полиномиального вероятностного алгоритма.

Доказательство. Пусть $P_1, P_2, \ldots$ — все полиномиальные вероятностные алгоритмы. Из предположения о том, что не существует полиномиального вероятностного алгоритма для вычисления root, следует, что для любого алгоритма P_n существует бесконечно много групп $\mathbb{Z}/(m)$, в которых он не может вычислить функцию root для некоторых степеней e. Из этого следует, что можно выбрать последовательность $\mu' = \{(e_1, m_1), (e_2, m_2), \ldots\}$ так, чтобы алгоритм P_n не вычислял root в группе $\mathbb{Z}/(m_n)$ со степенью e_n и для любого n выполнялось бы $m_{n+1} > 2m_n$. Последовательность μ' можно расширить до экспоненциальной последовательности μ, добавив где нужно новые члены. Заметим теперь, что root_μ и есть та функция, для вычисления которой не существует полиномиального алгоритма. ■

3. Основной результат

Следующая теорема говорит о том, что проблема извлечения корня в группах вычетов остается вычислительно трудной и в генерическом случае при условии её труднорешимости в худшем случае.

Теорема 1. Пусть μ — любая экспоненциальная последовательность. Если существует полиномиальный генерический алгоритм, вычисляющий функцию root_μ, то существует полиномиальный вероятностный алгоритм, вычисляющий root_μ для всех входов.

Доказательство. Пусть существует полиномиальный генерический алгоритм $\mathcal{A}$, вычисляющий функцию root_μ. Построим вероятностный полиномиальный алгоритм $\mathcal{B}$,

вычисляющий root_μ на всём множестве входов. Алгоритм $\mathcal{B}$ на входе (a, e, m) работает следующим образом:

1) Сгенерировать случайно и равномерно натуральное $b < m$.
2) Если $(b, m) = 1$, то положить $a' = ab^e$, затем:
   а) запустить алгоритм $\mathcal{A}$ на (a', e, m);
   б) если $\mathcal{A}(a', e, m) = y \in \mathbb{Z}/(m)$, то $a' = y^e = ab^e$, откуда $a = (yb^{-1})^e$ и корень степени e из a равен yb^{-1};
   в) если $\mathcal{A}(a', e, m) =?$, то выдать 1.
3) Если $(b, m) \neq 1$, то $(b, m) = p$, где p — один из двух делителей числа $m = pq$. Таким образом, можно легко посчитать $\varphi(m) = (p-1)(q-1)$ и найти корень как элемент a^d, где $d = e^{-1} \mod \varphi(m)$.

Заметим, что данный полиномиальный вероятностный алгоритм может выдать неправильный ответ только на шаге 2, *в*. Докажем, что вероятность этого меньше $1/2$. Действительно, $a' = ab^e$ при $b \in \mathbb{Z}/(m)$ и $(b, m) = 1$ пробегает все элементы группы $\mathbb{Z}/(m)$, поэтому множество

$$\{(ab^e, e, m) : \ b \in \mathbb{Z}/(m), \ (b, m) = 1\}$$

совпадает с множеством всех входов размера n. Но алгоритм $\mathcal{A}$ генерический, поэтому доля тех входов (a', e, m), на которых он выдаёт неопределённый ответ, стремится к 0 с ростом n и с некоторого момента становится меньше $1/2$. ■

Непосредственным следствием доказанной теоремы является

Теорема 2. Если для вычисления функции root не существует полиномиального вероятностного алгоритма, то существует экспоненциальная последовательность μ, такая, что для вычисления функции root_μ не существует генерического полиномиального алгоритма.

Автор выражает благодарность рецензенту за полезные замечания и предложения по улучшению текста статьи.

ЛИТЕРАТУРА

1. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks // J. Algebra. 2003. V. 264. No. 2. P. 665–694.
2. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Average-case complexity for the word and membership problems in group theory // Adv. Math. 2005. V. 190. P. 343–359.
3. *Hamkins J. D. and Miasnikov A. G.* The halting problem is decidable on a set of asymptotic probability one // Notre Dame J. Formal Logic. 2006. V. 47. No. 4. P. 515–524.
4. *Gilman R., Miasnikov A. G., Myasnikov A. D., and Ushakov A.* Report on generic case complexity // Herald of Omsk University. 2007. Special Issue. P. 103–110.
5. *Рыбалов А.* О генерической сложности проблемы распознавания квадратичных вычетов // Прикладная дискретная математика. 2015. № 2(28). С. 54–58.
6. *Рыбалов А.* О генерической сложности проблемы дискретного логарифма // Прикладная дискретная математика. 2016. № 3(33). С. 93–97.
7. *Rivest R., Shamir A., and Adleman L.* A method for obtaining digital signatures and public-key cryptosystems // Commun. ACM. 1978. V. 21. Iss. 2. P. 120–126.
8. *Mao B.* Современная криптография: теория и практика. М.: Вильямс, 2005. 768 с.
9. *Impagliazzo R. and Wigderson A.* P = BPP unless E has subexponential circuits: Derandomizing the XOR Lemma // Proc. 29th STOC. El Paso: ACM, 1997. P. 220–229.

10. *Myasnikov A. and Rybalov A.* Generic complexity of undecidable problems // J. Symbolic Logic. 2008. V. 73. No. 2. P. 656–673.
11. *Коблиц Н.* Курс теории чисел и криптографии. М.: ТВП, 2001. 254 с.

REFERENCES

1. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks. J. Algebra, 2003, vol. 264, no. 2, pp. 665–694.
2. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Average-case complexity for the word and membership problems in group theory. Adv. Math., 2005, vol. 190, pp. 343–359.
3. *Hamkins J. D. and Miasnikov A. G.* The halting problem is decidable on a set of asymptotic probability one. Notre Dame J. Formal Logic, 2006, vol. 47, no. 4, pp. 515–524.
4. *Gilman R., Miasnikov A. G., Myasnikov A. D., and Ushakov A.* Report on generic case complexity. Herald of Omsk University, 2007, Special Issue, pp. 103–110.
5. *Rybalov A.* O genericheskoy slozhnosti problemy raspoznavaniya kvadratichnykh vychetov [On generic complexity of the quadratic residuosity problem]. Prikladnaya Diskretnaya Matematika, 2015, no. 2(28), pp. 54–58. (in Russian)
6. *Rybalov A.* O genericheskoy slozhnosti problemy diskretnogo logarifma [On generic complexity of the discrete logarithm problem]. Prikladnaya Diskretnaya Matematika, 2016, no. 3(33), pp. 93–97. (in Russian)
7. *Rivest R., Shamir A., and Adleman L.* A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM, 1978, vol. 21, iss. 2, pp. 120–126.
8. *Mao V.* Sovremennaya kriptografiya: teoriya i praktika [Modern Cryptography: Theory and Practice]. Moscow, Wil'yams Publ., 2005. 768 p. (in Russian)
9. *Impagliazzo R. and Wigderson A.* P = BPP unless E has subexponential circuits: Derandomizing the XOR Lemma. Proc. 29th STOC, El Paso, ACM, 1997, pp. 220–229.
10. *Myasnikov A. and Rybalov A.* Generic complexity of undecidable problems. J. Symbolic Logic, 2008, vol. 73, no. 2, pp. 656–673.
11. *Koblits N.* Kurs teorii chisel i kriptografii [Course of Number Theory and Cryptography]. Moscow, TVP Publ., 2001. 254 p. (in Russian)

УДК 004.4

КОМПЛЕКСЫ В ЛЯПАСЕ[1]

В. О. Сафонов, Д. А. Стефанцов

Национальный исследовательский Томский государственный университет, г. Томск, Россия

Описаны уточнения, которые внесены в правила выполнения операций над комплексами в языке программирования ЛЯПАС. Цель этих нововведений — обеспечить более удобную и безопасную работу с комплексами. Представлена реализация новых правил в модуле транслятора.

Ключевые слова: *язык программирования ЛЯПАС, операции над комплексами, транслятор.*

DOI 10.17223/20710410/38/8

COMPLEXES IN LYAPAS

V. O. Safonov, D. A. Stefantsov

National Research Tomsk State University, Tomsk, Russia

E-mail: vsaffonov.1115@gmail.com, d.a.stephantsov@gmail.com

The changes in semantics of operations over the data structures called complexes in the LYaPAS programming language are discussed. The modifications include resizing the complexes, passing complexes that are created by a callee to the caller, and compile-time error reporting due to modifying operations being applied to input complexes. The goal is to make the work with complexes more convenient and less error prone, which is assumed to have a positive impact on the security of the programs written using the new language constructs. The implementation of the new semantics in the module of the translator is demonstrated.

Keywords: *LYaPAS, data structures, complex, translator.*

Введение

Отечественный язык программирования ЛЯПАС [1] возрождается с целью иметь высокопроизводительную доверенную вычислительную систему для создания доверенного ПО, разработки и исследования криптографических алгоритмов, безопасного управления сетевым оборудованием, критически важными объектами и технологическими процессами. Главные компоненты создаваемой системы — транслятор с ЛЯПАСа и операционная система (ОС) ЛЯПАС.

Модульный транслятор с ЛЯПАСа [2, 3] разрабатывается как вспомогательное средство для создания транслятора, работающего под управлением ОС ЛЯПАС [4]. *Модули* реализуют последовательные фазы процесса трансляции программы на ЛЯПАСе в программу на машинном языке. Результат работы каждого из модулей представляется на выходном языке соответствующего модуля. Данная работа посвящена

[1]Работа поддержана грантом РФФИ, проект № 17-01-00354.

разработке и реализации модуля для работы с комплексами. Далее этот модуль и его язык называются *комплексояз*. Описание этого языка можно найти в [5].

Комплекс — структура данных в виде набора равных по размеру элементов, расположенных в памяти непосредственно друг за другом. Комплекс имеет потенциальный и текущий размеры, которые называются *ёмкостью* и *мощностью* соответственно. Обращение к элементу за пределами комплекса приводит к аварийному завершению программы.

Комплекс, как и переменную, можно передать в подпрограмму. В текущей версии ЛЯПАСа [1] эта передача происходит по ссылке. Изменения комплекса (как входного, так и выходного) в подпрограмме видны в вызывающей программе.

Комплексы в существующей версии ЛЯПАСа обладают некоторыми недостатками, например, их ёмкость нельзя изменять динамически; проверка выхода за границы выполняется по ёмкости, а не по мощности. Деление на входные и выходные комплексы является формальным. За неизменность входного комплекса отвечает программист, транслятор не накладывает ограничения на работу с этим комплексом. Перенос такой ответственности с программиста на транслятор позволит внести в программы на ЛЯПАСе больше дисциплины. Для решения некоторых задач необходимо иметь возможность увеличивать размеры уже созданного комплекса, например, если нужно поместить в комплекс строку, которая считывается с клавиатуры, то заранее не известно, как много памяти потребуется.

В данной работе вносятся уточнения в правила выполнения в ЛЯПАСе операций над комплексами, обеспечивающие более удобную и безопасную работу с комплексами, и описываются их реализации в модуле транслятора.

Модуль транслирует операции работы с комплексами в набор операций более низкого уровня, часть которых присутствуют в предыдущих промежуточных языках. Первая часть новых операций — это операции захвата и освобождения памяти. Операция захвата присутствует в двух модификациях, одна из которых захватывает память с некоторым запасом для оптимизации дальнейшего расширения комплекса. Вторая часть — операции работы со строками и символами: размещение строки в памяти, чтение символа с консоли и запись символа в консоль.

Модульный транслятор поддерживает вывод ошибок трансляции и ошибок во время выполнения программ. На этапе трансляции обнаруживаются логические ошибки использования комплексов, например модификация входного комплекса в подпрограмме. К типичным ошибкам во время выполнения относятся выход за мощность комплекса и нехватка памяти для захвата. Аварийное завершение программы сопровождается сообщением, которое содержит информацию об ошибке и название операции.

1. Операции над комплексами в ЛЯПАСе

1.1. Динамическое изменение ёмкости

В ЛЯПАС вводится возможность изменять ёмкость комплекса. Увеличение ёмкости влечёт захват участка памяти с бóльшим размером, копирование элементов из старого участка памяти в новый, а затем освобождение старого участка. Для сокращения количества таких операций широко используется подход увеличения ёмкости с некоторым запасом [6]. Для выбора оптимальной ёмкости будем использовать неравенство $m \leqslant k^i$, где m — ёмкость, запрашиваемая пользователем, k — некоторый коэффициент больше единицы. После нахождения минимального i, для которого неравенство истинно, используем k^i в качестве оптимальной ёмкости.

Пусть n — текущая ёмкость комплекса, тогда запись значения m в ёмкость комплекса обладает следующими свойствами:

1) ёмкость не меняет свое значение, если $m \leqslant n$;
2) ёмкости будет присвоено значение не меньше m.

Из первого свойства следует, что запись в ёмкость комплекса никогда не уменьшает её. Для сокращения размера занимаемой комплексом памяти можно воспользоваться операцией сокращения ёмкости [1].

1.2. Правила передачи комплексов в подпрограммы

Необходимо обеспечить неизменность входных комплексов на этапе трансляции программы. Назовём операции над комплексами, которые могут менять мощность, ёмкость или элементы комплекса, операциями записи. Оставшиеся операции над комплексами назовём операциями чтения. Запишем следующие правила:

1) к входному комплексу можно применять только операции чтения, операции записи запрещены;
2) к выходному и входо-выходному комплексам можно применять операции записи и чтения.

Будем выдавать ошибку трансляции, если программа содержит операции, запрещённые по отношению к данному типу комплексов.

