МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА

Научный журнал

2014 № 3(25)

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

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Адрес редакции: 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 надежности, интеллектуальных системах.

Периодичность выхода журнала: 4 номера в год.

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

Подписано к печати 05.09.2014.
Формат $60 \times 84\frac{1}{8}$. Усл. п. л. 13,4. Уч.-изд. л. 15. Тираж 300 экз.

Издательство ТГУ. 634029, Томск, ул. Никитина, 4

СОДЕРЖАНИЕ

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ДИСКРЕТНОЙ МАТЕМАТИКЕ

ДИСКРЕТНЫЕ МОДЕЛИ РЕАЛЬНЫХ ПРОЦЕССОВ

CONTENTS

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРИКЛАДНОЙ ДИСКРЕТНОЙ МАТЕМАТИКИ

DOI 10.17223/20710410/25/1
УДК 512.53

О РЕГУЛЯРНОСТИ НЕКОТОРЫХ ПОДПОЛУГРУПП МОНОИДА ЭНДОМОРФИЗМОВ ОТНОШЕНИЯ ЭКВИВАЛЕНТНОСТИ

Е. А. Бондарь

Луганский национальный университет имени Тараса Шевченко, г. Луганск, Украина

E-mail: bondareug@gmail.com

Для графов отношения эквивалентности получен ответ на вопрос М. Беттчера и У. Кнауэра, при каких условиях множество полусильных (локально сильных, квазисильных) эндоморфизмов является полугруппой. Найдены условия регулярности таких полугрупп.

Ключевые слова: *регулярность, полугруппа, эндоморфизм, эквивалентность.*

Введение

Полугруппам эндоморфизмов графов различных классов посвящено множество исследований. К примеру, хорошо освещен в литературе вопрос об определяемости графов своими эндоморфизмами. Так, данную проблему изучал Л. М. Глускин [1] для квазипорядков; Л. Б. Шнеперман [2], Ю. М. Важенин [3], Б. В. Попов [4] — для рефлексивных графов; Ж. Араужо и Я. Конечны [5] — для так называемых плотных отношений. Условия регулярности полугрупп эндоморфизмов упорядоченного и квазиупорядоченного множеств исследованы в [6], а для конечных и счётных цепей — в [7, 8]. Копредставление моноида эндоморфизмов конечной цепи найдено в [9]. Как отмечено выше, определение эндоморфизма зачастую рассматривалось с некоторыми дополнительными условиями в зависимости от целей исследования. Различные типы эндоморфизмов собраны в [10] для определения спектра эндоморфизмов и эндотипа. С их помощью можно классифицировать графы [10, 11].

Целый ряд работ китайских математиков посвящен изучению отношений Грина и регулярных элементов полугрупп эндоморфизмов графов эквивалентностей и их подполугрупп: эндоморфизмы изучались в [12, 13], сильные эндоморфизмы в [14], изоморфизмы в [15, 16]. Точное представление моноида эндоморфизмов графа отношения эквивалентности описано в [17], а для 2-нильпотентных отношений — в [18]. Открытым в этом направлении остается вопрос М. Беттчера и У. Кнауэра [10], при каких условиях множество всех полусильных (локально сильных, квазисильных) эндоморфизмов неориентированного графа является полугруппой. В настоящей работе получен ответ на данный вопрос для графов отношения эквивалентности.

Работа построена следующим образом. В п. 1 приводятся необходимые определения и обозначения. В пп. 2–4 изучаются полусильные, локально сильные и квазисильные эндоморфизмы отношения эквивалентности. Получено описание соответствующих эндоморфизмов, найдены необходимые и достаточные условия, когда множество таких эндоморфизмов образует полугруппу, и доказана регулярность этих полугрупп.

1. Предварительные сведения

Пусть $\mathcal{T}(X)$ — симметрическая полугруппа на множестве X, $\varphi \in \mathcal{T}(X)$, $A \subseteq X$ — произвольное непустое подмножество. Через $\varphi|_A$ будем обозначать ограничение φ на множество A; множество всех константных отображений $\nu_t : A \to X : a \mapsto t$, $t \in X$, обозначим через $I(A)$.

Пусть $\rho \subseteq X \times X$ — произвольное отношение на X. Преобразование $f \in \mathcal{T}(X)$ называется *эндоморфизмом* реляционной системы (X, ρ), если для любых $a, b \in X$ из того, что $(a, b) \in \rho$, следует $(af, bf) \in \rho$. Множество всех эндоморфизмов реляционной системы (X, ρ) образует полугруппу относительно обычной композиции преобразований и обозначается $\mathrm{End}(X, \rho)$.

Эндоморфизм $f \in \mathrm{End}(X, \rho)$ называется *полусильным* эндоморфизмом, если для любых $x, y \in X$ из условия $(xf, yf) \in \rho$ следует, что существуют прообразы $x', y' \in X$, т. е. $xf = x'f$, $yf = y'f$, такие, что $(x', y') \in \rho$. Множество всех полусильных эндоморфизмов реляционной системы (X, ρ) обозначается $\mathrm{HEnd}(X, \rho)$.

Эндоморфизм $f \in \mathrm{End}(X, \rho)$ называется *локально сильным* эндоморфизмом, если для любых $x, y \in X$ из условия $(xf, yf) \in \rho$ следует, что для каждого прообраза $x' \in X$ элемента xf существует такой прообраз $y' \in X$ элемента yf, что $(x', y') \in \rho$, и аналогичное утверждение справедливо для каждого прообраза yf. Множество всех локально сильных эндоморфизмов реляционной системы (X, ρ) обозначается $\mathrm{LEnd}(X, \rho)$.

Эндоморфизм $f \in \mathrm{End}(X, \rho)$ называется *квазисильным* эндоморфизмом, если для любых $x, y \in X$ из условия $(xf, yf) \in \rho$ следует, что существует такой прообраз $x' \in X$ элемента xf, что для любого прообраза $y' \in X$ элемента yf выполняется $(x', y') \in \rho$, и аналогичное утверждение справедливо для каждого прообраза yf. Множество всех квазисильных эндоморфизмов реляционной системы (X, ρ) обозначается $\mathrm{QEnd}(X, \rho)$.

Эндоморфизм $f \in \mathrm{End}(X, \rho)$ называется *сильным* эндоморфизмом, если для любых $x, y \in X$ из условия $(xf, yf) \in \rho$ следует, что $(x, y) \in \rho$. Множество всех сильных эндоморфизмов реляционной системы (X, ρ) образует полугруппу относительно обычной композиции преобразований и обозначается $\mathrm{SEnd}(X, \rho)$.

Эндоморфизм $f \in \mathrm{End}(X, \rho)$ называется *автоморфизмом*, если f биективно и f^{-1} — эндоморфизм. Группа всех автоморфизмов реляционной системы (X, ρ) обозначается $\mathrm{Aut}(X, \rho)$. Таким образом, для реляционной системы (X, ρ) имеем цепочку включений

$$\mathrm{End}(X, \rho) \supseteq \mathrm{HEnd}(X, \rho) \supseteq \mathrm{LEnd}(X, \rho) \supseteq \mathrm{QEnd}(X, \rho) \supseteq \mathrm{SEnd}(X, \rho) \supseteq \mathrm{Aut}(X, \rho).$$

Множество всех отношений эквивалентности на X обозначим $\mathrm{Eq}(X)$. Для $\alpha \in \mathrm{Eq}(X)$ через X/α обозначим фактор-множество, а класс эквивалентности α, содержащий элемент $x \in X$, будем обозначать x_α.

Через i_X обозначается диагональное отношение на множестве X, а через w_X — универсальное:

$$i_X = \{(a, a) : a \in X\}, \quad w_X = X \times X.$$

Пусть $\mathcal{G}(X)$ — симметрическая группа на множестве X. Очевидно, что

$$\mathcal{T}(X) = \mathrm{End}(X, i_X) = \mathrm{HEnd}(X, i_X) = \mathrm{LEnd}(X, i_X) \supseteq \\ \supseteq \mathrm{QEnd}(X, i_X) = \mathrm{SEnd}(X, i_X) = \mathrm{Aut}(X, i_X) = \mathcal{G}(X),$$

$$\mathcal{T}(X) = \mathrm{End}(X, w_X) = \mathrm{HEnd}(X, w_X) = \mathrm{LEnd}(X, w_X) = \\ = \mathrm{QEnd}(X, w_X) = \mathrm{SEnd}(X, w_X) \supseteq \mathrm{Aut}(X, w_X) = \mathcal{G}(X).$$

Бинарное отношение называется тривиальным, если оно диагонально или универсально. Известно, что имеет место следующая

Лемма 1 [17]. Преобразование $f \in \mathcal{T}(X)$ является эндоморфизмом отношения $\alpha \in \mathrm{Eq}(X)$ тогда и только тогда, когда для любого $A \in X/\alpha$ существует $B \in X/\alpha$, такое, что $Af \subseteq B$.

2. Полусильные эндоморфизмы

Выясним, при каких условиях произвольный эндоморфизм графа отношения эквивалентности является полусильным.

Лемма 2. Эндоморфизм $f \in \mathrm{End}(X, \alpha)$ отношения $\alpha \in \mathrm{Eq}(X)$ является полусильным тогда и только тогда, когда для любого $B \in X/\alpha$, такого, что $B \cap \mathrm{im}(f) \neq \varnothing$, и любых $a, b \in B \cap \mathrm{im}(f)$ существует $A \in X/\alpha$, такой, что $a, b \in Af$.

Доказательство. Необходимость. Пусть f — полусильный эндоморфизм отношения эквивалентности α, $B \in X/\alpha$, $B \cap \mathrm{im}(f) \neq \varnothing$. Предположим, что $a, b \in B \cap \mathrm{im}(f)$, следовательно, $(a, b) \in \alpha$. Так как $f \in \mathrm{HEnd}(X, \alpha)$, то среди множества прообразов af^{-1}, bf^{-1} найдутся такие элементы a' и b' соответственно, что $(a', b') \in \alpha$, то есть $a', b' \in C$ для некоторого $C \in X/\alpha$. Следовательно, $a'f = a$, $b'f = b \in Cf$.

Достаточность. Пусть $f \in \mathrm{End}(X, \alpha)$ — произвольный эндоморфизм, $(a', b') \in \alpha$ для некоторых $a', b' \in \mathrm{im}(f)$. Тогда $a', b' \in B \cap \mathrm{im}(f)$ для некоторого $B \in X/\alpha$ и по условию леммы в X/α существует такой класс эквивалентности A, что $a', b' \in Af$. Следовательно, существуют прообразы $a \in a'f^{-1}$, $b \in b'f^{-1}$, для которых $(a, b) \in \alpha$: в самом деле, для любой пары (x, y) из $(a'f^{-1} \cap A) \times (b'f^{-1} \cap A)$, очевидно, $(x, y) \in \alpha$. Таким образом, $f \in \mathrm{HEnd}(X, \alpha)$. ■

Следствие 1. Любой эндоморфизм $f \in \mathrm{End}(X, \alpha)$ отношения $\alpha \in \mathrm{Eq}(X)$, область значений которого содержит не более чем по одному представителю из классов X/α, является полусильным.

Множество всех полусильных эндоморфизмов отношения эквивалентности в общем случае не является полугруппой. Действительно, пусть, например, $X = \{1, 2, 3, 4\}$, $\alpha = \{1, 2\}^2 \cup \{3\}^2 \cup \{4\}^2$. Тогда, согласно лемме 2, имеем $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 4 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 \end{pmatrix} \in \mathrm{HEnd}(X, \alpha)$, однако произведение $fg = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 2 \end{pmatrix} \notin \mathrm{HEnd}(X, \alpha)$.

Нетрудно убедиться, что если $|X| \leqslant 2$, то имеют место равенства

$$\mathrm{End}(X, \alpha) = \mathrm{HEnd}(X, \alpha) = \mathrm{LEnd}(X, \alpha).$$

Утверждение 1. Пусть $|X| > 2$, $\alpha \in \mathrm{Eq}(X)$. Множество $\mathrm{HEnd}(X, \alpha)$ всех полусильных эндоморфизмов отношения эквивалентности α является полугруппой тогда и только тогда, когда α — тривиальное отношение эквивалентности.

Доказательство. Пусть $\mathrm{HEnd}(X, \alpha)$ — полугруппа, α — нетривиальное отношение эквивалентности на X. Тогда в фактор-множестве X/α найдётся класс мощности больше 1, обозначим его через A. Пусть $B \in X/\alpha$ — произвольный фиксированный класс, а $\delta : X/\alpha \to X$ — отображение, которое ставит в соответствие каждому классу произвольный фиксированный элемент из этого класса. Обозначим через y элемент

из A, отличный от $A\delta$. Рассмотрим следующие полусильные эндоморфизмы f и g:

$$xf = C\delta, \text{ если } x \in C, \quad C \in X/\alpha,$$
$$xg = \begin{cases} x, & \text{если } x \in A, \\ y, & \text{если } x \in B, \\ x_\alpha\delta & \text{в остальных случаях.} \end{cases}$$

Нетрудно видеть, что $fg = \begin{pmatrix} A & B & C & D & \dots \\ A\delta & y & C\delta & D\delta & \dots \end{pmatrix}$. Таким образом, для класса A и элементов $A\delta, y \in A \cap \mathrm{im}(fg)$ не выполняется лемма 2. Следовательно, $fg \notin \mathrm{HEnd}(X, \alpha)$, что противоречит начальному предположению.

С другой стороны, если α — тривиально, то, как было отмечено в п. 1, $\mathrm{HEnd}(X, \alpha) = \mathcal{T}(X)$. ■

Хорошо известно, что $\mathcal{T}(X)$ регулярна, поэтому справедливо

Следствие 2. Для тривиального отношения $\alpha \in \mathrm{Eq}(X)$ полугруппа $\mathrm{HEnd}(X, \alpha)$ регулярна.

3. Локально сильные эндоморфизмы

Следующая лемма описывает критериальные условия, при которых обычные эндоморфизмы являются локально сильными.

Лемма 3. Эндоморфизм $f \in \mathrm{End}(X, \alpha)$ отношения $\alpha \in \mathrm{Eq}(X)$ является локально сильным эндоморфизмом тогда и только тогда, когда для любых $A, B, C \in X/\alpha$ из того, что $Af \subseteq C$ и $Bf \subseteq C$, следует $Af = Bf$.

Доказательство. Необходимость. Пусть $f \in \mathrm{LEnd}(X, \alpha)$ и выполняются включения $Af \subseteq C$, $Bf \subseteq C$ для некоторых $A, B, C \in X/\alpha$. Предположим, что $Af \neq Bf$, тогда $A \neq B$. Не нарушая общности рассуждений, можем считать, что $y \notin Af$, $y \in Bf$ для некоторого $y \in C$. Для любого $x \in Af$, очевидно, $(x, y) \in \alpha$, но для прообраза $x' \in xf^{-1} \cap A$ не существует такого прообраза y' элемента y, что $(x', y') \in \alpha$, а это противоречит условию $f \in \mathrm{LEnd}(X, \alpha)$.

Достаточность. Пусть $f \in \mathrm{End}(X, \alpha)$ — произвольный эндоморфизм и включения $Af \subseteq C$, $Bf \subseteq C$ для любых $A, B, C \in X/\alpha$ влекут $Af = Bf$. Тогда для любых $x, y \in C \cap \mathrm{im}(f)$ из условия $(x, y) \in \alpha$ следует, что для каждого прообраза $x' \in xf^{-1}$ существует $y' \in yf^{-1} \cap (x')_\alpha$, такой, что $(x', y') \in \alpha$. Аналогичное утверждение справедливо для каждого прообраза yf^{-1}. Таким образом, f — локально сильный эндоморфизм отношения $\alpha \in \mathrm{Eq}(X)$. ■

Множество всех локально сильных эндоморфизмов отношения эквивалентности в общем случае не является полугруппой. Чтобы убедиться в этом, рассмотрим на $X = \{1, 2, 3, 4\}$ эквивалентность $\alpha = \{1, 2\}^2 \cup \{3, 4\}^2$. Эндоморфизмы $f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 4 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 \end{pmatrix}$ удовлетворяют лемме 3, а их произведение $fg = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 2 \end{pmatrix}$ не является локально сильным эндоморфизмом.

Утверждение 2. Множество $\mathrm{LEnd}(X, \alpha)$ всех локально сильных эндоморфизмов отношения эквивалентности α на множестве $X \neq \varnothing$ является полугруппой тогда и только тогда, когда $\alpha = w_A \cup i_{X \setminus A}$ для некоторого $A \subseteq X$.

Доказательство. Необходимость. Пусть $\mathrm{LEnd}(X, \alpha)$ — полугруппа. Если $|X| < 4$, то α такое, как указано в условии данного утверждения. Пусть $|X| \geqslant 4$ и

$\alpha \in \mathrm{Eq}(X)$ такое, что $\alpha \neq w_A \cup i_{X\setminus A}$ для любого $A \subseteq X$. Тогда фактор-множество X/α содержит хотя бы два класса эквивалентности A, B с мощностью $\geqslant 2$. Пусть $a, a' \in A$, $a \neq a'$, $b \in B$. Рассмотрим следующие локально сильные эндоморфизмы:

$$x\varphi = \begin{cases} a, & \text{если } x \in A, \\ b, & \text{если } x \in B, \\ x & \text{в остальных случаях,} \end{cases} \qquad x\psi = \begin{cases} a, & \text{если } x = a \text{ или } x \in B, x \neq b, \\ a', & \text{если } x = b \text{ или } x \in A, x \neq a, \\ x & \text{в остальных случаях.} \end{cases}$$

Для любого $x \in X$ имеем

$$x(\varphi\psi) = \begin{cases} a, & \text{если } x \in A, \\ a', & \text{если } x \in B, \\ x & \text{в остальных случаях.} \end{cases}$$

Следовательно, $\varphi\psi \notin \mathrm{LEnd}(X, \alpha)$, что противоречит исходному предположению.

Достаточность. Пусть $\alpha = w_A \cup i_{X\setminus A}$ для некоторого $A \subseteq X$. Если $|A| \leqslant 1$ или $A = X$, то $\alpha = i_X$ или $\alpha = w_X$ и, следовательно, $\mathrm{LEnd}(X, \alpha) = \mathcal{T}(X)$ — полугруппа. Пусть $|A| \geqslant 2$. Учитывая, что X/α содержит единственный класс эквивалентности неединичной мощности, по лемме 3 все элементы из $\mathrm{LEnd}(X, \alpha)$ представляют собой объединение трёх попарно непересекающихся множеств:

$$\Phi_1 = \{\varphi \in \mathcal{T}(X) : \varphi|_A \in \mathcal{T}(A),\ \varphi|_A \notin I(A),\ \varphi|_{X\setminus A} \in \mathcal{T}(X\setminus A)\},$$
$$\Phi_2 = \{\varphi \in \mathcal{T}(X) : \varphi|_A \in I(A),\ \mathrm{im}(\varphi) \subseteq X\setminus A\},$$
$$\Phi_3 = \bigcup_{a\in A} \Phi^{(a)},\ \Phi^{(a)} = \{\varphi \in \mathcal{T}(X) : \varphi|_A \in I(A),\ \ a \in \mathrm{im}(\varphi),\ \ \mathrm{im}(\varphi) \subseteq (X\setminus A) \cup \{a\}\}.$$

Нетрудно видеть, что $\Phi_1 \cup \Phi_2 \cup \Phi_3$ замкнуто по умножению. Таким образом, множество $\mathrm{LEnd}(X, \alpha)$ образует подполугруппу $\mathcal{T}(X)$. ■

Следствие 3. Для любого $\alpha = w_A \cup i_{X\setminus A}$, $A \subseteq X$ полугруппа $\mathrm{LEnd}(X, \alpha)$ является регулярной.

Доказательство. Если $|A| \leqslant 1$ или $A = X$, то $\mathrm{LEnd}(X, \alpha) = \mathcal{T}(X)$, следовательно, $\mathrm{LEnd}(X, \alpha)$ регулярна. Пусть $\varphi \in \mathrm{LEnd}(X, \alpha)$ — произвольный локально сильный эндоморфизм. Построим такой $\psi \in \mathrm{LEnd}(X, \alpha)$, для которого $\varphi = \varphi\psi\varphi$. Рассмотрим возможные случаи.

Если $\varphi \in \Phi_1$, то определим преобразование ψ множества X следующим образом:

$$x\psi = \begin{cases} y, y \in x\varphi^{-1}, & \text{если } x \in \mathrm{im}(\varphi), \\ x & \text{в остальных случаях.} \end{cases}$$

Нетрудно видеть, что $\psi|_A \in \mathcal{T}(A)$, $\psi|_{X\setminus A} \in \mathcal{T}(X\setminus A)$ и $\varphi = \varphi\psi\varphi$. Поскольку ранг $\varphi|_A$ всегда $\geqslant 2$, имеем $\psi|_A \notin I(A)$. Таким образом, $\psi \in \Phi_1$.

Пусть $\varphi \in \Phi_2$, $b \in X\setminus A$ — произвольный фиксированный элемент. Положим

$$x\psi = \begin{cases} y, y \in x\varphi^{-1}, & \text{если } x \in \mathrm{im}(\varphi), \\ b, & \text{если } x \in A, \\ x & \text{в остальных случаях.} \end{cases}$$

Тогда $\psi \in \Phi_2 \cup \Phi_3$ и $\varphi = \varphi\psi\varphi$.

Пусть $\varphi \in \Phi^{(a)} \subseteq \Phi_3$, $a \in A$. Определим ψ так, что

$$x\psi = \begin{cases} y, y \in x\varphi^{-1}, & \text{если } x \in \mathrm{im}(\varphi), \\ a\psi, & \text{если } x \in A\setminus\{a\}, \\ x & \text{в остальных случаях.} \end{cases}$$

В этом случае $\psi \in \Phi_2 \cup \Phi_3$ и $\varphi = \varphi\psi\varphi$. ■

Отметим, что если X конечно, $|A| \leqslant 1$ или $A = X$, то $|\mathrm{LEnd}(X, \alpha)| = |X|^{|X|}$.

Как известно [19, с. 210], число всех сюръективных отображений $\mathrm{sur}\, n^m$ из m-элементного множества в n-элементное множество равно $n!S(m, n)$, где $S(m, n)$ — число Стирлинга второго рода.

Следствие 4. Пусть X — конечное множество, $\alpha = w_A \cup i_{X\setminus A}$ для некоторого $A \subset X$, $|A| \geqslant 2$. Тогда

$$|\mathrm{LEnd}(X, \alpha)| = (k^k - k)l^l + l^{l+1} + k \sum_{m=1}^{l+1} C_l^{m-1} \mathrm{sur}\, m^{(l+1)},$$

где k — мощность множества A; l — мощность множества $X\setminus A$.

Доказательство. Нетрудно видеть, что $|\Phi_1| = (k^k - k)l^l$, $|\Phi_2| = l^{l+1}$. Так как для произвольного преобразования из $\varphi \in \Phi^{(a)}$ ранга m, $m \leqslant l+1$, справедливо $\varphi|_A \in I(A)$, то $\Phi^{(a)}$ равномощно множеству всех сюръективных преобразований из $(l+1)$-элементного множества в m-элементное. Поскольку элемент a определён и фиксирован, остальные $(m-1)$ элементов можно выбрать C_l^{m-1} способами. Таким образом,

$$|\Phi^{(a)}| = \mathrm{sur}\, 1^{l+1} + C_l^1 \mathrm{sur}\, 2^{l+1} + C_l^2 \mathrm{sur}\, 3^{l+1} + \ldots$$
$$\ldots + C_l^{l-1} \mathrm{sur}\, l^{l+1} + \mathrm{sur}\, (l+1)^{l+1} = \sum_{m=1}^{l+1} C_l^{m-1} \mathrm{sur}\, m^{(l+1)},$$

и, следовательно, $|\Phi_3| = k|\Phi^{(a)}| = k \sum\limits_{m=1}^{l+1} C_l^{m-1} \mathrm{sur}\, m^{(l+1)}$. Поскольку $\Phi_1 \cap \Phi_2 \cap \Phi_3 = \varnothing$, получаем искомую формулу. ■

4. Квазисильные эндоморфизмы

Для квазисильных эндоморфизмов выполняется следующая

Лемма 4. Для всякого отношения $\alpha \in \mathrm{Eq}(X)$ справедливо равенство

$$\mathrm{QEnd}(X, \alpha) = \mathrm{SEnd}(X, \alpha).$$

Доказательство. Достаточно доказать включение $\mathrm{QEnd}(X, \alpha) \subseteq \mathrm{SEnd}(X, \alpha)$. Пусть $f \in \mathrm{QEnd}(X, \alpha)$. Отметим, что так как f — квазисильный, для любого $c \in \mathrm{im}(f)$ в cf^{-1} существует прообраз c', α-эквивалентный любому другому прообразу элемента c. Следовательно, все прообразы любого фиксированного элемента находятся в одном и том же классе эквивалентности, то есть несколько классов не могут одновременно отображаться в c.