1.3. Создание комплексов с помощью подпрограмм

Иногда возникает потребность создавать комплекс внутри подпрограммы, а в вызывающей программе использовать этот комплекс. Например, такой подпрограммой может быть построение последовательности (длины n) чисел Фибоначчи:

```
1 fibonacci(n/L1)
2 n⇒S1⇒Q1 ↑(n=0)2 0⇒L1.0⇒j ↑(n=1)2 1⇒L1.1⇒k 2⇒i
3 §1 ↑(i=n)2 L1j+L1k⇒L1i Δi Δk Δj →1
4 §2 **
```

В текущей версии ЛЯПАСа мы можем использовать такую подпрограмму только с созданным заранее комплексом:

```
1 main()
2 @+L1(10)
3 *fibonacci(10/L1)
4 **
```

В таком случае при создании комплекса можно ошибочно задать ёмкость больше или меньше, чем требуется функции `fibonacci`. Такая ошибка не критична, но может привести к лишним выделениям памяти.

Рассмотрим следующую функцию; она создаёт комплекс случайного размера и заполняет его числами 1, 2, ...:

```
1 makerandom(f,t/L1)
2 t-f⇒d X;d+f⇒S1⇒Q1 Oi
3 §1 ↑(i=Q1)2 i+1⇒L1i Δi →1
4 §2 **
```

Во время вызова подпрограммы недоступна информация о том, как много элементов будет содержать комплекс, это определяется внутри подпрограммы. Предлагается

разрешить передачу в подпрограмму ранее не созданного комплекса в качестве выходного. В этом случае будет создан комплекс с нулевой ёмкостью, а подпрограмма расширит его ёмкость до нужного размера:

```
1 main()
2 *fibonacci(10/L1)
3 *makerandom(100,4000/L2)
4 **
```

2. Программная реализация

Операции над комплексами транслируются в набор операций низкого уровня. В некоторых случаях количество операций превышает норму и ведёт к «разбуханию» кода [7]. В результате программа может не поместиться на устройствах с ограниченным количеством оперативной памяти, что приведёт к дополнительному обмену с жёстким диском, а также уменьшит коэффициент попадания команд в кэш процессора L1. Всё это может отрицательно повлиять на производительность программ. Предлагается операции над комплексами, которые транслируются в большой набор операций низкого уровня, заменять на вызов подпрограмм, которые назовём внутренними подпрограммами комплексояза. Список внутренних подпрограмм определяется экспериментальным путём и зависит от реализации транслятора, в частности рекомендуется оставить этот список пустым для устройств без дефицита оперативной памяти.

2.1. Разбиение на подзадачи

Работу комплексояза можно разделить на два этапа: проверку корректности входной программы и её трансляцию. Первый этап служит для выявления ошибок транслируемой программы и включает в себя две подзадачи; первая из них является общей для всех модулей транслятора — валидация операций и их операндов, вторая подзадача специфична для комплексояза и отвечает за соблюдение правил выполнения операций над комплексами. Второй этап состоит из подзадач трансляции операндов, а также трансляции операций над комплексами в операции более низкого уровня.

Для иллюстрации рассмотрим небольшой отрывок из входной программы комплексояза:

```
1 definition f1, a / F1
2 move L2[0], a
3 move 10, 15
4 read_complex F1
```

Первая подзадача первого этапа выявит, что третья строка содержит ошибку в первом операнде, так как операция `move` не может писать значения в константу. Вторая подзадача сообщит о попытке чтения из выходного комплекса в четвертой строке. Модуль комплексояз заканчивает работу на первом этапе, если найдена хотя бы одна ошибка, результатом является список найденных ошибок. После исправлений ошибок программа выглядит следующим образом:

```
1 definition f1, a, F1 /
2 move L2[0], a
3 read_complex F1
```

Программа после трансляции операндов:

```
1 definition f1, a, F1 /
2 move 8byte L2_buffer[0], a
3 read_complex F1
```

Программа после финального шага трансляции операций над комплексами:

```
1 definition f1, a, F1 /
2 move 8byte L2_buffer[0], a
3 move t1, 8byte F1_struct[0]
4 label 1
5 compare t1, 8byte F1_struct[1]
6 jump_⩾ 3
7 read_char 1byte F1_buffer[t1]
8 compare 1byte F1_buffer[t1], 10
9 jump_eq 2
10 inc t1
11 jump 1
12 label 2
13 inc t1
14 label 3
15 move 8byte F1_struct[0], t1
```

2.2. Вспомогательный язык для записи правил трансляции промежуточных языков

Правило трансляции записывается в следующем виде:

```
1 op <args>
2 =>
3 op_1 <args>
4 ...
5 op_n <args>
```

Здесь `op` — транслируемая операция; `op_1`, ... , `op_n` — список результирующих операций. Количество аргументов для каждой из операций не ограничено.

Операция `op` стоит из двух частей: тип операции и её значение. Если у операции есть тип, но отсутствует значение, то записывается только первая часть; например метка записывается следующим образом: `label`.

Типы аргументов

1) Аргумент {name_of_arg} сохраняет исходный тип:

```
1 cmd/store {arg}
2 =>
3 cmd/move {arg}, "acc"
```

Какой бы тип ни был у аргумента команды `store`, например строка или число, в результирующей команде `move` этот тип сохранится.

2) Аргумент <name_of_arg> позволяет выбрать результирующий тип:

```
1 cmd/some "<int>"
2 =>
3 cmd/some <int>, "<int>"
```

В этом случае происходит прямая подстановка, поэтому первый аргумент будет иметь целочисленный тип, а второй — строковый. Можно использовать более интересные подстановки:

```
1 cmd/swap_comp_el "<complex>", "<int:1>", "<int:2>"
2 =>
3 cmd/move "swaptemp", "<complex>[<int:1>]"
4 cmd/move "<complex>[<int:1>]", "<complex>[<int:2>]"
5 cmd/move "<complex>[<int:2>]", "swaptemp"
```

3) Вспомогательные аргументы, которые могут быть использованы только в результирующих операциях:

- <free_label = N> — подставить число, равное сумме номера первой свободной метки и N. Например, если метки 1, 2, 3 и 4 используются в процедуре, то <free_lable=2> будет транслирован в 7;
- <free_var = N> — подставить строку, которая получается в результате конкатенации двух строк: первая часть — t, вторая — сумма номера первой свободной переменной и числа N. Например, <free_var=3> будет транслироваться в t5 при условии, что в процедуре уже присутствуют переменные t1 и t2.

Пример:

```
1  cmd/clear_complex "<complex>"
2  =>
3  cmd/move "<free_var=1>", 0
4  label <free_label=1>
5  cmd/compare "<complex_cardinality>", "<free_var=1>"
6  cmd/jump_eq <free_label=2>
7  cmd/move "<complex_cell=<free_var=1>>", 0
8  cmd/inc "<free_var=1>"
9  cmd/jump <free_label=1>
10 label <free_label=2>
```

Параметризация аргументов

В аргументе могут присутствовать дополнительные параметры, например id аргумента. Это удобно использовать, если у аргументов одинаковые имена. Пример:

```
1 cmd/swap {variable:1}, {variable:2}
2 =>
3 cmd/move "swaptemp", {variable:1}
4 cmd/move {variable:1}, {variable:2}
5 cmd/move {variable:2}, "swaptemp"
```

2.3. Архитектура

Рассмотрим этап инициализации на диаграмме последовательности (рис. 1).

Для реализации класса `Translator` [8] используется паттерн «шаблонный метод» [9], что позволяет зафиксировать общее поведение алгоритма инициализации и оставляет возможность пользователям класса переопределить метод, который возвращает правила трансляции. Связка классов `CmdBuilder` [10] и `ArgBuilder` [11] рассматривается как паттерн «команда» [9], где `CmdBulder` — это класс, который хранит коллекцию `ArgBuilder`, которая позволяет сконструировать необходимый набор аргументов. Рабочий цикл транслятора приведён на рис. 2.

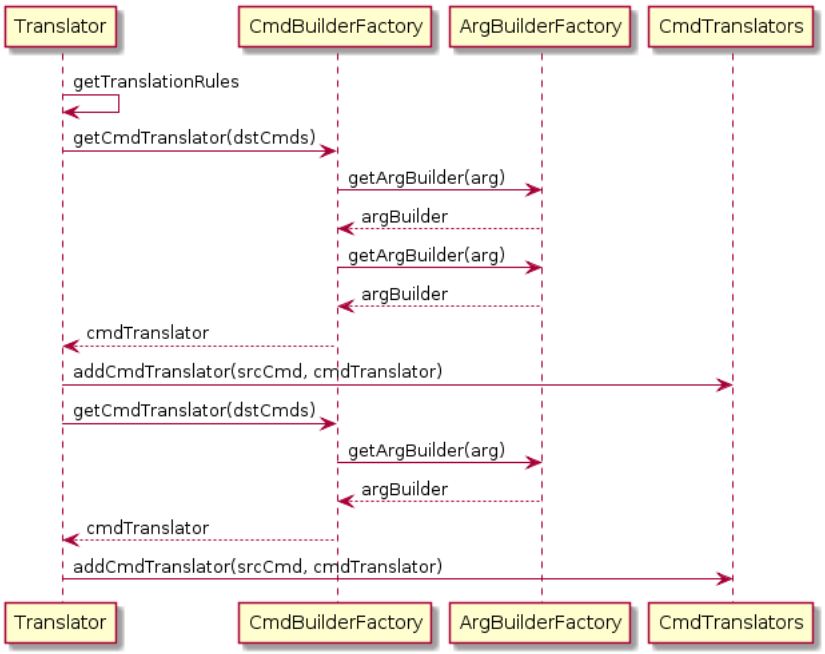

Рис. 1

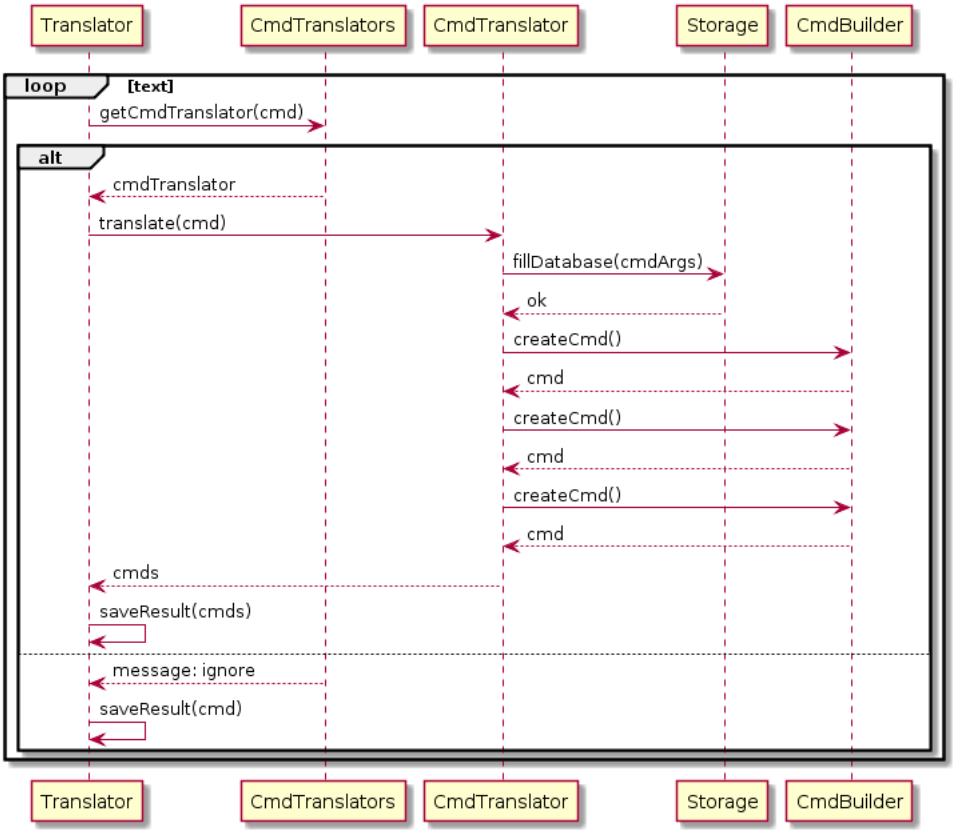

Рис. 2

Заключение

В качестве языка программирования для реализации модуля комплексояза выбран язык C++. Он удовлетворяет основным требованиям скорости разработки, а также

является достаточно популярным, что позволяет привлечь к разработке транслятора сторонних программистов. С переходом модульного транслятора на Open Source [3] последний пункт является особенно важным.

В язык ЛЯПАС внесены следующие доработки и изменения:

1) разработан механизм динамического изменения ёмкости комплексов, что позволяет расширять или уменьшать размеры комплексов в зависимости от нужд программиста;
2) разработан механизм неявного создания выходного комплекса во время вызова подпрограммы;
3) разработан механизм вывода ошибок трансляции и ошибок времени выполнения для комплексов;
4) реализована проверка на неизменность входных комплексов на этапе трансляции;
5) реализована проверка выхода за пределы комплекса по мощности.

ЛИТЕРАТУРА

1. *Агибалов Г. П., Липский В. Б., Панкратова И. А.* О криптографическом расширении и его реализации для русского языка программирования // Прикладная дискретная математика. 2013. № 3. С. 93–104.
2. *Стефанцов Д. А., Сафонов В. О., Першин В. В. и др.* Модульный транслятор с языка ЛЯПАС // Прикладная дискретная математика. Приложение. 2016. № 8. С. 122–126.
3. `https://github.com/tsu-iscd/lyapas-lcc` — LYaPAS Compiler Chain. 2017.
4. *Томских П. А., Стефанцов Д. А.* Разработка операционной системы на языке ЛЯПАС // Прикладная дискретная математика. Приложение. 2015. № 8. С. 134–135.
5. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/doc/cyaz.md` — LYaPAS Cyaz Documentation. 2017.
6. `http://en.cppreference.com/w/cpp/container/vector/reserve` — std::vector::reserve. 2017.
7. *Meyers S.* Effective C++: 55 Specific Ways to Improve Your Programs and Designs. Addison-Wesley Professional, 2005. 297 p.
8. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/sources/libs/translation_module/src/include/translation_module/translation_module.h` — LYaPAS class Translator. 2017.
9. *Gamma E., Helm R., Johnson R., et al.* Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Professional, 1994.
10. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/sources/libs/translation_module/src/include/translation_module/cmd_builder.h` — LYaPAS class CmdBuilder. 2017.
11. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/sources/libs/translation_module/src/include/translation_module/arg_builders.h` — LYaPAS class ArgBuilder. 2017.