Пусть $a, b \in \mathrm{im}(f)$ и $(a, b) \in \alpha$. По определению в af^{-1} существует такой элемент x, что $(x, y) \in \alpha$ для любого $y \in bf^{-1}$. Так как $af^{-1} \subseteq x_\alpha$, последнее равносильно условию: для любых $x \in af^{-1}$ и $y \in bf^{-1}$ выполняется $(x, y) \in \alpha$. Таким образом, f — сильный эндоморфизм. ■

Если X — конечное множество, моноид $\mathrm{SEnd}(X, \alpha)$ регулярен (см., например, [20]). Таким образом, имеет место

Следствие 5. Для любой эквивалентности $\alpha \in \mathrm{Eq}(X)$, где X — конечное множество, полугруппа $\mathrm{QEnd}(X, \alpha)$ является регулярной.

В случае если множество X бесконечное, согласно [21] и лемме 4, $\mathrm{QEnd}(X, \alpha)$ — нерегулярная полугруппа.

ЛИТЕРАТУРА

1. *Глускин Л. М.* Полугруппы изотонных преобразований // Успехи математических наук. 1961. Т. 16. № 5. С. 157–162.
2. *Шнеперман Л. Б.* Полугруппы эндоморфизмов квазиупорядоченных множеств // Учёные записки ЛГПИ им. А. И. Герцена. 1962. Т. 238. С. 21–37.
3. *Важенин Ю. М.* Об элементарной определяемости и элементарной характеризуемости классов рефлексивных графов // Изв. вузов. Математика. 1972. Т. 7. С. 3–11.
4. *Попов Б. В.* Полугруппы эндоморфизмов рефлексивных бинарных отношений // Учёные записки ЛГПИ им. А. И. Герцена. 1967. № 302. С. 116–123.
5. *Araújo J. and Konieczny J.* Dense relations are determined by their endomorphism monoids // Semigroup Forum. 2005. No. 70. P. 302–306.
6. *Кожухов И. Б., Ярошевич В. А.* Полугруппы отображений, сохраняющих бинарное отношение // Фундаментальная и прикладная математика. 2008. Т. 14. № 7. С. 129–135.
7. *Айзенштат А. Я.* Регулярные полугруппы эндоморфизмов упорядоченных множеств // Учёные записки ЛГПИ им. А. И. Герцена. 1968. Т. 387. С. 3–11.
8. *Ким В. И., Кожухов И. Б.* Условия регулярности полугрупп изотонных преобразований счетных цепей // Фундаментальная и прикладная математика. 2006. Т. 12. № 8. С. 97–104.
9. *Айзенштат А. Я.* Определяющие соотношения полугруппы эндоморфизмов конечного линейного упорядоченного множества // Сиб. мат. журн. 1962. Т. 3. № 2. С. 161–169.
10. *Böttcher M. and Knauer U.* Endomorphism spectra of graphs // Discrete Mathematics. 1992. No. 109. P. 45–57.
11. *Böttcher M. and Knauer U.* Postscript: Endomorphism spectra of graphs // Discrete Mathematics. 2003. No. 270. P. 329–331.
12. *Pei H. S. and Dingyu Z.* Green's equivalences on semigroups of transformations preserving order and an equivalence relation // Semigroup Forum. 2005. No. 71. P. 241–251.
13. *Ma M., You T., Luo S., et al.* Regularity and Green's relations for finite E-order-preserving transformations semigroups // Semigroup Forum. 2010. No. 80. P. 164–173.
14. *Deng L., Zeng J., and You T.* Green's relations and regularity for semigroups of transformations that preserve reverse direction equivalence // Semigroup Forum. 2011. No. 83. P. 489–498.
15. *Deng L., Zeng J., and Xu B.* Green's relations and regularity for semigroups of transformations that preserve double direction equivalence // Semigroup Forum. 2010. No. 80. P. 416–425.
16. *Deng L., Zeng J., and You T.* Green's relations and regularity for semigroups of transformations that preserve order and a double direction equivalence // Semigroup Forum. 2012. No. 84. P. 59–68.
17. *Жучок Ю. В.* Ендоморфізми відношень еквівалентності // Вісн. Київ. унів. Сер. Фіз.-мат. науки. 2007. Т. 3. С. 22–26.
18. *Жучок Ю. В.* Полугруппы эндоморфизмов 2-нильпотентных бинарных отношений // Фундаментальная и прикладная математика. 2008. Т. 14. № 6. С. 75–83.
19. *Новиков Ф. А.* Дискретная математика. 2-е изд. Стандарт третьего поколения. СПб.: Питер, 2013. 432 с.
20. *Knauer U. and Nieporte M.* Endomorphisms of graphs I. The monoid of strong endomorphisms // Arch. Math. 1989. V. 52. P. 607–614.
21. *Fan S.* Graphs whose strong endomorphism monoids are regular // Arch. Math. 1999. V. 73. P. 419–421.

2014 Теоретические основы прикладной дискретной математики № 3(25)

DOI 10.17223/20710410/25/2

УДК 519.716.32+519.854

О КЛАССЕ ВАРИАЦИОННО-КООРДИНАТНО-ПОЛИНОМИАЛЬНЫХ ФУНКЦИЙ НАД ПРИМАРНЫМ КОЛЬЦОМ ВЫЧЕТОВ[1]

М. В. Заец

ФГУП НИИ «КВАНТ», г. Москва, Россия

E-mail: mirzaets@hotmail.com

Работа посвящена изучению нового класса функций над примарным кольцом вычетов, который получил название класса функций с вариационно-координатной полиномиальностью. Этот класс обобщает класс полиномиальных функций и наряду с ним обладает тем свойством, что системы уравнений, составленные из таких функций, могут быть решены методом покоординатной линеаризации.

Ключевые слова: *примарное кольцо вычетов, полиномиальные функции, формальные производные, системы уравнений, ВКП-функции.*

Введение

Известно, что системы полиномиальных уравнений над кольцом Галуа — Эйзенштейна (т. е. конечным коммутативным цепным кольцом) могут быть решены методом покоординатной линеаризации [1]. Частным случаем такого кольца является примарное кольцо вычетов $\mathbb{Z}_{p^m}$, $m \in \mathbb{N}$. Суть рассматриваемого метода над $\mathbb{Z}_{p^m}$ заключается в последовательном нахождении p-ичных координат неизвестных переменных, при этом нахождение $(i+1)$-х координат при известных координатах меньшего порядка сводится к решению системы линейных уравнений над полем $\mathrm{GF}(p)$. В работе [2] показано, что класс функций над кольцом вычетов $\mathbb{Z}_{2^m}$, обладающий таким свойством, шире класса полиномиальных при $m \geqslant 3$. Построенный класс назван классом «вариационно-координатно-полиномиальных функций» (ВКП-функций). Данная работа продолжает изучение ВКП-функций и обобщает результаты, полученные ранее в [2, 3], на произвольное кольцо вычетов $\mathbb{Z}_{p^m}$.

1. Свойства полиномиальных функций над $\mathbb{Z}_{p^m}$

Сформулируем и докажем некоторые свойства полиномиальных и треугольных функций над примарным кольцом вычетов, которые необходимы для описания свойств ВКП-функций. Напомним, что функция называется полиномиальной над кольцом вычетов $\mathbb{Z}_k$, $k > 1$, если она представима формулой над классом $\{x_1x_2, x_1+x_2, 1\}$, или, что то же самое, представима некоторым многочленом из $\mathbb{Z}_k[x_1, \ldots, x_n]$. Обозначим класс всех полиномиальных функций от $n \in \mathbb{N}$ переменных над кольцом $\mathbb{Z}_k$ через $\mathcal{P}_k(n)$. Договоримся функции от переменных $x_1, \ldots, x_n$ записывать кратко $f(\mathbf{x})$, класс всех функций от n переменных над кольцом вычетов $\mathbb{Z}_k$ обозначим $\mathcal{F}_k(n)$. При этом равенства $f(\mathbf{x}) = g(\mathbf{x})$ или сравнения вида $f(\mathbf{x}) \equiv g(\mathbf{x}) \pmod{p^j}$ будем понимать соответственно как равенство и сравнение, выполнимые при всех $\mathbf{x}$. Всюду далее считаем, если не оговорено иное, что m, n — произвольные натуральные числа и $m > 1$.

Любой элемент a примарного кольца вычетов $\mathbb{Z}_{p^m}$, где $m \in \mathbb{N}, m > 1$, можно однозначно представить в виде

$$a = a^{(0)} + pa^{(1)} + \ldots + p^{m-1}a^{(m-1)},\ j = 0, \ldots, m-1,$$

[1]Работа выполнена при поддержке гранта Президента РФ (НШ № 6260.2012.10).

где $a^{(j)} \in \mathcal{B} = \{0, \ldots, p-1\} \subset \mathbb{Z}_{p^m}$, называемом разложением элемента a в p-ичном координатном множестве $\mathcal{B}$. Отображения

$$\gamma_j \colon \mathbb{Z}_{p^m} \to \mathcal{B},\ \gamma_j(a) = a^{(j)},\ j = 0, \ldots, m-1,$$

называются координатными функциями в координатном множестве $\mathcal{B}$, а элементы $a^{(j)} = \gamma_j(a) \in \mathcal{B}$ — координатами j-го порядка элемента a в координатном множестве $\mathcal{B}$. В частности, любой вектор $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_{p^m}^n$ однозначно представляется в виде суммы

$$\mathbf{x} = \mathbf{x}^{(0)} + p\mathbf{x}^{(1)} + \ldots + p^{m-1}\mathbf{x}^{(m-1)},$$

где $\mathbf{x}^{(j)} = (x_1^{(j)}, \ldots, x_n^{(j)}) \in \mathcal{B}^n$. Если ввести на $\mathcal{B}$ операции сложения $\oplus$ и умножения $\otimes$ по правилу

$$a \oplus b = \gamma_0(a+b),\ a \otimes b = \gamma_0(a \cdot b),\ a, b \in \mathcal{B},$$

то алгебра $(\mathcal{B}, \oplus, \otimes) \cong \mathbb{Z}_{p^m}/p\mathbb{Z}_{p^m} \cong \mathrm{GF}(p)$ будет являться полем из p элементов.

Определение 1. Для функции $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$ и $j \in \{0, \ldots, m-1\}$ отображение $\gamma_j f \colon \mathbb{Z}_{p^m}^n \to \mathcal{B}$, определяемое по правилу

$$\gamma_j f(\alpha) = \gamma_j(f(\alpha))$$

для всех $\alpha \in \mathbb{Z}_{p^m}^n$, будем называть её *$j$-й координатной функцией* или *j-м координатным отображением.*

Другими словами, если $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$, то она представима в виде суммы

$$f(\mathbf{x}) = \sum_{j=0}^{m-1} p^j \gamma_j f(\mathbf{x}).$$

При этом любую координатную функцию $\gamma_j f$, $j = 0, \ldots, m-1$, можно рассматривать в то же время как функцию $\gamma_j f \colon \mathcal{B}^{nm} \to \mathcal{B}$ от nm переменных над полем $\mathcal{B}$, в роли которых выступают координаты $\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(m-1)}$, при этом в таком случае будем предполагать, что координаты переменных расположены в указанном порядке, т. е. $\gamma_j f = \gamma_j f(\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(m-1)})$. А следовательно, любая такая координатная функция может быть представлена многочленом над полем $\mathcal{B}$ от указанных переменных [4].

Определение 2. Функцию $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$ будем называть *T-функцией*, или *треугольной функцией*, если для любого $j \in \{0, \ldots, m-1\}$ её j-я координатная функция зависит только от координат переменных $\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j)}$, т. е. если $f(\mathbf{x})$ имеет вид

$$f(\mathbf{x}) = \sum_{i=0}^{m-1} p^j \gamma_j f(\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j)}).$$

Примерами треугольных функций над кольцом $\mathbb{Z}_{p^m}$ являются полиномиальные функции. Для объяснения данного факта потребуется ввести еще несколько определений.

Определение 3. Будем говорить, что наборы целых чисел $\alpha = (a_1, \ldots, a_n)$ и $\beta = (b_1, \ldots, b_n)$ *сравнимы по модулю d* (или $\alpha \equiv \beta \pmod{d}$), если $a_i \equiv b_i \pmod{d}$ для всех $i \in \{1, \ldots, n\}$.

Определение 4. Функция $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$ *сохраняет отношение сравнимости по модулю $d \,|\, p^m$*, если на сравнимых по модулю d наборах она принимает сравнимые значения по модулю d.

Обозначим через $\mathcal{D}_{p^m}(n)$ класс всех функций над $\mathbb{Z}_{p^m}$ от n переменных, сохраняющих отношение сравнимости по любому делителю p^m, или, что то же самое, сохраняющих любую конгруэнцию кольца $\mathbb{Z}_{p^m}$. Из простейших свойств сравнений следует, что любая полиномиальная функция $f(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$ сохраняет отношение сравнимости по любому делителю p^m, и поэтому справедливо включение $\mathcal{P}_{p^m}(n) \subseteq \mathcal{D}_{p^m}(n)$.

Следующая теорема устанавливает связь между классом треугольных функций и классом $\mathcal{D}_{p^m}(n)$. Её доказательство несложно получить, используя работу [5].

Теорема 1. Пусть $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$. Равносильны следующие утверждения:

1) $f(\mathbf{x}) \in \mathcal{D}_{p^m}(n)$;
2) $f(\mathbf{x})$ является T-функцией.

Таким образом, классы треугольных функций и функций, сохраняющих отношение сравнимости по любому делителю p^m, совпадают. Отсюда следует, что полиномиальные функции являются треугольными.

Пусть $f(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$. Полиномиальную вектор-функцию $\operatorname{grad} f(\mathbf{x}) = \left(\dfrac{\partial f}{\partial x_1}(\mathbf{x}), \ldots, \dfrac{\partial f}{\partial x_n}(\mathbf{x})\right)$ будем называть градиентом многочлена $f(\mathbf{x})$, где $\dfrac{\partial f}{\partial x_i}(\mathbf{x})$ — формальная частная производная многочлена $f(\mathbf{x})$ по переменной x_i, $i = 1, \ldots, n$. Следующая теорема является основной для дальнейших рассуждений.

Теорема 2 (формула Тейлора [1]). Для любого многочлена $f(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$ и любых $j \in \{1, \ldots, m-1\}$, $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{Z}_{p^m}^n$ справедливо сравнение

$$f(\mathbf{x} + p^j\mathbf{h}) \equiv f(\mathbf{x}) + p^j \operatorname{grad} f(\mathbf{x}) \cdot \mathbf{h} \pmod{p^{j+1}}, \tag{1}$$

где $\operatorname{grad} f(\mathbf{x}) \cdot \mathbf{h} = \sum\limits_{i=1}^{n} \dfrac{\partial f}{\partial x_i}(\mathbf{x}) h_i$.

Теорему 2 можно в некотором смысле уточнить. Пусть $\operatorname{grad} f(\mathbf{x})$ — градиент многочлена $f(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$. Приведём каждую его компоненту (формальную частную производную) по модулю p. Тогда в силу свойств многочленов получим полиномиальную вектор-функцию над полем $\mathcal{B}$ от переменных $\mathbf{x}^{(0)}$:

$$\operatorname{grad} f(\mathbf{x}) \equiv \operatorname{grad} f(\mathbf{x}^{(0)}) \pmod{p}.$$

В дальнейшем будем её обозначать $\operatorname{grad} f(\mathbf{x}) \pmod{p}$. Докажем простое следствие.

Следствие 1. Для любого многочлена $f(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$ и любых $j \in \{1, \ldots, m-1\}$, $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{Z}_{p^m}^n$ справедливо сравнение

$$f(\mathbf{x} + p^j\mathbf{h}) \equiv f(\mathbf{x}) + p^j \operatorname{grad} f(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} \pmod{p^{j+1}}, \tag{2}$$

где $\operatorname{grad} f(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} = \sum\limits_{i=1}^{n} \dfrac{\partial f}{\partial x_i}(\mathbf{x}^{(0)}) h_i^{(0)}$.

Доказательство. Достаточно воспользоваться формулой 1 и тем, что $\dfrac{\partial f}{\partial x_i}(\mathbf{x})$ является также многочленом, а значит, $\dfrac{\partial f}{\partial x_i}(\mathbf{x}) h_i \equiv \dfrac{\partial f}{\partial x_i}(\mathbf{x}^{(0)}) h_i^{(0)} \pmod{p}$, откуда и следует сравнение $p^j \operatorname{grad} f(\mathbf{x}) \cdot \mathbf{h} \equiv p^j \operatorname{grad} f(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} \pmod{p^{j+1}}$. ■

Лемма 1. Если $a = x + p^j y$, где $x, y \in \mathbb{Z}_{p^m}$ и $j \in \{0, \ldots, m-1\}$, то

$$\gamma_j(a) = \gamma_j(x) \oplus \gamma_0(y).$$

Доказательство. Легко видеть, что
$\gamma_j(a) = \gamma_j(x + p^j y) = \gamma_j(x + p^j(\gamma_0(y) + p\gamma_1(y) + \ldots + p^{m-1}\gamma_{m-1}(y))) = \gamma_j(x) \oplus \gamma_0(y)$. ■

Теперь, если применить результаты леммы 1 к следствию 1, получим ещё одно

Следствие 2. Для любого многочлена $f(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$ и любых $j \in \{1, \ldots, m-1\}$, $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{Z}_{p^m}^n$ справедливо сравнение

$$\gamma_j f(\mathbf{x} + p^j \mathbf{h}) \equiv \gamma_j f(\mathbf{x}) + \operatorname{grad} f(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} \pmod{p}. \tag{3}$$

Обозначим через $\theta_i = (\delta_{i,1}, \ldots, \delta_{i,n}) \in \mathcal{B}^n$, $i \in \{1, \ldots, n\}$, вектор, i-я компонента которого равна 1, а остальные равны 0 ($\delta_{i,j}$ — символ Кронекера). Используем сравнение (3) при $\mathbf{h} = \theta_i$:

$$\gamma_j f(\mathbf{x} + p^j \theta_i) \equiv \gamma_j f(x) + \frac{\partial f}{\partial x_i}(\mathbf{x}^{(0)}) \pmod{p}.$$

Отсюда

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^{(0)}) \equiv \gamma_j f(\mathbf{x} + p^j \theta_i) - \gamma_j f(x) \pmod{p}.$$

Следовательно, если $\mathbf{x} = \mathbf{x}^{(0)}$, то

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^{(0)}) \equiv \gamma_j f(\mathbf{x}^{(0)} + p^j \theta_i) - \gamma_j f(\mathbf{x}^{(0)}) \pmod{p}. \tag{4}$$

Как видно, в полученном сравнении левая часть не зависит от j и оно выполняется при всех $j \in \{1, \ldots, m-1\}$. Это доказывает следующее важное утверждение.

Утверждение 1. Для любой полиномиальной функции $f(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$ и любого $i \in \{1, \ldots, n\}$ значение формальной частной производной $\dfrac{\partial g}{\partial x_i}(\mathbf{x}^{(0)}) \pmod{p}$ не зависит от представляющего $f(\mathbf{x})$ многочлена $g(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$ и для любого $j \in \{1, \ldots, m-1\}$ верно сравнение (4).

Теперь, если подставить сравнение (4) в (3), получим

$$\gamma_j f(\mathbf{x} + p^j \mathbf{h}) \equiv \gamma_j f(\mathbf{x}) + \sum_{i=1}^{n} \left(\gamma_j f(\mathbf{x}^{(0)} + p^j \theta_i) - \gamma_j f(\mathbf{x}^{(0)})\right) h_i^{(0)} \pmod{p}.$$

Таким образом, переходя к равенству в поле $\mathcal{B}$, докажем следующую теорему.

Теорема 3. Для любой полиномиальной функции $f(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$ и любых $j \in$ $\in \{1, \ldots, m-1\}$, $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{Z}_{p^m}^n$ справедливо равенство

$$\gamma_j f(\mathbf{x} + p^j \mathbf{h}) = \gamma_j f(\mathbf{x}) \oplus \sum_{i=1}^{n} \left(\gamma_j f(\mathbf{x}^{(0)} + p^j \theta_i) \ominus \gamma_j f(\mathbf{x}^{(0)})\right) \otimes h_i^{(0)},$$

где $\theta_i = (\delta_{i,1}, \ldots, \delta_{i,n})$, $i \in \{1, \ldots, n\}$; $\ominus$ — операция взятия противоположного элемента в аддитивной группе поля $\mathcal{B}$.

Из теоремы 3 вытекает, что для любой полиномиальной функции $f(\mathbf{x})$ значение $\gamma_j f(\mathbf{x} + p^j \mathbf{h})$, $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{Z}_{p^m}^n$, можно вычислить, зная значения $\gamma_j f(\mathbf{x})$ и $\gamma_j f(\mathbf{x}^{(0)} + p^j \theta_i) \ominus \gamma_j f(\mathbf{x}^{(0)})$, $i = 1, \ldots, n$. Это приводит к следующему утверждению.

Утверждение 2. Если функция $f(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$, то её j-я координатная функция $\gamma_j f(\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j)}) : \mathcal{B}^{nj} \to \mathcal{B}$, $j \in \{1, \ldots, m-1\}$, однозначно определяется по значениям на множестве $\{0, \ldots, p^j - 1\}^n$ и значениям $\gamma_j f(\theta + p^j \theta_i)$, $\theta \in \mathcal{B}^n$, $\theta_i = (\delta_{i,1}, \ldots, \delta_{i,n})$, $i \in \{1, \ldots, n\}$.

Доказательство. Пусть $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}_{p^m}^n$. Покажем, как, зная указанные в условии величины, вычислить $\gamma_j f(\alpha)$. Разделим каждое a_k на p^j, $k = 1, \ldots, n$, с остатком и представим вектор α в виде

$$\alpha = \beta + p^j \nu,$$

где $\beta = (b_1, \ldots, b_n) \in \{0, \ldots, p^j - 1\}^n$; $\nu = (v_1, \ldots, v_n) \in \mathbb{Z}_{p^m}^n$. Аналогично разделим каждое полученное v_k на p, $k = 1, \ldots, n$, с остатком и представим ν в виде

$$\nu = \theta + p\nu_1,$$

где $\theta = (h_1, \ldots, h_n) \in \mathcal{B}^n$; $\nu_1 \in \mathbb{Z}_{p^m}^n$.

Имеем

$$\alpha = \beta + p^j \nu = \beta + p^j(\theta + p\nu_1) = \beta + p^j\theta + p^{j+1}\nu_1.$$

В силу теоремы 1

$$\gamma_j f(\alpha) = \gamma_j f(\beta + p^j\theta + p^{j+1}\nu_1) = \gamma_j f(\beta + p^j\theta).$$

Тогда по теореме 3

$$\gamma_j f(\alpha) = \gamma_j f(\beta + p^j\theta) = \gamma_j f(\beta) \oplus \sum_{i=1}^{n} \left(\gamma_j f(\beta^{(0)} + p^j\theta_i) \ominus \gamma_j f(\beta^{(0)})\right) \otimes h_i^{(0)}.$$

При этом по условию утверждения известны значения $\gamma_j f(\beta)$, $\gamma_j f(\beta^{(0)})$ и $\gamma_j f(\beta^{(0)} + p^j\theta_i)$, а значит, используя полученное равенство, находим $\gamma_j f(\alpha)$. ■

Сформулируем утверждение о мощности класса полиномиальных функций над кольцом вычетов $\mathbb{Z}_{p^2}$. Его доказательство нетрудно получить, используя работу [6].