REFERENCES

1. *Agibalov G. P., Lipskiy V. B., and Pankratova I. A.* O kriptograficheskom rasshirenii i ego realizatsii dlya russkogo yazyka programmirovaniya [Cryptographic extension and its implementation for Russian programming language]. Prikladnaya Diskretnaya Matematika, 2013, no. 3, pp. 93–104. (in Russian)

2. *Stefantsov D. A., Safonov V. O., Pershin V. V., et al.* Modul'nyy translyator s yazyka LYaPAS [Modular translator from LYaPAS]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2016, no. 8, pp. 122–126. (in Russian)
3. `https://github.com/tsu-iscd/lyapas-lcc` — LYaPAS Compiler Chain, 2017.
4. *Tomskikh P. A. and Stefantsov D. A.* Razrabotka operatsionnoy sistemy na yazyke LYaPAS [The development of an operating system in LYaPAS]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2015, no. 8, pp. 134–135. (in Russian)
5. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/doc/cyaz.md` — LYaPAS Cyaz Documentation, 2017.
6. `http://en.cppreference.com/w/cpp/container/vector/reserve` — std::vector::reserve, 2017.
7. *Meyers S.* Effective C++: 55 Specific Ways to Improve Your Programs and Designs. Addison-Wesley Professional, 2005. 297 p.
8. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/sources/libs/translation_module/src/include/translation_module/translation_module.h` — LYaPAS class Translator, 2017.
9. *Gamma E., Helm R., Johnson R., et al.* Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Professional, 1994.
10. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/sources/libs/translation_module/src/include/translation_module/cmd_builder.h` — LYaPAS class CmdBuilder, 2017.
11. `https://github.com/tsu-iscd/lyapas-lcc/blob/73b21bcd5f674bc6762a379bc32f71f61ee51164/sources/libs/translation_module/src/include/translation_module/arg_builders.h` — LYaPAS class ArgBuilder, 2017.

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ДИСКРЕТНОЙ МАТЕМАТИКЕ

УДК 519.8

ИССЛЕДОВАНИЕ L-СТРУКТУРЫ МНОГОГРАННИКА СМЕШАННОЙ ЗАДАЧИ МАКСИМАЛЬНОЙ ВЫПОЛНИМОСТИ[1]

А. В. Адельшин, А. К. Кучин

Омский филиал Института математики им. С. Л. Соболева СО РАН, г. Омск, Россия

Исследуется смешанная задача максимальной выполнимости на основе моделей целочисленного линейного программирования и метода регулярных разбиений. Установлена зависимость мощности произвольного L-комплекса многогранника указанной смешанной задачи с мощностью L-комплекса соответствующей задачи выполнимости, использование которой позволяет создавать и анализировать алгоритмы решения смешанной задачи, основанные на методе перебора L-классов.

Ключевые слова: *логические ограничения, смешанная задача максимальной выполнимости, целочисленное программирование, L-разбиение.*

DOI 10.17223/20710410/38/9

ANALYSIS OF L-STRUCTURE OF POLYHEDRON IN THE PARTIAL MAX SAT PROBLEM

A. V. Adelshin, A. K. Kuchin

Sobolev Institute of Mathematics SB RAS, Omsk, Russia

E-mail: adelshin@ofim.oscsbras.ru, dr--on@yandex.ru

In many decision-making problems, related to design, planning, management etc., the logical constraints are used. These constraints are often described in the terms of mathematical logic and lead to the satisfiability problem (SAT) and its generalizations. Most known problems are the maximum satisfiability problem (MAX SAT) and the partial maximum satisfiability problem. The latter problem includes two types of constraints that are used: the "hard" constraints (that should be satisfied anyway) and the "soft" constraints (that can be violated under certain conditions). In this paper, we analyze the partial maximum satisfiability problem as discrete optimization problem based on integer linear programming models and L-partition approach. In previous papers, estimates of the cardinality of L-complexes of polyhedrons of the SAT and the MAX SAT problems were obtained. In this paper, we prove a new property of the polyhedron of the partial MAX SAT problem, namely a relation of cardinality of the L-complexes of the indicated problem and the corresponding SAT problem is obtained. Using this result, it is possible to obtain theoretical estimates of the cardinality of the L-complex of the polyhedron of the partial MAX SAT problem

[1]Работа поддержана грантом РФФИ № 16-01-00740.

on the basis of similar estimates for the SAT and the MAX SAT problems. In particular, it is established that if hard constraints form an instance of 2-SAT problem, then the cardinality of any L-complex of the partial MAX SAT problem does not exceed $n-1$. In addition, we can construct families of logical formulas for which the cardinality of L-complex of the polyhedron of partial MAX SAT problem grows exponentially with increasing number of variables in the formulas.

Keywords: *logical constraints, partial maximum satisfiability problem, integer programming, L-partition.*

Введение

Учёт логических ограничений необходим во многих задачах дискретной оптимизации и криптографических приложениях [1 – 5]. Данные ограничения могут описываться с помощью логических формул и приводить к задаче выполнимости, одной из центральных в теории сложности, а также к её известным обобщениям — задаче максимальной выполнимости или смешанной задаче максимальной выполнимости. Практическая значимость задач с логическими ограничениями и соответствующих им задач выполнимости стимулирует разработку различных методов для их анализа и решения [6 – 11]. Одним из известных подходов является использование моделей целочисленного линейного программирования (ЦЛП) и L-разбиения [12]. В рамках этого подхода исследованы некоторые структурные свойства многогранников рассматриваемых задач и предложены семейства труднорешаемых задач для определённых классов алгоритмов [13 – 15]. В работе продолжены исследования в данном направлении. Получено новое свойство, которое отражает зависимость структуры многогранника указанной смешанной задачи со структурой многогранника соответствующей задачи выполнимости. Полученное свойство позволяет создавать и анализировать алгоритмы решения смешанной задачи, основанные на методе перебора L-классов, в частности оценивать их трудоёмкость.

Рассмотрим постановку задачи выполнимости (SAT). Пусть $x_1, \ldots, x_n$ — переменные, принимающие значение *истина* или *ложь*. Под литералом понимается либо переменная x_j, $j = 1, \ldots, n$, либо её отрицание. Пусть логическая формула F представляет собой конъюнкцию формул (скобок) D_k, $k = 1, \ldots, l$, каждая из которых является дизъюнкцией литералов. Требуется определить, выполнима ли формула F, т. е. существует ли такой набор значений переменных, при котором F принимает значение *истина*.

Важным обобщением задачи SAT является задача максимальной выполнимости (MAX SAT). Пусть логическая формула F представляет собой конъюнкцию формул (скобок) C_i, $i = 1, \ldots, m$, и каждой скобке C_i соответствует неотрицательный вес c_i. Задача MAX SAT состоит в отыскании набора значений логических переменных, при котором суммарный вес выполненных скобок будет наибольшим. Особенностью смешанной задачи является наличие так называемых «жёстких» и «мягких» логических ограничений. «Жёсткие» ограничения обязательны для выполнения и формируют некоторую задачу SAT. Выполнимость или невыполнимость «мягких» ограничений влияет на значение целевой функции и приводит к некоторой задаче MAX SAT. Такие постановки, в которых логические ограничения обоих типов присутствуют одновременно, представляют большой теоретический и практический интерес.

1. Модели ЦЛП для задач с логическими ограничениями

Перейдем к рассмотрению моделей ЦЛП рассматриваемых задач. Приведём сначала модель для задачи выполнимости. Введём булевы переменные $y_1, \ldots, y_n$, такие, что y_j соответствует переменной x_j, а $(1 - y_j)$ — её отрицанию, т. е. $y_j = 1$ тогда и только тогда, когда переменная x_j принимает значение *истина*, $j = 1, \ldots, n$. Нетрудно показать, что условие выполнимости логической формулы F эквивалентно существованию решения следующей системы:

$$\sum_{j \in D_k^+} y_j - \sum_{j \in D_k^-} y_j \geqslant 1 - |D_k^-|, \ i = 1, \ldots, l; \tag{1}$$

$$0 \leqslant y_j \leqslant 1, \ j = 1, \ldots, n, \tag{2}$$

$$y_j \in \mathbb{Z}, \ j = 1, \ldots, n.$$

Здесь D_k^- и D_k^+ — множества индексов переменных, входящих в скобку D_k с отрицанием и без него соответственно; $|D_k^-|$ — мощность множества D_k^-.

Для формулировки задачи SAT в терминах ЦЛП необходимо ввести целевую функцию. В качестве такой функции может быть выбрана, например, $f(y) = y_1 \to \max$ или $f(y) = \sum\limits_{j=1}^{n} y_j \to \max$, где $y = (y_1, \ldots, y_n)$.

Определим множества C_i^- и C_i^+ аналогично множествам D_k^- и D_k^+. Тогда модель ЦЛП для задачи MAX SAT будет выглядеть следующим образом:

$$\sum_{i=1}^{m} c_i z_i \to \max;$$

$$\sum_{j \in C_i^-} y_j - \sum_{j \in C_i^+} y_j + z_i \leqslant |C_i^-|, \ i = 1, \ldots, m; \tag{3}$$

$$0 \leqslant y_j \leqslant 1, \ j = 1, \ldots, n; \tag{4}$$

$$0 \leqslant z_i \leqslant 1, \ i = 1, \ldots, m; \tag{5}$$

$$y_j, \ z_i \in \mathbb{Z}, \ j = 1, \ldots, n, \ i = 1, \ldots, m.$$

Здесь каждой скобке C_i поставлена в соответствие переменная z_i, причём в любом допустимом целочисленном решении $z_i = 1$ только в том случае, когда C_i выполнена. Таким образом, оптимальное значение целевой функции равно наибольшему суммарному весу выполненных скобок. Заметим, что если в последнюю модель добавить ограничения вида (1), то получится модель ЦЛП для смешанной задачи с «жёсткими» ограничениями (1) и «мягкими» ограничениями (3). В результате модель ЦЛП для смешанной задачи MAX SAT будет иметь следующий вид:

$$\sum_{i=1}^{m} c_i z_i \to \max; \tag{6}$$

$$\sum_{j \in C_i^-} y_j - \sum_{j \in C_i^+} y_j + z_i \leqslant |C_i^-|, \ i = 1, \ldots, m; \tag{7}$$

$$\sum_{j \in D_k^-} y_j - \sum_{j \in D_k^+} y_j \leqslant |D_k^-| - 1, \ k = 1, \ldots, l; \tag{8}$$

$$0 \leqslant y_j \leqslant 1, \ j = 1, \ldots, n; \tag{9}$$

$$0 \leqslant z_i \leqslant 1, \ i = 1, \ldots, m; \tag{10}$$

$$y_j, z_i \in \mathbb{Z}, \ j = 1, \ldots, n, \ i = 1, \ldots, m. \tag{11}$$

Далее в п. 3 проводится анализ структуры многогранника смешанной задачи MAX SAT на основе модели ЦЛП (6)–(11).

2. О методе регулярных разбиений

Для исследования структуры задач целочисленного программирования, построения и анализа алгоритмов их решения был предложен метод регулярных разбиений [12], получивший развитие в большом количестве работ А. А. Колоколова и его учеников. Идея данного подхода заключается в выделении семейства специальных разбиений релаксационного множества задачи.

Приведём необходимые определения и обозначения. В первую очередь нам потребуется понятие лексикографического порядка. Для этого рассмотрим функцию

$$\eta(x,y) = \min\{i : x_i \neq y_i,\ i = 1,\dots,n\},\ x,y \in \mathbb{R}^n,\ x \neq y.$$

Вектор x *лексикографически больше* (*меньше*) вектора y, т. е. $x \succ y$ ($x \prec y$), если $x \neq y$ и $x_w > y_w$ ($x_w < y_w$) для $w = \eta(x,y)$. Отношение $\succ$ представляет собой отношение строгого линейного порядка. Запись $x \succeq y$ означает, что либо $x \succ y$, либо $x = y$. Аналогично понимается $x \preceq y$.

Для множеств $X, Y \subset \mathbb{R}^n$ полагаем: $X \succ Y$ ($X \prec Y$), если $x \succ y$ ($x \prec y$) для любых $x \in X$ и $y \in Y$.

Пусть $x, y \in \mathbb{R}^n$ и $x \succ y$. Будем говорить, что точки x и y *отделимы*, если найдётся точка $z \in \mathbb{Z}^n$, для которой выполняется $x \succeq z \succeq y$. Точку z назовём *отделяющей*.

Далее рассмотрим L-разбиение, которое является наиболее изученным среди разбиений. Точки $x, y \in \mathbb{R}^n$ ($x \succ y$) называются *L-эквивалентными*, если не существует отделяющей их точки $z \in \mathbb{Z}^n$. Это отношение эквивалентности порождает разбиение любого множества $X \subset \mathbb{R}^n$ на непересекающиеся *L-классы*. Фактор-множество X/L называется *L-разбиением* множества X. Отметим основные свойства L-разбиения, применяемые при разработке и исследовании алгоритмов целочисленного программирования:

1) каждая точка $z \in \mathbb{Z}^n$ образует отдельный L-класс, остальные классы состоят из нецелочисленных точек и называются *дробными*;
2) если X — ограниченное множество, то фактор-множество X/L конечно;
3) любой дробный L-класс $V \in X/L$ можно представить в виде

$$V = X \cap \{x : x_1 = a_1, \dots, x_{r-1} = a_{r-1}, a_r < x_r < a_r + 1\},$$

где $a_i \in \mathbb{Z}$ $(i = 1,\dots,r)$, $1 \leqslant r \leqslant n$.

Рангом L-класса $V \subset \mathbb{R}^n$ называется величина

$$r(V) = \begin{cases} \min\{i : x_i \neq [x_i],\ i = 1,\dots,n,\ x \in V\}, & \text{если } V \text{ — дробный,} \\ n+1 & \text{в противном случае.} \end{cases}$$

Подмножество K дробных L-классов из X/L называется *L-комплексом*, если для любых $V, V' \in K$, $V \succ V'$, не существует точки $z \in X \cap \mathbb{Z}^n$, отделяющей V и V', т. е. $V \succ z \succ V'$. Говорят, что множество X имеет *альтернирующую* L-структуру, если мощность любого её L-комплекса не превосходит 1, а лексикографически максимальный и минимальный L-классы являются целочисленными.

Далее рассмотрим лексикографическую задачу ЦЛП следующего вида: найти лексикографически максимальную точку y^* множества $(M \cap \mathbb{Z}^n)$, т. е.

$$\text{найти } y^* = \operatorname{lexmax}(M \cap \mathbb{Z}^n), \tag{12}$$

где M — некоторый выпуклый многогранник. Важную роль в исследовании задачи (12) и методов её решения играет множество

$$M_* = \{y \in M : \forall z \in M \cap \mathbb{Z}^n \ (y \succ z)\},$$

которое называется *дробным накрытием* задачи (12). Фактор-множество M_*/L называется L-*накрытием* задачи.

Выделение этой части релаксационного множества задачи связано с тем, что в некоторых методах ЦЛП (отсечения, перебора L-классов) происходит последовательное исключение точек из M_*, т. е. эти методы можно рассматривать как определённые способы «снятия» дробных накрытий.

С использованием метода регулярных разбиений получен ряд теоретических результатов при исследовании задач с логическими условиями [10, 13, 14]. Проведён анализ L-структуры многогранников (1)–(2) и (3)–(5) задач выполнимости и максимальной выполнимости. Построены семейства логических формул, для которых мощности L-накрытий задач выполнимости и максимальной выполнимости растут экспоненциально с увеличением числа переменных. Получены оценки числа итераций некоторых алгоритмов ЦЛП при решении задач из построенных семейств. Новизна результатов исследований, представленных в данной работе, состоит в применении указанного подхода к смешанной задаче максимальной выполнимости и получении свойств L-структуры соответствующего многогранника (7)–(10).

3. Анализ L-структуры многогранника смешанной задачи максимальной выполнимости

Рассмотрим модель ЦЛП смешанной задачи максимальной выполнимости (6)–(11). Пусть ограничения (8) и (9) задают многогранник D задачи выполнимости в пространстве $\mathbb{R}^n$, а релаксационный многогранник всей задачи обозначим через M. Покажем, как связаны L-структуры многогранников M и D.

Определим два отображения ψ и φ следующего вида. Пусть отображение $\psi : c \to P$, $c = (c_1, \ldots, c_n) \in D \cap \mathbb{Z}^n$, $P = \{(c_1, \ldots, c_n, p_{n+1}, \ldots, p_m) \in M : 0 \leqslant p_i \leqslant 1, i = n+1, \ldots, m\}$ задаёт грань многогранника M. Отображение $\varphi : V \to \overline{V}$ для множеств

$$V = \{(v_1, \ldots, v_n) : \ 0 < v_k < 1, \ 0 \leqslant v_j \leqslant 1, \ j = k+1, \ldots, n\} \in D/L,$$
$$\overline{V} = \{(v_1, \ldots, v_{n+m}) : \ 0 < v_k < 1, \ 0 \leqslant v_j \leqslant 1, \ j = k+1, \ldots, n+m\} \in M/L$$

связывает между собой L-классы многогранников D и M, $k \leqslant n$.

Далее будем говорить, что отображение $f : X \to Y$ сохраняет лексикографический порядок, если для любых $x_1, \ x_2 \in X$ выполняется $x_1 \prec x_2$ тогда и только тогда, когда $f(x_1) \prec f(x_2)$.

Лемма 1. Отображения ψ и φ взаимно-однозначны и сохраняют лексикографический порядок.

Доказательство данного свойства очевидно. Кроме этого, для отображения ψ верно следущее утверждение.

Лемма 2. Пусть отображение $\psi : c \to P$, $c = (c_1, \ldots, c_n) \in D \cap \mathbb{Z}^n$, $P = \{(c_1, \ldots, c_n, p_{n+1}, \ldots, p_m) \in M : 0 \leqslant p_i \leqslant 1, i = n+1, \ldots, m\}$. Тогда множество P имеет альтернирующую L-структуру.

Доказательство. Заметим, что $(c_1, \ldots, c_n, 0, \ldots, 0) \in M$, а значит, множество P не пусто. Пусть $c = (c_1, \ldots, c_n) \in D \cap \mathbb{Z}$, $P = \psi(c)$.

Очевидно, что $\text{lexmin}(P) = (c_1, \ldots, c_n, 0, \ldots, 0)$. Докажем, что $\text{lexmax}(P)$ — целочисленная точка. Допустим, $\text{lexmax}(P) = (c_1, \ldots, c_n, v_{n+1}, \ldots, v_{n+m})$ — нецелочисленная точка. Нетрудно показать, что $(c_1, \ldots, c_n, \lceil v_{n+1} \rceil, \ldots, \lceil v_{n+m} \rceil) \in P$ и лексикографически больше $(c_1, \ldots, c_n, v_{n+1}, \ldots, v_{n+m})$. Противоречие.

Рассмотрим два различных дробных L-класса

$$V = \{(c_1, \ldots, c_n, v_{n+1}, \ldots, v_{n+m}) : \ 0 \leqslant v_i \leqslant 1, \ i = n+1, \ldots, m\},$$
$$W = \{(c_1, \ldots, c_n, w_{n+1}, \ldots, w_{n+m}) : \ 0 \leqslant w_i \leqslant 1, \ i = n+1, \ldots, m\},$$

принадлежащих P/L. Будем считать, что $V \prec W$. Докажем, что существует целочисленная точка из P, отделяющая эти два L-класса:

1) Если $n < r(V) < r(W)$, то точка

$$(c_1, \ldots, c_n, w_1, \ldots, w_{r(W)-1}, \lfloor w_{r(W)} \rfloor, \ldots, \lfloor w_{n+m} \rfloor)$$

является отделяющей.

2) Если $n < r(W) \leqslant r(V)$, то точка

$$(c_1, \ldots, c_n, v_1, \ldots, v_{r(V)-1}, \lceil v_{r(V)} \rceil, \ldots, \lceil v_{n+m} \rceil)$$

является отделяющей.

Таким образом, в силу произвольности V и W, множество P имеет альтернирующую L-структуру. ■

Лемма 3. Для любого целочисленного L-класса $V \in D/L$ и любого дробного L-класса $W \in D/L$, таких, что $V \prec W$ ($V \succ W$), справедливо $\psi(V) \prec \varphi(W)$ ($\psi(V) \succ \succ \varphi(W)$).

Доказательство. Пусть $V = \{(v_1, \ldots, v_n)\} \in D/L$, где $v_j \in \{0, 1\}$, $j = 1, \ldots, n$; $W = \{(w_1, \ldots, w_n) : 0 \leqslant w_j \leqslant 1, j = r+1, \ldots, n, 0 < w_r < 1\} \in D/L$, где $w_j \in \{0, 1\}$, $j = 1, \ldots, r-1$, $r \leqslant n$, и известно, что $V \prec W$ ($V \succ W$). Тогда

$$\psi(V) = \{(v_1, \ldots, v_n, \overline{v}_{n+1}, \ldots, \overline{v}_{n+m}) \in M : 0 \leqslant \overline{v}_j \leqslant 1, \ j = n+1, \ldots, n+m\},$$
$$\varphi(W) = \{(w_1, \ldots, w_{r-1}, \overline{w}_r, \ldots, \overline{w}_{n+m}) \in M : 0 \leqslant \overline{w}_j \leqslant 1, j = r+1, \ldots, n+m, 0 < \overline{w}_r < 1\}.$$

Видно, что $\psi(V) \prec \varphi(W)$ (соответственно $\psi(V) \succ \varphi(W)$). ■

Леммы 1–3 использованы при анализе L-структуры многогранника смешанной задачи максимальной выполнимости, результаты представлены в теоремах 1 и 2.

Теорема 1. Пусть мощность любого L-комплекса многогранника (8)–(9) не превосходит величины $t(n)$. Тогда мощность любого L-комплекса многогранника (7)–(10) также не превосходит $t(n)$.

Доказательство. Используем те же обозначения: D — многогранник (8)–(9), M — многогранник (7)–(10).

Рассмотрим L-комплекс $\Omega = \{V_1, \ldots, V_k\}$ многогранника D, где $V_1, \ldots, V_k \notin \mathbb{Z}^n$, $V_1 \prec \ldots \prec V_k$, $k \leqslant t(n)$. Из леммы 1 получаем $\varphi(V_1) \prec \ldots \prec \varphi(V_k)$. Предположим, что существует $\overline{V} \in M/L$, такое, что $\varphi(V_i) \prec \overline{V} \prec \varphi(V_{i+1})$, $1 \leqslant i \leqslant k-1$. Если $r(\overline{V}) \leqslant n$, то по лемме 1 $V_i \prec \varphi^{-1}(\overline{V}) \prec V_{i+1}$; если $r(\overline{V}) > n$, то $V_i \prec \psi^{-1}(\overline{V}) \prec V_{i+1}$. Противоречие.

Пусть существует $V_0 \in D \cap \mathbb{Z}^n$, такое, что $V_0 \prec V_1$ ($V_k \prec V_0$) и между ними нет других L-классов. Из леммы 3 известно, что $\psi(V_0) \prec \varphi(V_1)$ ($\varphi(V_k) \prec \psi(V_0)$). Предположим,

что существует $\overline{V} \in M/L$, для которого $\psi(V_0) \prec \overline{V} \prec \varphi(V_1)$ ($\varphi(V_k) \prec \overline{V} \prec \psi(V_0)$). Если $r(\overline{V}) \leqslant n$, то $V_0 \prec \varphi^{-1}(\overline{V}) \prec V_1$ ($V_k \prec \varphi^{-1}(\overline{V}) \prec V_0$); если $r(\overline{V}) > n$, то $V_0 \prec \psi^{-1}(\overline{V}) \prec V_1$ ($V_k \prec \psi^{-1}(\overline{V}) \prec V_0$). Противоречие. Для случая, когда L-класса V_0 не существует, справедливы аналогичные рассуждения.

Таким образом, L-комплексу $\Omega = \{V_1, \ldots, V_k\}$ многогранника D соответствует L-комплекс $\overline{\Omega} = \{\varphi(V_1), \ldots, \varphi(V_k)\}$ многогранника M той же мощности. По лемме 2 множество $\varphi(V_0)$, где $V_0 \in D \cap \mathbb{Z}^n$, имеет альтернирующую L-структуру, а значит, мощность любого L-комплекса из $\varphi(V_0)$ не превосходит 1. Следовательно, размер любого L-комплекса множества M не превосходит $t(n)$. ■

Используя аналогичные рассуждения, можно доказать также следующую теорему.

Теорема 2. Пусть многогранник (8)–(9) содержит L-комплекс мощности не меньше $t(n)$. Тогда многогранник (7)–(10) также содержит L-комплекс мощности не меньше $t(n)$.

Таким образом, структура многогранника D является определяющей для L-структуры смешанной задачи максимальной выполнимости. В [13, 14] построены семейства задач выполнимости, у которых мощность L-накрытия растёт экспоненциально с увеличением числа переменных в формуле. Это дало возможность генерировать семейства труднорешаемых задач для определённого класса алгоритмов, основанных на непрерывной оптимизации (методы ветвей и границ, отсечения, перебора L-классов). На основе указанных результатов и теоремы 2 нетрудно построить такие смешанные задачи максимальной выполнимости, мощности L-накрытий которых также будут экспоненциальными от числа переменных. С другой стороны, в [15] показано, что мощность любого L-комплекса многогранника задачи 2-выполнимости с выполнимой формулой не превосходит $n-1$. В некоторых прикладных задачах (например, задачах проектирования [10]) набор логических ограничений представляет собой задачу 2-выполнимости, а значит, теорема 1 гарантирует, что многогранник указанной задачи содержит L-комплексы, мощности которых ограничены полиномом от числа переменных в формуле. Таким образом, переход от одного допустимого решения задачи к следующему в лексикографическом порядке осуществляется достаточно быстро. В связи с этим для смешанной задачи MAX SAT возможно построение эффективных алгоритмов, основанных на методе перебора L-классов, позволяющих применять указанный подход для некоторых постановок прикладных задач с логическими ограничениями.

ЛИТЕРАТУРА

1. *Колоколов А. А., Ярош А. В.* Автоматизация проектирования сложных изделий с использованием дискретной оптимизации и информационных технологий // Омский научный вестник. 2010. № 2(90). С. 234–238.
2. *Посыпкин М. А., Заикин О. С., Беспалов Д. В., Семенов А. А.* Решение задач криптоанализа поточных шифров в распределенных вычислительных средах // Труды ИСА РАН. 2009. Т. 46. С. 119–137.
3. *Massacci F. and Marraro L.* Logical cryptanalysis as a SAT problem // J. Automated Reasoning. 2000. No. 24. P. 165–203.
4. *Marathe M. V. and Ravi S. S.* On approximation algorithms for the minimum satisfiability problem // Inform. Proc. Lett. 1996. V. 58. P. 23–29.
5. *Заикин О. С., Отпущенников И. В., Семенов А. А.* Применение SAT-подхода к решению квадратичной задачи о назначениях // XV Байкальская Междунар. школа-семинар «Методы оптимизации и их приложения». Иркутск, 2011. С. 111–116.