Утверждение 3. Для любого $n \in \mathbb{N}$ справедливо равенство

$$|\mathcal{P}_{p^2}(n)| = p^{p^n(n+2)}.$$

2. Класс ВКП-функций над кольцом вычетов

Введём понятие функций с вариационно-координатной полиномиальностью над кольцом вычетов, а также опишем некоторые их общие свойства. Дадим оценку мощности класса ВКП-функций над примарным кольцом вычетов и докажем утверждения о его соотношении с классом полиномиальных функций.

2.1. Определение класса ВКП-функций и его простейшие свойства

Определение 5. Функцию $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$ назовём *ВКП-функцией*, если для любого $j \in \{0, \ldots, m-1\}$ существует полиномиальная функция $p_j(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$, j-я координатная функция которой совпадает с j-й координатной функцией функции $f(\mathbf{x})$, т. е. выполняется равенство

$$\gamma_j f(\mathbf{x}) = \gamma_j p_j(\mathbf{x}),\ j = 0, \ldots, m-1. \tag{5}$$

В таком случае будем говорить, что $p_j(\mathbf{x})$ является *многочленом j-й координаты функции* $f(\mathbf{x})$ или её *j-м координатным многочленом.*

При этом в условиях определения 5 будем говорить, что функция $f(\mathbf{x})$ обладает свойством вариационно-координатной полиномиальности. Класс всех ВКП-функций от n переменных над $\mathbb{Z}_{p^m}$ обозначим через $\mathcal{CP}_{p^m}(n)$.

Поясним введённое определение. Произвольная функция $f(\mathbf{x}) \in \mathcal{F}_{p^m}(n)$ является вариационно-координатно-полиномиальной, если существуют такие многочлены, или полиномиальные функции $p_0(\mathbf{x}), p_1(\mathbf{x}), \ldots, p_{m-1}(\mathbf{x})$ над кольцом $\mathbb{Z}_{p^m}$, что выполнено равенство

$$f(\alpha) = \sum_{j=0}^{m-1} p^j \gamma_j p_j(\alpha) \tag{6}$$

для всех $\alpha \in \mathbb{Z}_{p^m}^n$. Использование при этом термина «вариационно» подчёркивает тот факт, что данные координатные многочлены могут быть разными для различных координат, т. е. могут меняться от координаты к координате. Если же все координатные многочлены одинаковы, то такая функция полиномиальна, поэтому справедливо включение

$$\mathcal{P}_{p^m}(n) \subseteq \mathcal{CP}_{p^m}(n).$$

Следующая теорема устанавливает, что ВКП-функции, так же, как и полиномиальные функции, сохраняют отношение сравнимости по любому делителю p^m.

Теорема 4. При любом $n \in \mathbb{N}$ все ВКП-функции $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$ сохраняют отношение сравнимости по любому делителю p^m, т. е. справедливо включение

$$\mathcal{CP}_{p^m}(n) \subseteq \mathcal{D}_{p^m}(n).$$

Доказательство. Если $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$, то существуют полиномиальные функции $p_0(\mathbf{x}), p_1(\mathbf{x}), \ldots, p_{m-1}(\mathbf{x})$, что выполнено равенство (6). В соответствии с теоремой 1 справедливо

$$\gamma_j f(\mathbf{x}) = \gamma_j p_j(\mathbf{x}) = \gamma_j p_j(\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j)}),\ j = 0, \ldots, m-1,$$

а значит, $\gamma_j f(\mathbf{x}) = \gamma_j f(\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j)})$ и $f(\mathbf{x})$ является треугольной функцией, поэтому по теореме 1 $f(\mathbf{x}) \in \mathcal{D}_{p^m}(n)$. ■

Следствие 3. Для любой $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$ справедливо сравнение

$$\gamma_0 f(\mathbf{x}) \equiv p_0(\mathbf{x}^{(0)}) \pmod{p}. \tag{7}$$

Используя (7), можно в некотором смысле считать, что многочлен нулевой координаты ВКП-функции является многочленом над полем $\mathcal{B}$ и при этом задаёт функцию от младших координат аргументов:

$$\gamma_0 p_0(\mathbf{x}^{(0)}) \colon \mathcal{B}^n \to \mathcal{B}.$$

Как известно, любая функция над полем $\mathcal{B}$ полиномиальна и при этом однозначно представима многочленом, у которого степени входящих в него переменных изменяются от 0 до $p-1$ [4]. Поэтому можно считать, что многочлен нулевой координаты определяется однозначно. Более того, по значениям самой функции $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$ его легко восстановить из сравнений $f(\alpha) \equiv p_0(a_1^{(0)}, \ldots, a_n^{(0)}) \pmod{p}$, $\alpha \in \mathbb{Z}_{p^m}^n$.

Следующая теорема говорит о строении координатных функций $\gamma_j f$ ВКП-функций, рассматриваемых над полем $\mathcal{B}$; при этом, как и ранее, будем предполагать, что порядок следования их переменных идёт с возрастанием номеров координат.

Теорема 5 (о строении координатных функций). Если $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$, то для любого $j \in \{1,\dots,m-1\}$ существуют полиномиальные функции $g_{ji}\colon \mathcal{B}^n \to \mathcal{B}$, $g_j\colon \mathcal{B}^{jn} \to \mathcal{B}$, $i = 1,\dots,n$, над полем $\mathcal{B}$, такие, что выполнено равенство

$$\gamma_j f(\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j)}) = \sum_{i=1}^{n} g_{ji}(\mathbf{x}^{(0)}) \otimes x_i^{(j)} \oplus g_j(\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j-1)}).$$

Доказательство. Действительно, согласно равенству (3),

$$\gamma_j f(\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j)}) = \gamma_j p_j(\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j)}) = \gamma_j p_j(\mathbf{x}^{(0)} + \dots + p^{j-1}\mathbf{x}^{(j-1)} + p^j\mathbf{x}^{(j)}) \equiv$$
$$\equiv \gamma_j p_j(\mathbf{x}^{(0)} + \dots + p^{j-1}\mathbf{x}^{(j-1)}) + \operatorname{grad} p_j(\mathbf{x}^{(0)}) \cdot \mathbf{x}^{(j)} \pmod p.$$

Функция $\gamma_j p_j(\mathbf{x}^{(0)} + \dots + p^{j-1}\mathbf{x}^{(j-1)})\colon \mathcal{B}^{jn} \to \mathcal{B}$, рассматриваемая как функция над полем $\mathcal{B}$, представима над ним некоторым многочленом g_j от переменных $\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j-1)}$. Аналогично, существуют полиномиальные функции $g_{ji}(\mathbf{x}^{(0)})\colon \mathcal{B}^n \to \mathcal{B}$, $i = 1,\dots,n$, над полем $\mathcal{B}$, такие, что $\dfrac{\partial p_j}{\partial x_i}(\mathbf{x}^{(0)}) \equiv g_{ji}(\mathbf{x}^{(0)}) \pmod p$. Отсюда имеем сравнение

$$\operatorname{grad} p_j(\mathbf{x}^{(0)}) \cdot \mathbf{x}^{(j)} = \sum_{i=1}^{n} \frac{\partial p_j}{\partial x_i}(\mathbf{x}^{(0)}) x_i^{(j)} \equiv \sum_{i=1}^{n} g_{ji}(\mathbf{x}^{(0)}) \otimes x_i^{(j)} \pmod p,$$

которое, при переходе к равенству в поле $\mathcal{B}$, завершает доказательство теоремы. ■

В частном случае, когда функция $f(\mathbf{x})$ полиномиальна, для функций g_{ji} справедливо сравнение

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^{(0)}) \equiv g_{ji}(\mathbf{x}^{(0)}) \pmod p,$$

из которого следует, что g_{ji} не зависят от j. Это доказывает

Следствие 4. Если $f(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$, то существуют полиномиальные функции $g_i\colon \mathcal{B}^n \to \mathcal{B}$, $i = 1,\dots,n$, над полем $\mathcal{B}$ и для любого $j \in \{1,\dots,m-1\}$ существуют полиномиальные функции $g_j\colon \mathcal{B}^{jn} \to \mathcal{B}$ над полем $\mathcal{B}$, такие, что выполнено равенство

$$\gamma_j f(\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j)}) = \sum_{i=1}^{n} g_i(\mathbf{x}^{(0)}) \otimes x_i^{(j)} \oplus g_j(\mathbf{x}^{(0)},\dots,\mathbf{x}^{(j-1)}).$$

Пример 1. Рассмотрим ВКП-функцию $f(x)$ (см. таблицу) над $\mathbb{Z}_8$ с координатными многочленами $p_0(x) = x$, $p_1(x) = 3x^3 + 2$, $p_2(x) = 5x^3 + x + 7$. Её координатные функции над полем $\mathcal{B} = \{0,1\}$ имеют вид

$$\gamma_0 f(x) = x^{(0)},\ \gamma_1 f(x) = x^{(0)} \otimes x^{(1)} \oplus x^{(0)} \oplus 1,\ \gamma_2 f(x) = (x^{(0)} \oplus 1) \otimes x^{(2)} \oplus x^{(1)} \oplus 1.$$

x	0	1	2	3	4	5	6	7
$f(x)$	6	5	2	3	2	5	6	3

Теорема 5 свидетельствует о свойстве ВКП-функций, играющем важную роль при решении систем уравнений

$$\begin{cases} f_1(x_1,\dots,x_n) = y_1, \\ \qquad\vdots \\ f_l(x_1,\dots,x_n) = y_l, \end{cases}$$

где $y_i \in \mathbb{Z}_{p^m}$; $f_i \in \mathcal{CP}_{p^m}(n)$, $i = 1,\dots,l$ (в дальнейшем такие системы будем называть системами ВКП-уравнений). Идея использования данного свойства заключается

в следующем. При нахождении координат неизвестных переменных до j-го порядка включительно $(j+1)$-я координатная функция каждой из ВКП-функций, входящих в систему, становится аффинной относительно неизвестных $(j+1)$-х координат, а значит, можно свести задачу их нахождения к решению некоторой системы линейных уравнений над полем $\mathcal{B}$.

Сформулируем и докажем формулу Тейлора для ВКП-функций, обобщающую формулу (1). Этот результат в определённом смысле оправдывает терминологию «полиномиальность» в названии ВКП-функций.