6. *Колоколов А. А., Адельшин А. В., Ягофарова Д. И.* Решение задачи выполнимости с использованием метода перебора L-классов // Информационные технологии. 2009. № 2. С. 54–59.
7. *Gu J., Purdom P., Franco J., and Wah B.* Algorithms for the satisfiability (SAT) problem: a survey // DIMACS Series in Discr. Math. and Theor. Comput. Sci. 1996. V. 35. P. 19–152.
8. *Hamadi Y., Jabbour S., and Sais L.* ManySAT: a parallel SAT solver // J. Satisfiability, Boolean Modeling and Computation. 2009. No. 6. P. 245–262.
9. *Thornton J., Bain S., Sattar A., and Pham D. N.* A two level local search for MAX-SAT problems with hard and soft constraints // Proc. 15th Australian Joint Conf. on Artificial Intelligence, AustAI-2002. Canberra, Australia, 2002. P. 603–614.
10. *Колоколов А. А., Адельшин А. В., Ягофарова Д. И.* Исследование задач дискретной оптимизации с логическими ограничениями на основе метода регулярных разбиений // Прикладная дискретная математика. 2013. № 1(19). С. 99–109.
11. *Адельшин А. В., Кучин А. К.* Разработка алгоритмов решения смешанной задачи максимальной выполнимости // V Всерос. конф. «Проблемы оптимизации и экономические приложения». Омск, 2012. С. 99.
12. *Колоколов А. А.* Регулярные разбиения и отсечения в целочисленном программировании // Сиб. журнал исследования операций. 1994. № 2. С. 18–39.
13. *Адельшин А. В.* Исследование задач максимальной и минимальной выполнимости с использованием L-разбиения // Автоматика и телемеханика. 2004. № 1. С. 35–42.
14. *Kolokolov A., Adelshin A., and Yagofarova D.* Analysis and solving SAT and MAX-SAT problems using an L-partition approach // J. Math. Modeling. Algorithms. 2013. No. 12(2). P. 201–212.
15. *Колоколов А. А., Адельшин А. В.* Анализ и решение задач дискретной оптимизации с логическими ограничениями на основе L-разбиения // Прикладная дискретная математика. 2015. № 4(30). С. 100–108.

REFERENCES

1. *Kolokolov A. A. and Yarosh A. V.* Avtomatizatsiya proyektirovaniya slozhnykh izdeliy s ispol'zovaniyem diskretnoy optimizatsii i informatsionnykh tekhnologiy [Automation of designing complex products using discrete optimization and information technology]. Omsk Scientific Bulletin, 2010, no. 2(90), pp. 234–238. (in Russian)
2. *Posypkin M. A., Zaikin O. S., Bespalov D. V., and Semenov A. A.* Resheniye zadach kriptoanaliza potochnykh shifrov v raspredelennykh vychislitel'nykh sredakh [Solving the problems of cryptanalysis of stream ciphers in distributed computing environments]. Proc. ISA RAS, 2009, vol. 46, pp. 119–137. (in Russian)
3. *Massacci F. and Marraro L.* Logical cryptanalysis as a SAT problem. J. Automated Reasoning, 2000, no. 24, pp. 165–203.
4. *Marathe M. V. and Ravi S. S.* On approximation algorithms for the minimum satisfiability problem. Inform. Proc. Lett., 1996, vol. 58, pp. 23–29.
5. *Zaikin O. S., Otpushchennikov I. V., and Semenov A. A.* Primeneniye SAT-podkhoda k resheniyu kvadratichnoy zadachi o naznacheniyakh [Application of the SAT approach to the solution of the quadratic assignment problem]. 15th Baikal Intern. School-Seminar on Optimization Methods and their Applications, Irkutsk, 2011, pp. 111–116. (in Russian)
6. *Kolokolov A. A., Adelshin A. V., and Yagofarova D. I.* Resheniye zadachi vypolnimosti s ispol'zovaniyem metoda perebora L-klassov [Solving the SAT problem using the L-partition approach]. Information Technologies, 2009, no. 2. pp. 54–59. (in Russian)
7. *Gu J., Purdom P., Franco J., and Wah B.* Algorithms for the satisfiability (SAT) problem: a survey. DIMACS Series in Discr. Math. and Theor. Comput. Sci., 1996, vol. 35, pp. 19–152.

8. *Hamadi Y., Jabbour S., and Sais L.* ManySAT: a parallel SAT solver. J. Satisfiability, Boolean Modeling and Computation, 2009, no. 6, pp. 245–262.
9. *Thornton J., Bain S., Sattar A., and Pham D. N.* A two level local search for MAX-SAT problems with hard and soft constraints. Proc. 15th Australian Joint Conf. on Artificial Intelligence, AustAI-2002. Canberra, Australia, 2002, pp. 603–614.
10. *Kolokolov A. A., Adelshin A. V., and Yagofarova D. I.* Issledovaniye zadach diskretnoy optimizatsii s logicheskimi ogranicheniyami na osnove metoda regulyarnykh razbiyeniy [Study of dicrete optimization problems with logical constraints based on regular partitions]. Prikladnaya Diskretnaya Matematika, 2013, no. 1(19), pp. 99–109. (in Russian)
11. *Adelshin A. V. and Kuchin A. K.* Razrabotka algoritmov resheniya smeshannoy zadachi maksimal'noy vypolnimosti [Development of algorithms for solving the partial MAX SAT problem]. V Russian Conf. "Optimization Problems and their Economical Applications", Omsk, 2012, pp. 99. (in Russian)
12. *Kolokolov A. A.* Regulyarnye razbieniya i otsecheniya v tselochislennom programmirovanii [Regular partitions and cuts in integer programming]. Siberian J. Operations Research, 1994, no. 2, pp. 18–39. (in Russian)
13. *Adel'shin A. V.* Investigation of maximum and minimum satisfiability problems using L-partition. Automation and Remote Control, 2004, vol. 65, no. 3, pp. 388–395.
14. *Kolokolov A., Adelshin A., and Yagofarova D.* Analysis and solving SAT and MAX-SAT problems using an L-partition approach. J. Math. Modeling. Algorithms, 2013, no. 12(2), pp. 201–212.
15. *Kolokolov A. A. and Adelshin A. V.* Analiz i reshenie zadach diskretnoy optimizatsii s logicheskimi ogranicheniyami na osnove L-razbieniya [Analysis and solving discrete optimization problems with logical constraints on the basis of the L-partition approach]. Prikladnaya Diskretnaya Matematika, 2015, no. 4(30), pp. 100–108. (in Russian)

УДК 519.714.1

УТОЧНЕНИЕ НИЖНЕЙ ОЦЕНКИ СЛОЖНОСТИ ВОЗВЕДЕНИЯ В СТЕПЕНЬ

В. В. Кочергин, Д. В. Кочергин

Московский государственный университет имени М. В. Ломоносова, г. Москва, Россия

Для величины $l(x^n)$ — минимального числа операций умножения, достаточного для вычисления по переменной x степени x^n — уточнена нижняя оценка. Установлено, что для любого $\varepsilon > 0$ доля чисел k, не превосходящих n и удовлетворяющих условию

$$l(x^k) > \log_2 n + \frac{\log_2 n}{\log_2 \log_2 n}\left(1 - (2+\varepsilon)\frac{\log_2 \log_2 \log_2 n}{\log_2 \log_2 n}\right),$$

стремится к 1 при $n \to \infty$.

Ключевые слова: *аддитивные цепочки, возведение в степень, нижние оценки сложности.*

DOI 10.17223/20710410/38/10

IMPROVEMENT OF THE LOWER BOUND FOR THE COMPLEXITY OF EXPONENTIATION

V. V. Kochergin, D. V. Kochergin

Lomonosov Moscow State University, Moscow, Russia

E-mail: vvkoch@yandex.ru, dk@kocherg.in

Let $l(x^n)$ be the minimal number of multiplications sufficient for computing x^n. In the paper, we improve the lower bound of $l(x^n)$. We establish that for all $\varepsilon > 0$ the fraction of the numbers k, $k \leqslant n$, satisfying the relation

$$l(x^k) > \log_2 n + \frac{\log_2 n}{\log_2 \log_2 n}\left(1 - (2+\varepsilon)\frac{\log_2 \log_2 \log_2 n}{\log_2 \log_2 n}\right),$$

tends to 1 as $n \to \infty$.

Keywords: *addition chains, exponentiation, lower bounds of complexity.*

1. Определения, обсуждения, формулировка основного результата

В работе исследуется возможность уточнения нижней оценки для известной (см., например, [1, разд. 4.6.3]) задачи о сложности возведения в степень, т. е. задачи о нахождении величины $l(x^n)$ — минимального числа операций умножения, достаточного для вычисления по переменной x степени x^n.

Эту задачу (а также её обобщения) обычно рассматривают в аддитивной постановке — это так называемая задача об аддитивных цепочках [2], которая формулируется следующим образом. *Аддитивной цепочкой* для натурального числа n называется всякая последовательность целых чисел

$$a_0 = 1, a_1, \ldots, a_m = n,$$

удовлетворяющая следующему свойству: для каждого k, $1 \leqslant k \leqslant m$, найдутся два целых числа (не обязательно различных) i и j, $0 \leqslant i \leqslant k-1$, $0 \leqslant j \leqslant k-1$, таких, что $a_k = a_i + a_j$. Минимальная длина m аддитивной цепочки для n называется *аддитивной сложностью* числа n и обозначается $l(n)$. Очевидно, что величины $l(n)$ и $l(x^n)$ равны.

Различным аспектам классической задачи об эффективном вычислении степеней (задачи о длине аддитивных цепочек) посвящено большое число публикаций (см., например, работы [1, 3–11], являющиеся обзорами или содержащие обзорную часть). Кроме того, в связи с активным применением аппарата аддитивных цепочек в криптографических алгоритмах и других приложениях [12–14] в последнее время объём литературы по этой тематике серьёзно увеличился. В значительной части публикаций приводятся разные эвристические алгоритмы возведения в степень (построения аддитивных цепочек), но принципиальных улучшений оценок величины $l(x^n)$, доказанных в середине прошлого века, практически не получено.

В 1939 г. А. Брауэром [15] для величины $l(n)$ при $n \to \infty$ установлена асимптотическая формула

$$l(n) \sim \log n$$

и получена верхняя оценка

$$l(n) \leqslant \log n + \frac{\log n}{\log\log n}\left(1 + \frac{2\log\log\log n}{\log\log n} + \frac{1+\mathrm{o}(1)}{\log\log n}\right). \tag{1}$$

Здесь и далее $\log x$ — это $\log_2 x$, а запись $f(n) \sim g(n)$ означает, что при $n \to \infty$ отношение $f(n)/g(n)$ стремится к 1.

В 1960 г. П. Эрдёш [16] показал, что для любого $\varepsilon > 0$ найдётся такая константа c, $1 < c < 2$, что доля чисел n, не превосходящих N и удовлетворяющих условию

$$l(n) > \log N + (1-\varepsilon)\frac{\log N}{\log\log N},$$

а, следовательно, и условию

$$l(n) > \log n + (1-\varepsilon)\frac{\log n}{\log\log n}, \tag{2}$$

при всех достаточно больших значениях n не меньше величины $1 - (c/2)^{\log n}$ и, как следствие, стремится к 1 при $n \to \infty$ (похожий, но более слабый результат установлен в [17]).

Стоит отметить разную природу слагаемых в правой части неравенства (2) — слагаемое $\log n$ связано с величиной числа n и должно присутствовать для любого значения n, а «мощностное» (отношение логарифма количества чисел, не превосходящих n, к повторному логарифму) слагаемое зависит от «строения» числа n и присутствует для «почти всех» n. Однако несмотря на то, что для почти всех значений n величина $l(n) - \log n$ достаточно велика, предъявить явным образом бесконечную последовательность таких значений не удаётся. Конструктивных нижних оценок, качественно сильнее неравенства

$$l(n) \geqslant \log n + \log s(n) - 2{,}13$$

(здесь $s(n)$ — число единиц в двоичной записи числа n), установленного в 1975 г. А. Шенхаге [18], до сих пор, по-видимому, не получено.

Такая ситуация (принципиальная разница между типичным значением сложности и известными нижними оценками сложности для конкретных представителей) характерна для схемных вычислений [19, 20].

Сравним результаты о типичных значениях сложности вычисления степеней с аналогичными результатами в хорошо изученной задаче о сложности реализации булевых функций схемами из функциональных элементов (логическими схемами, булевыми схемами).

О. Б. Лупановым найдена (см., например, [21], где можно найти все необходимые определения) асимптотика роста функции Шеннона $L_B(n)$, характеризующей сложность (минимальное число элементов) реализации схемами в произвольном полном конечном базисе B самой сложнореализуемой функции от n переменных. Из этого результата для произвольного полного базиса B_1, состоящего из функций, зависящих не более чем от двух переменных, можно извлечь следующие оценки:

$$L_{B_1}(n) \leqslant \frac{H_1(n)}{\log H_1(n)} \left(1 + (3 + \mathrm{o}(1)) \frac{\log\log H_1(n)}{\log H_1(n)}\right),$$

$$L_{B_1}(n) \geqslant \frac{H_1(n)}{\log H_1(n)} \left(1 + (1 + \mathrm{o}(1)) \frac{\log\log H_1(n)}{\log H_1(n)}\right), \tag{3}$$

где $H_1(n)$ — двоичный логарифм числа объектов из реализуемого класса (множества булевых функций от n переменных), т. е. $H_1(n) = 2^n$. Отметим, что нижняя оценка величины $L_B(n)$ выводится стандартным мощностным методом, восходящим применительно к рассматриваемой тематике, по-видимому, к работе [22], и справедлива для «почти всех» функций от n переменных, в то время как все известные нижние оценки для эффективно задаваемых функций не более чем линейны.