Теорема 6 (формула Тейлора). Если функция $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$ и $p_0(\mathbf{x}), \ldots, p_{m-1}(\mathbf{x})$ — её координатные многочлены, то для любых $j \in \{1, \ldots, m-1\}$ и $\mathbf{h} \in \mathbb{Z}_{p^m}^n$ справедливо сравнение

$$f(\mathbf{x} + p^j\mathbf{h}) \equiv f(\mathbf{x}) + p^j \operatorname{grad} p_j(\mathbf{x}) \cdot \mathbf{h} \pmod{p^{j+1}}. \tag{8}$$

Доказательство. В соответствии с формулой (3) имеем сравнение

$$\gamma_j f(\mathbf{x} + p^j\mathbf{h}) = \gamma_j p_j(\mathbf{x} + p^j\mathbf{h}) \equiv \gamma_j p_j(\mathbf{x}) + \operatorname{grad} p_j(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} \pmod{p}.$$

В силу сравнимости $\operatorname{grad} p_j(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} \equiv \operatorname{grad} p_j(\mathbf{x}) \cdot \mathbf{h} \pmod{p}$ запишем

$$\gamma_j p_j(\mathbf{x} + p^j\mathbf{h}) \equiv \gamma_j p_j(\mathbf{x}) + \operatorname{grad} p_j(\mathbf{x}) \cdot \mathbf{h} \pmod{p}. \tag{9}$$

Теперь воспользуемся равенством (6) и приведём его по модулю p^{j+1}, в результате получим сравнение

$$f(\mathbf{x} + p^j\mathbf{h}) \equiv \sum_{i=0}^{j} p^i \gamma_i p_i(\mathbf{x} + p^j\mathbf{h}) \pmod{p^{j+1}}. \tag{10}$$

Так как $p_i(\mathbf{x}) \in \mathcal{D}_{p^m}(n)$, $i = 0, \ldots, j-1$, то из свойств функций, сохраняющих отношение сравнимости, очевидно, следует равенство $\gamma_i p_i(\mathbf{x} + p^j\mathbf{h}) = \gamma_i p_i(\mathbf{x})$. Подставив данные равенства и сравнение (9) в (10), получим

$$f(\mathbf{x} + p^j\mathbf{h}) \equiv \sum_{i=0}^{j-1} p^i \gamma_i p_i(\mathbf{x}) + p^j(\gamma_j p_j(\mathbf{x}) + \operatorname{grad} p_j(\mathbf{x}) \cdot \mathbf{h} \pmod{p^{j+1}}.$$

Отсюда $f(\mathbf{x} + p^j\mathbf{h}) \equiv \sum\limits_{i=0}^{j} p^i \gamma_i p_i(\mathbf{x}) + p^j \operatorname{grad} p_j(\mathbf{x}) \cdot \mathbf{h} \pmod{p^{j+1}}$. Осталось заметить, что $\sum\limits_{i=0}^{j} p^i \gamma_i p_i(\mathbf{x}) \equiv f(\mathbf{x}) \pmod{p^{j+1}}$ и, следовательно, справедливо сравнение (8). ■

Следствие 5. Если функция $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$ и $p_0(\mathbf{x}), \ldots, p_{m-1}(\mathbf{x})$ — её координатные многочлены, то для любых $j \in \{1, \ldots, m-1\}$ и $\mathbf{h} \in \mathbb{Z}_{p^m}^n$ справедливо сравнение

$$f(\mathbf{x} + p^j\mathbf{h}) \equiv f(\mathbf{x}) + p^j \operatorname{grad} p_j(\mathbf{x}^{(0)}) \cdot \mathbf{h}^{(0)} \pmod{p^{j+1}}. \tag{11}$$

В заключение данного подраздела дадим определение ВКП-функции над произвольным кольцом вычетов $\mathbb{Z}_k$, $k \in \mathbb{N}$, $k > 1$. При этом при $m = 1$ положим по определению, что над полем вычетов по модулю p класс ВКП-функций равен классу полиномиальных (или, что то же самое, совпадает с классом всех функций $\mathcal{F}_p(n)$). Другими словами,

$$\mathcal{CP}_p(n) = \mathcal{P}_p(n).$$

Определение 6. Функцию $f(\mathbf{x}) \in \mathcal{F}_k(n)$ будем называть *ВКП-функцией* над кольцом вычетов $\mathbb{Z}_k$, где $k = p_1^{m_1} \cdot \ldots \cdot p_t^{m_t}$ — каноническое разложение числа k, если одновременно выполняются следующие два условия:

— $f(\mathbf{x})$ сохраняет отношение сравнимости по модулю $p_i^{m_i}$, $i = 1, \ldots, t$;
— $f_i(\mathbf{x}) \in \mathcal{F}_{p_i^{m_i}}(n)$, где $f_i(\mathbf{x}) \equiv f(\mathbf{x}) \pmod{p_i^{m_i}}$, является ВКП-функцией над примарным кольцом вычетов $\mathbb{Z}_{p_i^{m_i}}$, $i = 1, \ldots, t$.

В частности, как легко видеть, непосредственно из данного определения следует, что класс полиномиальных над $\mathbb{Z}_k$ функций содержится в классе ВКП-функций над этим кольцом. Класс всех ВКП-функций от n переменных над кольцом вычетов $\mathbb{Z}_k$ обозначим через $\mathcal{CP}_k(n)$.

2.2. Оценка числа ВКП-функций от n переменных над $\mathbb{Z}_{p^m}$

Поскольку каждая ВКП-функция однозначно определяется соответствующими координатными отображениями полиномиальных функций, то их количество равно произведению мощностей классов таких отображений. При $j = 0$ количество различных координатных отображений $\gamma_0 p(\mathbf{x})$ (где $p(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$) определяется числом всех функций от n переменных над полем $\mathcal{B}$ (см. формулу (7)) и равно p^{p^n}. Для каждого $j \in \{1, \ldots, m-1\}$ оценим сверху количество различных $\gamma_j p(\mathbf{x})$ (где $p(\mathbf{x}) \in \mathcal{P}_{p^m}(n)$), что даст верхнюю оценку мощности класса $\mathcal{CP}_{p^m}(n)$.

Утверждение 4. Для любого $n \in \mathbb{N}$ верна оценка числа ВКП-функций от n переменных над $\mathbb{Z}_{p^m}$:

$$\log_p |\mathcal{CP}_{p^m}(n)| \leqslant p^n + (m-1)np^n + \frac{p^n(p^{n(m-1)} - 1)}{p^n - 1}. \tag{12}$$

Доказательство. Согласно утверждению 2, j-е координатное отображение, $j \in \{1, \ldots, m-1\}$, любой полиномиальной функции однозначно определяется по значениям на множестве $\{0, \ldots, p^j - 1\}^n$ и векторах вида $\theta + p^j\theta_i$, $\theta \in \mathcal{B}^n$, $\theta_i = (\delta_{i,1}, \ldots, \delta_{i,n})$, $i \in \{1, \ldots, n\}$. Количество таких значений равно в точности $p^{jn} + np^n$. А значит, число всех различных j-х координатных отображений полиномиальных функций от n переменных над $\mathbb{Z}_{p^m}$ не превосходит $p^{p^{jn}+n\cdot p^n}$. Отсюда получим оценку мощности класса $\mathcal{CP}_{p^m}(n)$:

$$|\mathcal{CP}_{p^m}(n)| \leqslant p^{p^n} \prod_{j=1}^{m-1} p^{p^{jn}+np^n} = p^{p^n} p^{\sum\limits_{j=1}^{m-1}(p^{jn}+np^n)} = p^{p^n+(m-1)np^n+\frac{p^n(p^{n(m-1)}-1)}{p^n-1}}.$$

Утверждение доказано. ■

2.3. Соотношение между классами полиномиальных и ВКП-функций

Ранее было отмечено, что класс полиномиальных функций над кольцом вычетов вкладывается в класс ВКП-функций. Ответим на вопрос о строгости этого вложения. Сначала рассмотрим случай примарных колец вычетов.

Теорема 7. Для любого $n \in \mathbb{N}$ классы полиномиальных и ВКП-функций над $\mathbb{Z}_{p^2}$ от n переменных совпадают, т. е.

$$\mathcal{P}_{p^2}(n) = \mathcal{CP}_{p^2}(n).$$

Доказательство. Согласно утверждению 3,

$$|\mathcal{P}_{p^2}(n)| = p^{p^n(n+2)}.$$

С другой стороны, согласно неравенству (12),

$$|\mathcal{CP}_{p^2}(n)| \leqslant p^{p^n+np^n+p^n} = p^{p^n(n+2)}.$$

А значит, $|\mathcal{CP}_{p^2}(n)| \leqslant |\mathcal{P}_{p^2}(n)|$. Но при этом $\mathcal{P}_{p^2}(n) \subseteq \mathcal{CP}_{p^2}(n)$. Следовательно, имеет место равенство $\mathcal{P}_{p^2}(n) = \mathcal{CP}_{p^2}(n)$. ■

Следствие 6. Если $k = p_1^{m_1} \cdot \ldots \cdot p_t^{m_t}$ — каноническое разложение числа $k \in \mathbb{N}$, $k > 1$, и $m_i \in \{1, 2\}$, $i = 1, \ldots, t$, то для любого $n \in \mathbb{N}$ справедливо равенство

$$\mathcal{P}_k(n) = \mathcal{CP}_k(n).$$

Доказательство. Если $f(\mathbf{x}) \in \mathcal{CP}_k(n)$, то функция $f_i(\mathbf{x}) \equiv f(\mathbf{x}) \pmod{p_i^{m_i}}$, $f_i \in \mathcal{F}_{p_i^{m_i}}(n)$, $i = 1, \ldots, t$, в силу теоремы 7 при $m_i = 2$ и равенства $\mathcal{CP}_p(n) = \mathcal{P}_p(n)$, полиномиальна. А значит, и $f(\mathbf{x})$ полиномиальна. ■

Чтобы показать, что при $m \geqslant 3$ существуют не полиномиальные ВКП-функции над $\mathbb{Z}_{p^m}$, докажем следующую теорему о достаточном условии отсутствия полиномиального представления для функции из данного класса.

Теорема 8. Пусть функция $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$, $m \geqslant 3$, и $p_0(\mathbf{x}), \ldots, p_{m-1}(\mathbf{x})$ — её координатные многочлены. Если существуют $i, j \in \{1, \ldots, m-1\}$, $i \neq j$, и $\alpha \in \mathcal{B}^n$, такие, что

$$\operatorname{grad} p_i(\alpha) \not\equiv \operatorname{grad} p_j(\alpha) \pmod{p},$$

то $f(\mathbf{x}) \notin \mathcal{P}_{p^m}(n)$.

Доказательство. Так как $\operatorname{grad} p_i(\alpha) \not\equiv \operatorname{grad} p_j(\alpha) \pmod{p}$, то существует $s \in \{1, \ldots, n\}$, такое, что

$$\frac{\partial p_i}{\partial x_s}(\alpha) \not\equiv \frac{\partial p_j}{\partial x_s}(\alpha) \pmod{p}. \tag{13}$$

Предположим, что функция $f(\mathbf{x})$ полиномиальна и представима некоторым многочленом $g(\mathbf{x}) \in \mathbb{Z}_{p^m}[x_1, \ldots, x_n]$. Тогда в соответствии с утверждением 1 (сравнение (4)) справедливо сравнение

$$\frac{\partial g}{\partial x_s}(\alpha) \equiv \gamma_k f(\alpha + p^k\theta_s) - \gamma_k f(\alpha) \pmod{p},\ k = 1, \ldots, m-1.$$

В частности, при $k = i$ имеем

$$\frac{\partial g}{\partial x_s}(\alpha) \equiv \gamma_i f(\alpha + p^i\theta_s) - \gamma_i f(\alpha) = \gamma_i p_i(\alpha + p^i\theta_s) - \gamma_i p_i(\alpha) \equiv \frac{\partial p_i}{\partial x_s}(\alpha) \pmod{p}.$$

Таким образом, справедливо сравнение $\dfrac{\partial g}{\partial x_s}(\alpha) \equiv \dfrac{\partial p_i}{\partial x_s}(\alpha) \pmod{p}$. Применяя те же рассуждения при $k = j$, получим

$$\frac{\partial g}{\partial x_s}(\alpha) \equiv \frac{\partial p_i}{\partial x_s}(\alpha) \equiv \frac{\partial p_j}{\partial x_s}(\alpha) \pmod{p},$$

что противоречит (13), а значит, и полиномиальности функции $f(\mathbf{x})$. ■

Пример 2. Вернемся к рассмотрению ВКП-функции $f(x)$ над $\mathbb{Z}_8$ из примера 1. Найдём формальные производные многочленов $p_1(x)$ и $p_2(x)$:

$$p_1'(x) = x^2 \equiv x \pmod 2, \quad p_2'(x) = 7x^2 + 1 \equiv x + 1 \pmod 2.$$

Ясно, что при любом $\alpha \in \mathcal{B} = \{0, 1\}$ $p_1'(\alpha) \not\equiv p_2'(\alpha) \pmod 2$, поэтому выполнено условие теоремы 8, а значит, $f(x)$ не является полиномиальной.

Используя теорему 8, можно доказать утверждение о соотношении классов полиномиальных и ВКП-функций над $\mathbb{Z}_{p^m}$ при $m \geqslant 3$.

Утверждение 5. Для любых $n \in \mathbb{N}$ и $m \geqslant 3$ класс ВКП-функций $\mathcal{CP}_{p^m}(n)$ не совпадает с классом полиномиальных $\mathcal{P}_{p^m}(n)$.