С. А. Ложкиным в работе [23] дано краткое схематичное описание метода, позволяющего следующим образом усилить верхнюю оценку:

$$L_{B_1}(n) \leqslant \frac{H_1(n)}{\log H_1(n)} \left(1 + (1 + \varkappa_{B_1} + \mathrm{o}(1)) \frac{\log\log H_1(n)}{\log H_1(n)}\right),$$

где $\varkappa_B = 1$ в случае, когда базис B симметричный [3], и $\varkappa_B = 0$ в остальных случаях.

В [24] анонсировано получение такой верхней оценки, которая вместе с мощностной нижней оценкой (1) даёт равенство

$$L_{B_1}(n) = \frac{H_1(n)}{\log H_1(n)} \left(1 + (1 + \mathrm{o}(1)) \frac{\log\log H_1(n)}{\log H_1(n)}\right). \tag{4}$$

Таким образом, если следовать [24], то для функции Шеннона сложности реализации функций алгебры логики схемами над произвольным полным конечным базисом мощностная нижняя оценка, записанная в виде (3), принципиально неулучшаема.

Однако для ещё одной модели схемных вычислений — сборки слов схемами конкатенации [25] — выявлена возможность усиления аналога мощностной нижней оценки из соотношения (4).

Схемы конкатенации можно рассматривать как схемы, имеющие два входа, на которые подаются символы 0 и 1, а каждому элементу схемы соответствует операция конкатенации. Под величиной $L^c(S)$ понимается общее число элементов схемы S, сложность двоичного слова $\tilde{\alpha}$ определяется равенством $L^c(\tilde{\alpha}) = \min L^c(S)$, где минимум берётся по всем схемам конкатенации, реализующим слово $\tilde{\alpha}$, а соответствующая функция

Шеннона — равенством $L^c(n) = \max L^c(\tilde{\alpha})$, где максимум берётся по всем двоичным словам $\tilde{\alpha}$ длины n.

При аккуратном применении известных методов можно получить следующие верхнюю и нижнюю оценки:

$$L^c(n) \leqslant \frac{H_2(n)}{\log H_2(n)} \left(1 + (2 + \mathrm{o}(1)) \frac{\log\log H_2(n)}{\log H_2(n)}\right),$$

$$L^c(n) \geqslant \frac{H_2(n)}{\log H_2(n)} \left(1 + (1 + \mathrm{o}(1)) \frac{\log\log H_2(n)}{\log H_2(n)}\right), \tag{5}$$

где $H_2(n)$ — двоичный логарифм числа объектов из реализуемого класса (слов длины n), т. е. $H_2(n) = n$.

Мощностная нижняя оценка (5) имеет такой же вид, что и нижняя оценка (3). Однако в отличие от случая реализации булевых функций, для задачи о сложности сборки слов схемами конкатенации нижняя оценка может быть усилена [25]: при $n \to \infty$ справедливо равенство

$$L^c(n) = \frac{H_2(n)}{\log H_2(n)} \left(1 + (2 + \mathrm{o}(1)) \frac{\log\log H_2(n)}{\log H_2(n)}\right). \tag{6}$$

Возвращаясь к исходной задаче об аддитивных цепочках, заметим, что из результатов Брауэра (неравенство (1)) и Эрдёша для соответствующей функции Шеннона $L^a(n)$ аддитивной сложности множества натуральных чисел, не превосходящих n, определяемой равенством

$$L^a(n) = \max_{k:k\leqslant n} l(k),$$

непосредственно следуют оценки

$$L^a(n) \leqslant H_3(n) + \frac{H_3(n)}{\log H_3(n)} \left(1 + (2 + \mathrm{o}(1)) \frac{\log\log H_3(n)}{\log H_3(n)}\right),$$

$$L^a(n) \geqslant H_3(n) + (1 - \mathrm{o}(1)) \frac{H_3(n)}{\log H_3(n)}, \tag{7}$$

где $H_3(n)$ — двоичный логарифм числа натуральных чисел, не превосходящих n, т. е. $H_3(n) = \log n$.

Возникает вопрос: можно ли в нижней оценке (7) величины $L^a(n)$ мощностную составляющую усилить подобно нижней оценке (6) или хотя бы подобно нижним оценкам (3) и (5)? Этот вопрос оказался весьма трудным. Трудности связаны с наличием в оценке (7) слагаемых двух типов, которое приводит к бóльшему разнообразию схем с заданной «мощностной» составляющей сложности. В этом направлении удалось получить следующее продвижение.

Обозначим через $R(m, \varepsilon)$ число возрастающих аддитивных цепочек

$$1 = a_0, a_1, a_2, \ldots, a_r, \tag{8}$$

удовлетворяющих условиям

$$\lfloor \log a_r \rfloor = m; \tag{9}$$

$$r \leqslant m + \frac{m}{\log m} - (2 + \varepsilon) \frac{m \log\log m}{(\log m)^2}. \tag{10}$$

Лемма 1. Для любого положительного значения ε при $m \to \infty$ выполняется соотношение

$$\frac{R(m,\varepsilon)}{2^m} \to 0.$$

Эта лемма является основным содержательным утверждением работы, её доказательство приводится в п. 2. Непосредственно из леммы 1 следует

Теорема 1. Пусть $n \to \infty$. Тогда

$$L^a(n) \geqslant \log n + \frac{\log n}{\log\log n}\left(1 - (2+\mathrm{o}(1))\frac{\log\log\log n}{\log\log n}\right).$$

Следствие 1. При $n \to \infty$ выполняется соотношение

$$\left|L^a(n) - \left(\log n + \frac{\log n}{\log\log n}\right)\right| \leqslant (2+\mathrm{o}(1))\frac{\log n \log\log\log n}{(\log\log n)^2}.$$

Последнее неравенство, непосредственно вытекающее из теоремы 1 и верхней оценки Брауэра, может быть переписано в удобном для сравнения с оценками функций Шеннона сложности реализации булевых функций и двоичных наборов виде таким образом:

$$\left|L^a(n) - \left(H_3(n) + \frac{H_3(n)}{\log H_3(n)}\right)\right| \leqslant (2+\mathrm{o}(1))\frac{H_3(n)\log\log H_3(n)}{(\log H_3(n))^2}. \tag{11}$$

Итак, установлена нижняя оценка, имеющая для почти всех значений n точно такое же по абсолютной величине отклонение от $\log n + (\log n)/(\log\log n)$, как и отклонение в верхней оценке Брауэра из (1).

2. Доказательство основной леммы

В основе доказательства леммы 1 лежат рассуждения из [16] (см. также [1, разд. 4.6.3]).

Положим

$$\delta = \delta(m) = 2^{\frac{\varepsilon}{4}\frac{\log\log m}{\log m}} - 1. \tag{12}$$

Поскольку $\delta \to 0$ при $m \to \infty$, будем считать, что

$$\delta < \sqrt{2} - 1. \tag{13}$$

Разобьём r шагов произвольной возрастающей цепочки (8), удовлетворяющей условиям (9) и (10), на три класса.

Шаг i отнесём к *первому классу*, если этот шаг является удвоением, т. е. $a_i = 2a_{i-1}$.

Шаг i отнесём ко *второму классу*, если

$$a_i < 2a_{i-1},$$
$$a_i \geqslant (1+\delta)^{i-j}a_j \text{ для всех } j \ \ (0 \leqslant j < i).$$

Шаг i отнесём к *третьему классу*, если

$$a_i < 2a_{i-1},$$
$$a_i < (1+\delta)^{i-j}a_j \text{ для некоторого } j \ \ (0 \leqslant j < i). \tag{14}$$

Шаги, отнесённые к первому классу, будем называть удвоениями, а к третьему — мини-шагами. Количество шагов в классах обозначим соответственно через u_1, u_2 и u_3. Очевидно, что $u_1 + u_2 + u_3 = r$.

Оценим сверху величину $u_2 + u_3$. Если шаг i отнесён ко второму или третьему классу, то $a_i \neq 2a_{i-1}$, т. е. $a_i \leqslant a_{i-1} + a_{i-2} \leqslant 3a_{i-2}$. Используя (9) и (13), получаем

$$2^m \leqslant a_r < 3^{(u_2+u_3)/2} 2^{u_1} = \frac{2^{u_1+u_2+u_3}}{(4/3)^{(u_2+u_3)/2}} = \frac{2^r}{2^{(u_2+u_3+u_4)\log(2/\sqrt{3})}}.$$

Поэтому в силу (10)

$$u_2 + u_3 < \frac{1}{\log(2/\sqrt{3})}(r-m) \leqslant \frac{1}{\log(2/\sqrt{3})}\frac{2m}{\log m} < 10\frac{m}{\log m}.$$

Теперь оценим сверху величину u_3. Для каждого мини-шага i_s ($s = 1, \ldots, u_3$) при соответствующем j_s ($0 \leqslant j_s \leqslant i_s$) в силу (14) выполняется неравенство $a_{i_s} < a_{j_s}(1+\delta)^{i_s - j_s}$. Пусть $I_1, \ldots, I_{u_3}$ — полуинтервалы $(j_1, i_1], \ldots, (j_{u_3}, i_{u_3}]$, где $(j, i]$ — множество целых чисел ρ, таких, что $j < \rho \leqslant i$. Построим систему неперекрывающихся полуинтервалов $J_1 = (j'_1, i'_1], \ldots, J_h = (j'_h, i'_h]$, такую, что

$$I_1 \cup \ldots \cup I_{u_2} = J_1 \cup \ldots \cup J_h,$$
$$a_{i'_s} < a_{j'_s}(1+\delta)^{2(i'_s - j'_s)} \text{ для } 1 \leqslant s \leqslant h.$$

Пусть $i_1 < i_2 < \ldots < i_{u_2}$. Удалим все полуинтервалы I_t, которые можно удалить, не сужая множество $I_1 \cup \ldots \cup I_{u_2}$. Каждую систему перекрывающихся полуинтервалов $(j_c, i_c], \ldots, (j_d, i_d]$ объединим в один полуинтервал $(j', i'] = (j_c, i_d]$. Заметим, что

$$a_{i'} < a_{j'}(1+\delta)^{i_c - j_c + \ldots + i_d - j_d} \leqslant a_{j'}(1+\delta)^{2(i'-j')},$$

так как каждая точка полуинтервала $(j', i']$ покрыта полуинтервалами $(j_c, i_c]$, $\ldots$, $(j_d, i_d]$ не более чем дважды.

Положим $q = (i'_1 - j'_1) + \ldots + (i'_h - j'_h)$. Обозначим через u'_1 и u'_2 число шагов, относящихся соответственно к первому и второму классам и имеющих номера, принадлежащие множеству $J_1 \cup \ldots \cup J_h$. Отметим, что $u'_1 + u'_2 + u_3 = q$. Тогда справедливы соотношения

$$2^m \leqslant a_r \leqslant 2^{u_1 - u'_1} 3^{(u_2 - u'_2)/2} (1+\delta)^{2q} =$$
$$= 2^{u_1-u'_1}\left(\frac{\sqrt{3}}{2}\right)^{u_2-u'_2} 2^{u_2-u'_2}\left(\frac{(1+\delta)^2}{2}\right)^{u'_1+u'_2+u_3} 2^{u'_1+u'_2+u_3} = 2^r\left(\frac{\sqrt{3}}{2}\right)^{u_2-u'_2}\left(\frac{(1+\delta)^2}{2}\right)^{u'_1+u'_2+u_3}.$$

Отсюда и из (9) и (12) при всех достаточно больших значениях m получаем

$$u_3 \leqslant \frac{r-m}{1-2\log(1+\delta)} - \frac{1-(\log 3)/2}{1-2\log(1+\delta)}(u_2 - u'_2) - u'_1 - u'_2 \leqslant$$
$$\leqslant \frac{m/\log m - (2+\varepsilon)m\log\log m/(\log m)^2}{1-\varepsilon\log\log m/(4\log m)} - \left(1 - \frac{1}{2}\log 3\right)(u_2 - u'_2) - u'_1 - u'_2 <$$
$$< \left(1 + \frac{\varepsilon}{3}\frac{\log\log m}{\log m}\right)\left(\frac{m}{\log m} - (2+\varepsilon)\frac{m\log\log m}{(\log m)^2}\right) - \left(1 - \frac{1}{2}\log 3\right)u_2 - u'_1 \leqslant$$
$$\leqslant \frac{m}{\log m} - (2+\varepsilon/2)\frac{m\log\log m}{(\log m)^2} - \left(1 - \frac{1}{2}\log 3\right)u_2 - u'_1.$$

Таким образом, если для произвольной возрастающей цепочки (8) выполняются неравенства (9) и (10), то эта цепочка удовлетворяет следующей системе условий:

$$\begin{gathered} u_3 \leqslant \frac{m}{\log m} - (2+\varepsilon/2)\frac{m\log\log m}{(\log m)^2} - \left(1 - \frac{1}{2}\log 3\right) u_2, \\ u_2 + u_3 \leqslant 10\frac{m}{\log m}, \\ u_1 + u_2 + u_3 \leqslant m + \frac{m}{\log m} - (2+\varepsilon)\frac{m\log\log m}{(\log m)^2}. \end{gathered} \tag{15}$$

Отметим, что везде, за исключением выкладок в (36), вместо первого неравенства из совокупности (15) достаточно использовать следующую более грубую оценку величины u_3:

$$u_3 \leqslant \frac{m}{\log m} - (2+\varepsilon/2)\frac{m\log\log m}{(\log m)^2}.$$

Пусть теперь заданы значения u_1, u_2, u_3, удовлетворяющие условиям (15). Тогда существует не более

$$\binom{u_1+u_2+u_3}{u_2+u_3}\binom{u_2+u_3}{u_2} \tag{16}$$

способов отнести каждый шаг к тому или иному классу шагов. После того как распределение шагов задано, оценим число возможностей выбора самих шагов. Отметим, что шаги, отнесённые к первому классу, полностью определяются своими номерами.