Доказательство. Действительно, достаточно рассмотреть ВКП-функцию $f(\mathbf{x}) \in \mathcal{CP}_{p^m}(n)$, у которой координатные многочлены $p_1(\mathbf{x}) = x_1$, $p_2(\mathbf{x}) = 0$, а остальные многочлены произвольны. Тогда ясно, что при любом $\alpha \in \mathcal{B}^n$

$$\operatorname{grad} p_1(\alpha) = (1, 0, \ldots, 0) \not\equiv \operatorname{grad} p_2(\alpha) = (0, \ldots, 0).$$

Поэтому $f(\mathbf{x}) \notin \mathcal{P}_{p^m}(n)$. ■

Следствие 7. Если $k = p_1^{m_1} \cdot \ldots \cdot p_t^{m_t}$ — каноническое разложение числа $k \in \mathbb{N}$, $k > 1$, и $m_j \geqslant 3$ для некоторого $j \in \{1, \ldots, t\}$, то для любого $n \in \mathbb{N}$ справедливо строгое включение

$$\mathcal{P}_k(n) \subsetneq \mathcal{CP}_k(n).$$

Доказательство. Рассмотрим произвольные ВКП-функции $f_i(\mathbf{x}) \in \mathcal{CP}_{p_i^{m_i}}(n)$, $i = 1, \ldots, t$, при этом, так как $m_j \geqslant 3$, можно выбрать ВКП-функцию $f_j(\mathbf{x}) \notin \mathcal{P}_{p_j^{m_j}}(n)$.

Построим ВКП-функцию $f(\mathbf{x})$ над $\mathbb{Z}_k$ следующим образом. Для любого $\alpha \in \mathbb{Z}_k^n$ если выполнена система сравнений

$$\begin{cases} \alpha \equiv \alpha_1 \pmod{p_1^{m_1}}, \\ \quad \vdots \\ \alpha \equiv \alpha_t \pmod{p_t^{m_t}}, \end{cases}$$

где $\alpha_i \in \mathbb{Z}_{p_i^{m_i}}^n$, $i = 1, \ldots, t$, то положим значение $f(\alpha) \in \mathbb{Z}_k$ таким, чтобы была выполнена система сравнений

$$\begin{cases} f(\alpha) \equiv f_1(\alpha_1) \pmod{p_1^{m_1}}, \\ \quad \vdots \\ f(\alpha) \equiv f_t(\alpha_t) \pmod{p_t^{m_t}}. \end{cases}$$

Легко проверить, что определённая таким образом функция $f(\mathbf{x})$ сохраняет отношение сравнимости по модулю $p_i^{m_i}$, $i = 1, \ldots, t$, и $f(\mathrm{x}) \equiv f_i(\mathbf{x}) \pmod p_i^{m_i}$, $i = 1, \ldots, t$. Следовательно, $f(\mathbf{x})$ является ВКП-функцией и при этом $f(\mathbf{x}) \notin \mathcal{P}_k(n)$ (иначе получим противоречие с тем, что $f_j(\mathbf{x}) \notin \mathcal{P}_{p_j^{m_j}}(n)$). ■

3. Метод покоординатной линеаризации для решения систем ВКП-уравнений

Опишем алгоритм решения систем ВКП-уравнений над примарным кольцом вычетов, являющийся обобщением метода покоординатной линеаризации для решения

систем полиномиальных уравнений (см. также [7]). Будем при этом предполагать, что для каждой ВКП-функции $f_i(\mathbf{x})$, $i = 1, \ldots, l$, из системы уравнений

$$\begin{cases} f_1(\mathbf{x}) = y_1, \\ \quad \vdots \\ f_l(\mathbf{x}) = y_l, \end{cases} \tag{14}$$

где $y_i \in \mathbb{Z}_{p^m}$; $f_i \in \mathcal{CP}_{p^m}(n)$, $i = 1, \ldots, l$, известны её координатные многочлены $p_{ij}(\mathbf{x})$, $j = 0, \ldots, m-1$.

1. Приведём систему (14) по модулю p. В силу сравнения (7) получим систему полиномиальных уравнений над полем $(\mathcal{B}, \oplus, \otimes)$ относительно младших координат неизвестных переменных:

$$\begin{cases} \gamma_0 f_1(\mathbf{x}^{(0)}) \equiv y_1^{(0)}, \\ \quad \vdots \\ \gamma_0 f_l(\mathbf{x}^{(0)}) \equiv y_l^{(0)} \end{cases} \pmod p \;\Leftrightarrow\; \begin{cases} p_{10}(\mathbf{x}^{(0)}) \equiv y_1^{(0)}, \\ \quad \vdots \\ p_{l0}(\mathbf{x}^{(0)}) \equiv y_l^{(0)} \end{cases} \pmod p. \tag{15}$$

Полученную систему необходимо решить каким-либо образом и найти все возможные значения $\mathbf{c}^{(0)} = (c_1^{(0)}, \ldots, c_n^{(0)}) \in \mathcal{B}^n$ координат $\mathbf{x}^{(0)}$. Если система несовместна, то алгоритм заканчивает работу и исходная система (14) не имеет решений.

2. Пусть при $j \in \{1, \ldots, m\}$ найдены все значения $\mathbf{c}^{(0)}$, …, $\mathbf{c}^{(j-1)}$ координат $\mathbf{x}^{(0)}$, …,$\mathbf{x}^{(j-1)}$ соответственно, которые в силу того, что f_i являются треугольными функциями, т. е. $f_i \in \mathcal{D}_{p^m}(n)$, $i = 1, \ldots, l$, удовлетворяют системам сравнений вида

$$\begin{cases} f_1(\mathbf{x}) \equiv y_1, \\ \quad \vdots \\ f_l(\mathbf{x}) \equiv y_l \end{cases} \pmod{p^j} \;\Leftrightarrow\; \begin{cases} f_1(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \equiv y_1, \\ \quad \vdots \\ f_l(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \equiv y_l \end{cases} \pmod{p^j}. \tag{16}$$

Если $j = m$, то перейти к п. 3. Иначе при любом таком наборе координат неизвестных $\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}$ выполним следующие действия. Приведём систему (14) по модулю p^{j+1}, в результате получим систему сравнений вида

$$\begin{cases} f_1(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}, \mathbf{x}^{(j)}) \equiv y_1, \\ \quad \vdots \\ f_l(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}, \mathbf{x}^{(j)}) \equiv y_l \end{cases} \pmod{p^{j+1}}. \tag{17}$$

Рассмотрим i-е уравнение полученной системы, $i \in \{1, \ldots, l\}$:

$$f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}, \mathbf{x}^{(j)}) \equiv y_i \pmod{p^{j+1}}.$$

Воспользуемся сравнением (11):

$$\begin{aligned} f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}, \mathbf{x}^{(j)}) = f_i(\mathbf{c}^{(0)} + \ldots + p^{j-1}\mathbf{c}^{(j-1)} + p^j\mathbf{x}^{(j)}) \equiv \\ \equiv f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) + p^j \operatorname{grad} p_{ij}(\mathbf{c}^{(0)}) \cdot \mathbf{x}^{(j)} \equiv y_i \pmod{p^{j+1}}. \end{aligned}$$

В силу (16) справедливо сравнение $f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \equiv y_i \pmod{p^j}$, а значит,

$$f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \equiv y_i^{(0)} + \ldots + p^{j-1}y_i^{(j-1)} + p^j\gamma_j f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \pmod{p^{j+1}}.$$

Отсюда имеем

$$y_i^{(0)} + \ldots + p^{j-1} y_i^{(j-1)} + p^j \gamma_j f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) + p^j \operatorname{grad} p_{ij}(\mathbf{c}^{(0)}) \cdot \mathbf{x}^{(j)} \equiv y_i \pmod{p^{j+1}}.$$

Приходим к равенству в поле $\mathcal{B}$:

$$p^j \gamma_j f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) + p^j \operatorname{grad} p_{ij}(\mathbf{c}^{(0)}) \cdot \mathbf{x}^{(j)} \equiv p^j y_i^{(j)} \pmod{p^{j+1}} \Leftrightarrow$$
$$\Leftrightarrow \operatorname{grad} p_{ij}(\mathbf{c}^{(0)}) \otimes \mathbf{x}^{(j)} =_{\mathcal{B}} y_i^{(j)} \ominus \gamma_j f_i(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}).$$

Применив таким образом указанные преобразования и рассуждения к каждому уравнению системы (17), получим равносильную ей систему линейных уравнений относительно неизвестных координат $\mathbf{x}^{(j)}$ над полем $\mathcal{B}$:

$$\begin{pmatrix} \operatorname{grad} p_{1j}(\mathbf{c}^{(0)}) \\ \vdots \\ \operatorname{grad} p_{lj}(\mathbf{c}^{(0)}) \end{pmatrix} \otimes \mathbf{x}^{(j)} =_{\mathcal{B}} \begin{pmatrix} y_1^{(j)} \ominus \gamma_j f_1(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \\ \vdots \\ y_l^{(j)} \ominus \gamma_j f_l(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j-1)}) \end{pmatrix}. \tag{18}$$

Матрица полученной системы линейных уравнений является матрицей Якоби $\mathcal{J}_{F_j}(\mathbf{c}^{(0)}) \pmod{p}$ (т. е. приведённой по модулю p) полиномиальной вектор-функции $F_j = (p_{1j}, \ldots, p_{lj})$ в точке $\mathbf{c}^{(0)} \in \mathcal{B}^n$. Решая полученную линейную систему над $\mathcal{B}$, найдём все возможные значения координат $\mathbf{x}^{(j)}$ (при заданных координатах $\mathbf{x}^{(0)} = \mathbf{c}^{(0)}$, $\ldots$, $\mathbf{x}^{(j-1)} = \mathbf{c}^{(j-1)}$).

При всех $\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j-1)}$, удовлетворяющих системе (16), необходимо решить систему (18) и найти все возможные значения координат $\mathbf{x}^{(j)}$. Если таких $\mathbf{x}^{(j)}$ нет (т. е. система (18) несовместна при любых $\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(j-1)}$, удовлетворяющих системе (16)), то исходная система (14) несовместна и алгоритм заканчивает работу. Увеличить j на 1 и перейти к п. 2 алгоритма.

3. Если найдены все координаты переменных $\mathbf{x}^{(0)} = \mathbf{c}^{(0)}$, $\ldots$, $\mathbf{x}^{(m-1)} = \mathbf{c}^{(m-1)}$, которые удовлетворяют системе (16) при $j = m$, то решения $\mathbf{c} = (c_1, \ldots, c_n)$ системы вычисляются следующим образом:

$$\mathbf{c} = \sum_{j=0}^{m-1} p^j \mathbf{c}^{(j)},$$

и на этом алгоритм завершает свою работу.

Теорема 9. Алгоритм решения системы ВКП-уравнений (14) находит все решения или доказывает её несовместность.

Доказательство. Вектор $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{Z}_{p^m}^n$ является решением системы (14) тогда и только тогда, когда при любом $j \in \{0, \ldots, m-1\}$ координаты $\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j)}$ данного вектора удовлетворяют системе

$$\begin{cases} f_1(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j)}) \equiv y_1, \\ \qquad \vdots \\ f_l(\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j)}) \equiv y_l \end{cases} \pmod{p^{j+1}}.$$

Но, как видно, данный алгоритм последовательно при каждом $j \in \{0, \ldots, m-1\}$ находит все такие координаты $\mathbf{c}^{(0)}, \ldots, \mathbf{c}^{(j)}$, которые удовлетворяют указанной системе. А значит, находит все решения исходной системы (14) либо доказывает её несовместность. ■

Приведём некоторые сложностные оценки описанного алгоритма.

Утверждение 6. Пусть $\{\mathbf{c}_1^{(0)}, \ldots, \mathbf{c}_t^{(0)}\}$ — все решения системы (15), $t \in \mathbb{N}$, $k_{ij} = \operatorname{rang} \mathcal{J}_{F_j}(\mathbf{c}_i^{(0)})$, $i = 1, \ldots, t$, $F_j = (p_{1j}, \ldots, p_{lj})$ и $l = \mathrm{O}(n)$. Тогда сложность S приведённого алгоритма оценивается сверху величиной

$$S \leqslant T_0 + \mathrm{O}\left(n^3 t \sum_{j=0}^{m-2} p^{j(n-k)}\right),$$

где T_0 — сложность решения системы (15); $k = \min\{k_1, \ldots, k_t\}$, $k_i = \min\{k_{i1}, \ldots, k_{i,m-1}\}$, $i = 1, \ldots, t$.