Сначала оценим сверху число способов выбора шагов, отнесённых ко второму классу. Если i-й шаг отнесён ко второму классу, то в силу (13) выполняется неравенство $a_i \geqslant (1+\delta)a_{i-1}$. Пусть $a_i = a_j + a_k$, где $a_j \geqslant a_k$. Тогда

$$\delta a_{i-1} \leqslant a_k \leqslant a_j \leqslant a_{i-1}. \tag{17}$$

Кроме того, в силу (13)

$$a_j \leqslant \frac{a_i}{(1+\delta)^{i-j}} \leqslant \frac{2a_{i-1}}{(1+\delta)^{i-j}}, \quad a_k \leqslant \frac{a_i}{(1+\delta)^{i-k}} \leqslant \frac{2a_{i-1}}{(1+\delta)^{i-k}}. \tag{18}$$

Из (17) и (18) получаем $\delta \leqslant \dfrac{2}{(1+\delta)^{i-j}}$, $\delta \leqslant \dfrac{2}{(1+\delta)^{i-k}}$.

Пусть для некоторого натурального β выполняется неравенство $\delta \leqslant 2/(1+\delta)^{\beta}$. Тогда $\delta(1+\beta\delta) \leqslant 2$. Отсюда в силу (12) получаем

$$\beta \leqslant \frac{2-\delta}{\delta^2} \leqslant \frac{2}{\delta^2} = \frac{2}{\left(2^{\frac{\varepsilon}{4}\frac{\log\log m}{\log m}} - 1\right)^2} \leqslant \frac{32(\log m)^2}{\varepsilon^2(\ln 2)^2(\log\log m)^2}.$$

Таким образом, имеется не более

$$\left(\beta^2\right)^{u_2} \leqslant \left(\frac{8\log m}{\varepsilon \ln 2 \log\log m}\right)^{4u_2} \tag{19}$$

возможностей для выбора шагов из второго класса.

Оценим сверху число способов выбора шагов, отнесённых к третьему классу. После того как определены все шаги из второго класса, имеется не более

$$\binom{r^2}{u_3} \tag{20}$$

возможностей выбора u_3 пар индексов $(j_1, k_1), \ldots, (j_{u_3}, k_{u_3})$ для шагов третьего класса. В порядке возрастания номера i для каждого шага из третьего класса используем такую пару (j_h, k_h) из множества $\{(j_1, k_1), \ldots, (j_{u_3}, k_{u_3})\}$, для которой выполнены условия $j_h < i$, $k_h < i$ и величина $\max(j_h, k_h)$ принимает наименьшее значение. Если такой пары не существует, или после её использования цепочка перестанет быть возрастающей, или получившийся шаг не является мини-шагом, то аддитивная цепочка не будет получена. При этом любая возрастающая аддитивная цепочка с шагами из третьего класса на назначенных местах будет получена одним из указанных способов.

Учитывая это, получаем, что величина (20) даёт верхнюю оценку для числа вариантов выбора шагов, отнесённых к третьему классу. Таким образом, число возрастающих аддитивных цепочек с заданным количеством u_1, u_2 и u_3 шагов, отнесённых к первому, второму и третьему классам соответственно, не больше произведения величин (16), (19) и (20), т. е. не превосходит величины

$$\binom{u_1+u_2+u_3}{u_2+u_3}\binom{u_2+u_3}{u_2}\left(\frac{8\log m}{\varepsilon \ln 2 \log\log m}\right)^{4u_2}\binom{(u_1+u_2+u_3)^2}{u_3}.$$

Поэтому

$$R(m,\varepsilon) \leqslant \sum \binom{u_1+u_2+u_3}{u_2+u_3}\binom{u_2+u_3}{u_2}\left(\frac{8\log m}{\varepsilon \ln 2 \log\log m}\right)^{4u_2}\binom{(u_1+u_2+u_3)^2}{u_3}, \tag{21}$$

где сумма берется по всем u_1, u_2 и u_3, удовлетворяющим условиям (15).

Для завершения доказательства леммы 1 достаточно установить, что

$$\log R(m,\varepsilon) - m \to -\infty$$

при $m \to \infty$. Отдельно оценим сверху каждый множитель (точнее, его логарифм) в слагаемых из правой части неравенства (21).

Оценивая первый множитель, получаем

$$\binom{u_1+u_2+u_3}{u_2+u_3} \leqslant \frac{(u_1+u_2+u_3)^{u_2+u_3}}{(u_2+u_3)!} \leqslant 3^{u_2+u_3}\left(\frac{u_1+u_2+u_3}{u_2+u_3}\right)^{u_2+u_3}. \tag{22}$$

Далее рассмотрим два случая. Если выполняется неравенство

$$u_2+u_3 \leqslant \frac{m}{\log m \log\log m}, \tag{23}$$

то в силу (15), (22) и (23) справедливы соотношения

$$\begin{aligned}\log\binom{u_1+u_2+u_3}{u_2+u_3} &\leqslant (u_2+u_3)\log 3 + (u_2+u_3)\log(2m) \leqslant \\ &\leqslant \frac{m}{\log\log m} + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right).\end{aligned} \tag{24}$$

Если же выполняется неравенство

$$u_2 + u_3 > \frac{m}{\log m \log\log m}, \tag{25}$$

то, используя (15), (22) и (25), получаем

$$\begin{gathered}\log \binom{u_1+u_2+u_3}{u_2+u_3} \leqslant (u_2+u_3)\log 3 + (u_2+u_3)\log(2\log m \log\log m) \leqslant \\ \leqslant (u_2+u_3)\log\log m + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right).\end{gathered} \tag{26}$$

Логарифм второго множителя произвольного слагаемого из правой части неравенства (21) оценить, учитывая (15), совсем просто:

$$\log \binom{u_2+u_3}{u_2} \leqslant u_2 + u_3 = \mathrm{O}\left(\frac{m}{\log m}\right) = \mathrm{o}\left(\frac{m\log\log m}{\log m}\right). \tag{27}$$

Теперь для всех достаточно больших значений m оценим сверху логарифм третьего множителя:

$$\log\left(\left(\frac{8\log m}{\varepsilon \ln 2 \log\log m}\right)^{4u_2}\right) \leqslant 4u_2 \log\log m. \tag{28}$$

И наконец, оценим логарифм четвертого множителя произвольного слагаемого из правой части неравенства (21):

$$\binom{(u_1+u_2+u_3)^2}{u_3} \leqslant \frac{(u_1+u_2+u_3)^{2u_3}}{u_3!} \leqslant 3^{u_3}\left(\frac{(u_1+u_2+u_3)^2}{u_3}\right)^{u_3}. \tag{29}$$

Далее рассмотрим два случая. Если выполняется неравенство

$$u_3 \leqslant \frac{m}{\log m \log\log m}, \tag{30}$$

то в силу (15), (29) и (30) справедливы соотношения

$$\log \binom{(u_1+u_2+u_3)^2}{u_3} \leqslant u_3\log 3 + u_3\log(4m^2) \leqslant \frac{2m}{\log\log m} + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right). \tag{31}$$

Если же выполняется неравенство

$$u_3 > \frac{m}{\log m \log\log m}, \tag{32}$$

то, используя (15) и (32), получаем

$$\begin{gathered}\log \binom{(u_1+u_2+u_3)^2}{u_3} \leqslant u_3\log 3 + u_3\log(4m\log m\log\log m) = \\ = u_3\log m + u_3\log\log m + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right).\end{gathered} \tag{33}$$

Для оценки величины $R(m,\varepsilon)$ разобьём все слагаемые суммы из (21) на три группы в зависимости от того, каким условиям, помимо (15), удовлетворяют значения u_1, u_2 и u_3. Условием вхождения в первую группу является выполнение неравенства

$$u_3 \leqslant \frac{m}{2\log m};$$

во вторую группу — неравенств

$$u_3 > \frac{m}{2\log m}, \quad u_2 \leqslant \frac{m}{\log m \log\log\log m};$$

в третью группу — выполнение неравенств

$$u_3 > \frac{m}{2\log m}, \quad u_2 > \frac{m}{\log m \log\log\log m}.$$

Оценим сверху величины слагаемых в каждой группе.

Если слагаемое из правой части неравенства (21) отнесено к первой группе, то, применяя максимальную из оценок (24) и (26), оценки (27) и (28), а также максимальную из оценок (31) и (33), с использованием (15) и (2) имеем

$$\begin{gathered}
\log\left(\binom{u_1+u_2+u_3}{u_2+u_3}\binom{u_2+u_3}{u_2}\left(\frac{8\log m}{\varepsilon\ln 2\log\log m}\right)^{4u_2}\binom{(u_1+u_2+u_3)^2}{u_3}\right)\leqslant \\
\leqslant \max\left(\frac{m}{\log\log m},\,(u_2+u_3)\log\log m\right)+4u_2\log\log m+ \\
+\max\left(\frac{2m}{\log\log m},\,u_3\log m+u_3\log\log m\right)+\mathrm{o}\left(\frac{m\log\log m}{\log m}\right)\leqslant \\
\leqslant u_3\log m+5u_2\log\log m+2u_3\log\log m+\mathrm{O}\left(\frac{m}{\log\log m}\right)\leqslant\frac{m}{2}+\mathrm{O}\left(\frac{m}{\log\log m}\right).
\end{gathered} \tag{34}$$

Если слагаемое из правой части неравенства (21) отнесено ко второй группе, то выполняются неравенства (25) и (32). Применяя оценки (26), (27), (28) и (33), с использованием (15) и (2) имеем

$$\begin{gathered}
\log\left(\binom{u_1+u_2+u_3}{u_2+u_3}\binom{u_2+u_3}{u_2}\left(\frac{8\log m}{\varepsilon\ln 2\log\log m}\right)^{4u_2}\binom{(u_1+u_2+u_3)^2}{u_3}\right)\leqslant \\
\leqslant (u_2+u_3)\log\log m+4u_2\log\log m+u_3\log m+u_3\log\log m+\mathrm{o}\left(\frac{m\log\log m}{\log m}\right)\leqslant \\
\leqslant u_3\log m+5u_2\log\log m+2u_3\log\log m+\mathrm{o}\left(\frac{m\log\log m}{\log m}\right)\leqslant \\
\leqslant\left(\frac{m}{\log m}-(2+\varepsilon/2)\frac{m\log\log m}{(\log m)^2}\right)\log m+ \\
+2\left(\frac{m}{\log m}-(2+\varepsilon/2)\frac{m\log\log m}{(\log m)^2}\right)\log\log m+\mathrm{o}\left(\frac{m\log\log m}{\log m}\right)\leqslant \\
\leqslant m-\frac{\varepsilon}{2}\frac{m\log\log m}{\log m}+\mathrm{o}\left(\frac{m\log\log m}{\log m}\right).
\end{gathered} \tag{35}$$

Если слагаемое из правой части неравенства (21) отнесено к третьей группе, то также выполняются неравенства (25) и (32). Применяя оценки (26), (27), (28) и (33), с использованием (15) и (2) имеем

$$\log\left(\binom{u_1+u_2+u_3}{u_2+u_3}\binom{u_2+u_3}{u_2}\left(\frac{8\log m}{\varepsilon\ln 2\log\log m}\right)^{4u_2}\binom{(u_1+u_2+u_3)^2}{u_3}\right)\leqslant$$

$$\leqslant (u_2+u_3)\log\log m + 4u_2\log\log m + u_3\log m + u_3\log\log m + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right)\leqslant$$

$$\leqslant u_3\log m + \mathrm{O}\left(\frac{m\log\log m}{\log m}\right)\leqslant \tag{36}$$

$$\leqslant \left(\frac{m}{\log m} - (2+\varepsilon/2)\frac{m\log\log m}{(\log m)^2} - \left(1-\frac{1}{2}\log 3\right)u_2\right)\log m + \mathrm{O}\left(\frac{m\log\log m}{\log m}\right) <$$

$$< m - \left(1-\frac{1}{2}\log 3\right)\frac{m}{\log\log\log m} + \mathrm{O}\left(\frac{m\log\log m}{\log m}\right).$$

Таким образом, из (34), (35) и (36) следует, что двоичный логарифм любого слагаемого из (21) не превосходит величины

$$m - \frac{\varepsilon}{2}\frac{m\log\log m}{\log m} + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right).$$

Поэтому, учитывая, что в сумме из правой части неравенства (21) не более $8m^3$ слагаемых, получаем

$$\log R(m,\varepsilon) - m \leqslant -\frac{\varepsilon}{2}\frac{m\log\log m}{\log m} + \mathrm{o}\left(\frac{m\log\log m}{\log m}\right),$$

а последнее выражение стремится к $-\infty$ при $m\to\infty$. Лемма 1 доказана.

Заключение

Помимо классических аддитивных цепочек, в различных криптографических приложениях, в первую очередь в алгоритмах, связанных с быстрыми вычислениями на эллиптичских кривых, используется также аппарат цепочек из сложений и вычитаний [26, 27]. Утверждение теоремы 1 можно перенести и на такой класс вычислений.

Обозначим через $l^{as}(n)$ наименьшую длину цепочки из сложений и вычитаний для числа n (формальное определение цепочки из сложений и вычитаний отличается от определения аддитивной цепочки только тем, что произвольный шаг цепочки имеет вид либо $a_k = a_i + a_j$, либо $a_k = a_i - a_j$), а через $\hat{l}^{as}(n)$ — наименьшую длину цепочки для числа n, состоящей из шагов вида $a_k = a_i + a_j$, $a_k = a_i - a_j$ и $a_k = -a_i - a_j$. Положим

$$L^{as}(n) = \max_{k\,:\,k\leqslant n} l^{as}(k), \quad \hat{L}^{as}(n) = \max_{k\,:\,k\leqslant n} \hat{l}^{as}(k).$$

Теорема 2. Пусть $n\to\infty$. Тогда

$$L^{as}(n) \geqslant \hat{L}^{as}(n) \geqslant \log n + \frac{\log n}{\log\log n}\left(1-(2+\mathrm{o}(1))\frac{\log\log\log n}{\log\log n}\right).$$

Доказательство теоремы 2 несущественно сложнее доказательства теоремы 1. Технические особенности, связанные с наличием одной или двух дополнительных операций, можно проследить на примере доказательства нижней оценки из [28].

ЛИТЕРАТУРА

1. *Кнут Д. Е.* Искусство программирования. Т. 2. 3-е изд. М.: Издательский дом «Вильямс», 2000.
2. *Scholz A.* Jahresbericht // Deutsche Mathematiker-Vereinigung. 1937. B. 47. S. 41–42.
3. *Subbarao M. V.* Addition chains — some results and problems // Number Theory and Applications / ed. R. A. Mollin. NATO Advanced Science Institutes Series: Ser. C. Kluwer Academic Publisher Group, 1989. V. 265. P. 555–574.
4. *Bos J. and Coster M.* Addition chain heuristics // Crypto'89. LNCS. 1990. V. 435. P. 400–407.
5. *Gordon D. M.* A survey of fast exponentiation methods // J. Algorithms. 1998. V. 27. P. 129–146.
6. *Thurber E. G.* Efficient generation of minimal length addition chains // SIAM J. Comput. 1999. V. 28. P. 1247–1263.
7. *Bernstein D. J.* Pippenger's exponentiation algorithm. `http://cr.yp.to/papers/pippenger.pdf`, 2002.
8. *Гашков С. Б.* Задача об аддитивных цепочках и ее обобщения // Математическое просвещение. Третья серия. Вып. 15. М.: МЦНМО, 2011. С. 138–153.
9. *Clift N. M.* Calculating optimal addition chains // Computing. 2011. V. 91. P. 265–284.
10. *Кочергин В. В.* Уточнение оценок сложности вычисления одночленов и наборов степеней в задачах Беллмана и Кнута // Дискретный анализ и исследование операций. 2014. Т. 21. № 6. С. 51–72.
11. *Järvinen K, Dimitrov V., and Azarderakhsh R.* A generalization of addition chains and fast inversions in binary fields // IEEE Trans. Computers. 2015. V. 64(9). P. 2421–2432.
12. *Von zur Gathen J. and Nocker M.* Exponentiation in finite fields: theory and practice // LNCS. 1997. V. 1255. P. 88–113.
13. *Смарт Н.* Криптография. М.: Техносфера, 2005.
14. *Гашков С. Б., Сергеев И. С.* О применении метода аддитивных цепочек к инвертированию в конечных полях // Дискретная математика. 2006. Т. 18. № 4. С. 56–72.
15. *Brauer A.* On addition chains // Bull. Amer. Math. Soc. 1939. V. 45. P. 736–739.
16. *Erdos P.* Remarks on number theory, III: On addition chains // Acta Arith. 1960. V. 6. P. 77–81.
17. *Ильин А. М.* Об аддитивных цепочках чисел // Проблемы кибернетики. Вып. 13. М.: Физматлит, 1965. С. 245–248.
18. *Schönhage A. A.* Lower bound for the length of addition chains // Theor. Comput. Sci. 1975. V. 1. P. 1–12.
19. *Сэвидж Д. Е.* Сложность вычислений. М.: Факториал, 1998.
20. *Нигматуллин Р. Г.* Сложность булевых функций. М.: Наука, 1991.
21. *Лупанов О. Б.* Асимптотические оценки сложности управляющих систем. М.: Изд-во МГУ, 1984.
22. *Riordan J. and Shannon C. E.* The number of two-terminal series-parallel networks // J. Math. Phys. Mass. Inst. Tech. 1942. V. 21. No. 2. P. 83–93. Рус. пер.: Риодан Дж., Шеннон К. Число двухполюсных параллельно-последовательных сетей / сб. Шеннон К. Работы по теории информации и кибернетике. М.: ИЛ, 1963. С. 46–58.
23. *Ложкин С. А.* Оценки высокой степени точности для сложности управляющих систем из некоторых классов // Математические вопросы кибернетики. Вып. 6. М.: Наука, 1996. С. 189–214.
24. *Ложкин С. А.* Асимптотические оценки высокой степени точности для сложности реализации булевских функций схемами из функциональных элементов // Труды II Меж-

дунар. конф. «Дискретные модели в теории управляющих систем» (23–28 июня 1997 г.). М.: Диалог-МГУ, 1997. С. 37–39.

25. *Кочергин В. В., Кочергин Д. В.* Уточнение асимптотического поведения сложности сборки слов схемами конкатенации // Вестн. Моск. ун-та. Сер. 1. Матем. Механ. 2016. № 2. С. 12–18.
26. *Volger H.* Some results on addition/subtraction chains // Inform. Proc. Lett. 1985. V. 20. P. 155–160.
27. *Morain F. and Olivos J.* Speeding up the computation on an elliptic curve using addition-subtraction chains // Informatique Théorique et Applications. 1990. V. 24. P. 531–544.
28. *Кочергин В. В.* О сложности вычислений одночленов и наборов степеней // Дискретный анализ. Новосибирск: Изд-во Института математики СО РАН, 1994. С. 94–107.

REFERENCES

1. *Knuth D. E.* The art of computer programming, vol. 2, third edition. Reading, Massachusetts, Addison-Wesley, 1997.
2. *Scholz A.* Jahresbericht. Deutsche Mathematiker-Vereinigung, 1937, B. 47, S. 41–42.
3. *Subbarao M. V.* Addition chains — some results and problems. Number Theory and Applications. Ed. R. A. Mollin. NATO Advanced Science Institutes Series, Ser. C. Kluwer Academic Publisher Group, 1989, vol. 265, pp. 555–574.
4. *Bos J. and Coster M.* Addition chain heuristics. Crypto'89, LNCS, 1990, vol. 435, pp. 400–407.
5. *Gordon D. M.* A survey of fast exponentiation methods. J. Algorithms, 1998, vol. 27, pp. 129–146.
6. *Thurber E. G.* Efficient generation of minimal length addition chains. SIAM J. Comput., 1999, vol. 28, pp. 1247–1263.
7. *Bernstein D. J.* Pippenger's exponentiation algorithm. `http://cr.yp.to/papers/pippenger.pdf`, 2002.
8. *Gashkov S. B.* Zadacha ob additivnykh tsepochkakh i eye obobshcheniya [Addition chains problem and its generalizations]. Matematicheskoye Prosveshcheniye. Third series, vol. 15. Moscow, MTsNMO, 2011, pp. 138–153. (in Russian)
9. *Clift N. M.* Calculating optimal addition chains. Computing, 2011, vol. 91, pp. 265–284.
10. *Kochergin V. V.* Improvement of the estimates of the computational complexity for monomials and sets of powers in Bellman's and Knuth's problems. J. Appl. Industr. Math., 2015, vol. 9, no. 1, pp. 68–82.
11. *Järvinen K, Dimitrov V., and Azarderakhsh R.* A generalization of addition chains and fast inversions in binary fields. IEEE Trans. Computers, 2015, vol. 64(9), pp. 2421–2432.
12. *Von zur Gathen J. and Nocker M.* Exponentiation in finite fields: theory and practice. LNCS, 1997, vol. 1255, pp. 88–113.
13. *Smart N.* Cryptography: An Introduction (3rd edition). McGraw — Hill, 2003.
14. *Gashkov S. B. and Sergeev I. S.* An application of the method of additive chains to inversion in finite fields. Discr. Math. Appl., 2006, vol. 16, no. 6, pp. 601-618.
15. *Brauer A.* On addition chains. Bull. Amer. Math. Soc., 1939, vol. 45, pp. 736–739.
16. *Erdos P.* Remarks on number theory, III: On addition chains. Acta Arith., 1960, vol. 6, pp. 77–81.
17. *Il'in A. M.* Ob additivnykh tsepochkakh chisel [On addition chains of numbers]. Problemy Kibernetiki, vol. 13. Moscow, Fizmatlit Publ., 1965, pp. 245–248. (in Russian)
18. *Schönhage A. A.* Lower bound for the lenght of addition chains. Theor. Comput. Sci., 1975, vol. 1, pp. 1–12.
19. *Savage J. E.* The Complexity of Computing. New York, Wiley, 1976.

20. *Nigmatullin R. G.* Slozhnost' bulevykh funktsiy [The Complexity of Boolean Functions]. Moscow, Nauka Publ., 1991. (in Russian)
21. *Lupanov O. B.* Asimptoticheskie otsenki slozhnosti upravlyayushchikh sistem [Asymptotic Estimations of Complexity of Control Systems]. Moscow, MSU Publ., 1984. (in Russian)
22. *Riordan J. and Shannon C. E.* The number of two-terminal series-parallel networks. J. Math. Phys. Mass. Inst. Tech., 1942, vol. 21, no. 2, pp. 83–93.
23. *Lozhkin S. A.* Otsenki vysokoy stepeni tochnosti dlya slozhnosti upravlyayushchikh sistem iz nekotorykh klassov [More accurate estimations of complexity of control systems for some classes]. Matematicheskiye Voprosy Kibernetiki, vol. 6, Moscow, Nauka Publ., 1996, pp. 189–214. (in Russian)
24. *Lozhkin S. A.* Asimptoticheskiye otsenki vysokoy stepeni tochnosti dlya slozhnosti realizatsii bulevskikh funktsiy skhemami iz funktsional'nykh elementov [More accurate asymptotic estimations of complexity of Boolean functions realization by logic circuits]. Proc. II Intern. conf. "Diskretnye modeli v teorii upravlyayushchikh sistem" (22–23 June 1997), Moscow, Dialog-MSU Publ., 1997, pp. 37–39. (in Russian)
25. *Kochergin V. V. and Kochergin D. V.* Revision of asymptotic behavior of the complexity of word assembly by concatenation circuits. Moscow University Math. Bull., 2016, vol. 71, iss. 2, pp. 55–60.
26. *Volger H.* Some results on addition/subtraction chains. Inform. Proc. Lett., 1985, vol. 20, pp. 155–160.
27. *Morain F. and Olivos J.* Speeding up the computation on an elliptic curve using addition-subtraction chains. Informatique Théorique et Applications, 1990, vol. 24, pp. 531–544.
28. *Kochergin V. V.* On the complexity of computations of monomials and tuples of powers. Siberian Adv. Math., 1996, vol. 6, no. 1, pp. 71–86.

СВЕДЕНИЯ ОБ АВТОРАХ

АБРОСИМОВ Михаил Борисович — доктор физико-математических наук, профессор, профессор Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского, г. Саратов. E-mail: **mic@rambler.ru**

АГИБАЛОВ Геннадий Петрович — доктор технических наук, профессор, профессор кафедры защиты информации и криптографии Национального исследовательского Томского государственного университета, г. Томск. E-mail: **agibalov@isc.tsu.ru**

АДЕЛЬШИН Александр Владимирович — кандидат физико-математических наук, доцент, старший научный сотрудник Омского филиала Института математики им. С. Л. Соболева СО РАН, г. Омск. E-mail: **adelshin@ofim.oscsbras.ru**

КОЧЕРГИН Вадим Васильевич — доктор физико-математических наук, профессор кафедры дискретной математики механико-математического факультета Московского государственного университета имени М. В. Ломоносова, г. Москва.
E-mail: **vvkoch@yandex.ru**

КОЧЕРГИН Дмитрий Вадимович — студент механико-математического факультета Московского государственного университета имени М. В. Ломоносова, г. Москва. E-mail: **dk@kocherg.in**

КУЧИН Андрей Константинович — аспирант Омского филиала Института математики им. С. Л. Соболева СО РАН, г. Омск. E-mail: **dr- -on@yandex.ru**

МОДЕНОВА Ольга Владимировна — ведущий программист Научно-образовательного центра «Эрудит», г. Саратов. E-mail: **oginiel@rambler.ru**

ПАНОВ Николай Петрович — аспирант Ульяновского государственного университета, г. Ульяновск. E-mail: **nppanov@yandex.ru**

ПОПКОВ Кирилл Андреевич — кандидат физико-математических наук, младший научный сотрудник Института прикладной математики им. М. В. Келдыша РАН, г. Москва. E-mail: **kirill-formulist@mail.ru**

РЫБАЛОВ Александр Николаевич — кандидат физико-математических наук, доцент кафедры компьютерной математики и программирования Омского государственного университета им. Ф. М. Достоевского, г. Омск; старший научный сотрудник Института математики им. С. Л. Соболева СО РАН, г. Новосибирск.
E-mail: **alexander.rybalov@gmail.com**

САФОНОВ Вадим Олегович — студент кафедры защиты информации и криптографии Национального исследовательского Томского государственного университета, г. Томск. E-mail: **vsaffonov.1115@gmail.com**

СТЕФАНЦОВ Дмитрий Александрович — старший преподаватель кафедры защиты информации и криптографии Национального исследовательского Томского государственного университета, г. Томск. E-mail: **d.a.stephantsov@gmail.com**

ЧЕРЕДНИК Игорь Владимирович — преподаватель Московского технологического университета (МИРЭА), г. Москва. E-mail: **p.n.v.k.s@mail.ru**

ШЕВЛЯКОВ Артем Николаевич — кандидат физико-математических наук, Институт математики им. С. Л. Соболева; Омский государственный технический университет, г. Омск. E-mail: **a_shevl@mail.ru**

ШУЛЕЖКО Олеся Владимировна — кандидат физико-математических наук, доцент кафедры информатики Ульяновского государственного педагогического университета имени И. Н. Ульянова, г. Ульяновск. E-mail: **ol.shulezhko@gmail.com**