Доказательство. Пусть $\mathbf{c}_i^{(0)}$, $i \in \{1, \ldots, t\}$, — фиксированное решение системы (15). Если $k_{ij} = \operatorname{rang} \mathcal{J}_{F_j}(\mathbf{c}_i^{(0)})$, то в п. 2 алгоритма система (18) в случае совместности имеет $p^{n-k_{is}}$ решений для $\mathbf{x}^{(s)}$, $s = 1, \ldots, m-1$ (при фиксированных предыдущих значениях координат $\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(s-1)}$). Отсюда в «худшем» случае для нахождения $\mathbf{x}^{(1)}$ потребуется решить одну систему линейных уравнений над полем $\mathcal{B}$ со сложностью $\mathrm{O}(n^3)$, а для нахождения всех возможных значений координат $\mathbf{x}^{(j)}$, $j \in \{2, \ldots, m-1\}$, — решить $p^{n-k_{i1}} \cdot \ldots \cdot p^{n-k_{i,j-1}} = p^{n(j-1)-\sum\limits_{s=1}^{j-1} k_{is}}$ систем линейных уравнений. Тогда сложность нахождения всех возможных значений $\mathbf{x}^{(j)}$, $j \in \{2, \ldots, m-1\}$ (при заданном $\mathbf{c}_i^{(0)}$, $i \in \{1, \ldots, t\}$) составит в «худшем» случае $S_j = \mathrm{O}\left(n^3 p^{n(j-1)-\sum\limits_{s=1}^{j-1} k_{is}}\right)$. Если $k_i = \min\{k_{i1}, \ldots, k_{i,m-1}\}$, то при любом $j \in \{2, \ldots, m-1\}$ полученное S_j можно оценить следующим образом:

$$S_j \leqslant \mathrm{O}(n^3 p^{n(j-1)-\sum\limits_{s=1}^{j-1} k_{is}}) \leqslant \mathrm{O}(n^3 p^{(j-1)(n-k_i)}).$$

При этом данная оценка справедлива и при $j = 1$. Отсюда сложность S решения системы (14) не превосходит величины

$$S = T_0 + \sum_{i=1}^{t} \sum_{j=1}^{m-1} S_j \leqslant T_0 + \sum_{i=1}^{t} \sum_{j=0}^{m-2} \mathrm{O}(n^3 p^{j(n-k_i)}) = T_0 + \mathrm{O}\left(n^3 \sum_{i=1}^{t} \sum_{j=0}^{m-2} p^{j(n-k_i)}\right) \leqslant$$
$$\leqslant T_0 + \mathrm{O}\left(n^3 \sum_{i=1}^{t} \sum_{j=0}^{m-2} p^{j(n-k)}\right) = T_0 + \mathrm{O}\left(n^3 t \sum_{j=0}^{m-2} p^{j(n-k)}\right).$$

Утверждение доказано. ■

Следствие 8. В условиях утверждения 6 справедливо:

1) если $k \neq n$, то сложность S алгоритма оценивается сверху величиной

$$S \leqslant T_0 + \mathrm{O}\left(n^3 t \frac{p^{(m-1)(n-k)} - 1}{p^{n-k} - 1}\right);$$

2) если $k = n$, то сложность алгоритма оценивается сверху величиной

$$S \leqslant T_0 + \mathrm{O}(n^3 t(m-1)).$$

Заметим, что если $k = n$ в условиях утверждения 6, то и все $k_{ij} = n$ для $i = 1, \ldots, t$, $j = 1, \ldots, m-1$, а значит, система (18) на каждом шаге $j \in \{1, \ldots, m-1\}$ (при любом фиксированном решении системы (15)) имеет не более одного решения. И в «худшем» случае алгоритм сводится к t-кратному решению $(m-1)$ системы линейных уравнений над полем $\mathcal{B}$.

Следствие 9. Если система (15) имеет единственное решение $\mathbf{c}^{(0)}$, $l = n$ и $\operatorname{rang} \mathcal{J}_{F_j}(\mathbf{c}^{(0)}) = n$, $j = 1, \ldots, m-1$, то система (14) имеет единственное решение и сложность S его нахождения алгоритмом равна

$$S = T_0 + \mathrm{O}(n^3(m-1)).$$

Используя предложенный алгоритм, можно решать системы ВКП-уравнений над произвольным кольцом вычетов $\mathbb{Z}_k$, $k > 1$. Напомним, что задачу решения систем полиномиальных уравнений над кольцом вычетов $\mathbb{Z}_k$ можно свести к решению систем полиномиальных уравнений над его соответствующими примарными компонентами. Аналогичным образом можно решать и системы ВКП-уравнений над $\mathbb{Z}_k$

$$\begin{cases} f_1(\mathbf{x}) = y_1, \\ \quad \vdots \\ f_l(\mathbf{x}) = y_l, \end{cases}$$

где $y_i \in \mathbb{Z}_k$; $f_i \in \mathcal{CP}_k(n)$, $i = 1, \ldots, l$. Для этого в соответствии с определением 6 нужно перейти к функциям $f_{ij}(\mathbf{x}) \equiv f_i(\mathbf{x}) \pmod{p_j^{m_j}}$, где $k = p_1^{m_1} \cdot \ldots \cdot p_t^{m_t}$, и решить методом покоординатной линеаризации системы ВКП-уравнений

$$\begin{cases} f_{1j}(\mathbf{x}) = y_{1j}, \\ \quad \vdots \\ f_{lj}(\mathbf{x}) = y_{lj} \end{cases}$$

над примарными компонентами $\mathbb{Z}_{p_j^{m_j}}$ ($y_{ij} \equiv y_i \pmod{p_j^{m_j}}$, $j = 1, \ldots, t$, $i = 1, \ldots, l$), после чего найти искомое решение над $\mathbb{Z}_k$.

Заключение

В работе введено и обобщено (по сравнению с [2, 3]) понятие функции с вариационно-координатной полиномиальностью на произвольное кольцо вычетов. Показано, что в общем случае данный класс расширяет класс полиномиальных функций, приведены оценки его мощности. Обобщён метод покоординатной линеаризации для решения систем уравнений, составленных из таких функций.

ЛИТЕРАТУРА

1. *Михайлов Д. А., Нечаев А. А.* Решение системы полиномиальных уравнений над кольцом Галуа — Эйзенштейна с помощью канонической системы образующих полиномиального идеала // Дискретная математика. 2004. Т. 1. № 1. С. 21–51.
2. *Заец М. В., Никонов В. Г., Шишков А. Б.* Класс функций с вариационно-координатной полиномиальностью над кольцом $\mathbb{Z}_{2^m}$ и его обобщение // Математические вопросы криптографии. 2013. Т. 4. № 3. С. 19–45.
3. *Заец М. В., Никонов В. Г., Шишков А. Б.* Функции с вариационно-координатной полиномиальностью и их свойства // Открытое образование. 2012. № 3. С. 57–61.
4. *Глухов М. М., Шишков А. Б.* Математическая логика. Дискретные функции. Теория алгоритмов. М.: Лань, 2012. 400 с.
5. *Anashin V. and Khrennikov A.* Applied Algebraic Dynamics. Berlin, N. Y.: Walter de Gruyter, 2009. 533 p.
6. *Hungerbuhler N. and Specker E.* A generalization of the Smarandache function to several variables // Integers. 2006. V. 6. P. 1–14.

7. *Заец М. В.* Решение систем ВКП-уравнений методом покоординатной линеаризации над примарным кольцом вычетов // Тезисы XLI Междунар. конф., XI Междунар. конф. молодых учёных «Информационные технологии в науке, образовании, телекоммуникации и бизнесе IT+SE13». Вестник Московского университета им. С. Ю. Витте. 2013. Сер. 1 (приложение). С. 155–157.

2014 Теоретические основы прикладной дискретной математики № 3(25)

DOI 10.17223/20710410/25/3

УДК 519.7

ВЕРХНЯЯ ОЦЕНКА ЧИСЛА БЕНТ-ФУНКЦИЙ НА РАССТОЯНИИ 2^k ОТ ПРОИЗВОЛЬНОЙ БЕНТ-ФУНКЦИИ ОТ $2k$ ПЕРЕМЕННЫХ

Н. А. Коломеец

Институт математики им. С. Л. Соболева СО РАН, г. Новосибирск, Россия

E-mail: nkolomeec@gmail.com

Получена точная верхняя оценка числа бент-функций на расстоянии 2^k от произвольной бент-функции от $2k$ переменных. Установлено, что она достигается только для квадратичных бент-функций. Введено понятие полной аффинной расщепляемости булевой функции. Доказано, что полностью аффинно расщепляемыми могут быть только аффинные и квадратичные функции.

Ключевые слова: *булевы функции, бент-функции, квадратичные бент-функции.*

Введение

Рассматриваются метрические свойства класса бент-функций, а именно число бент-функций на минимальном возможном расстоянии от произвольной бент-функции. Бент-функции — булевы функции от чётного числа переменных, наиболее удалённые от множества всех аффинных функций. Они предложены О. Ротхаусом в 1966 г. в работе [1]. Бент-функции имеют приложения в криптографии, теории кодирования, комбинаторике, алгебре, теории символьных последовательностей (см., например, обзор [2]).

В работе [3] доказан критерий расположения двух бент-функций на минимальном возможном расстоянии друг от друга. В [4] построены все бент-функции на минимальном расстоянии от квадратичной бент-функции и подсчитано число таких бент-функций. В [5] получены возможные расстояния между двумя бент-функциями от $2k$ переменных, принадлежащие интервалу от 2^k до $2^{k+1} - 1$ (от минимального до удвоенного минимального). Заметим, что гипотеза о том, что любую булеву функцию степени не больше k можно представить как сумму двух бент-функций от $2k$ переменных, высказанная Н. Н. Токаревой в работе [6], также связана с метрическими свойствами класса бент-функций.

Работа построена следующим образом: в п. 1 вводятся необходимые определения; в п. 2 приводится обзор свойств булевых функций, связанных с аффинностью на подпространстве; в п. 3 вводится понятие полностью аффинно расщепляемой булевой функции. Доказывается, что полностью аффинно расщепляемыми являются только аффинные и квадратичные булевы функции. Отметим, что в работе [7] рассмотрен частный случай полной аффинной расщепляемости, а именно доказано, что только аффинные и квадратичные булевы функции от n переменных являются полностью аффинно расщепляемыми порядка $\lceil n/2 \rceil$. В п. 4 доказывается точная верхняя оценка числа бент-функций на расстоянии 2^k от бент-функции от $2k$ переменных. Данная оценка достигается только для полностью аффинно расщепляемых бент-функций, т. е. только для квадратичных бент-функций.

1. Определения

Введём необходимые определения. Функция вида $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ называется *булевой функцией* от n переменных; $x = (x_1, \ldots, x_n) \in \mathbb{Z}_2^n$ — *двоичным вектором* длины n.

Через $\mathcal{F}_n$ обозначим множество всех булевых функций от n переменных. *Расстоянием* между двумя булевыми функциями $f, g \in \mathcal{F}_n$ называется число векторов из $\mathbb{Z}_2^n$, на которых значения функций различаются.

Для $x, y \in \mathbb{Z}_2^n$ определим $x \oplus y = (x_1 \oplus y_1, \ldots, x_n \oplus y_n)$, где $\oplus$ — сложение по модулю 2. Введём аналог скалярного произведения векторов x и y:

$$\langle x, y \rangle = x_1 y_1 \oplus x_2 y_2 \oplus \ldots \oplus x_n y_n,$$

где $x_i y_i$ — умножение по модулю 2.

Весом $\mathrm{wt}(f)$ булевой функции $f \in \mathcal{F}_n$ называется число векторов из $\mathbb{Z}_2^n$, на которых она принимает значение 1.

Подфункцией $f_{i_1,\ldots,i_k}^{b_1,\ldots,b_k}$ функции f, где $0 \leqslant k \leqslant n$; $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$ и $b_1, \ldots, b_k \in \mathbb{Z}_2$, называется функция из $\mathcal{F}_{n-k}$, полученная из f подстановкой вместо $x_{i_1}, \ldots, x_{i_k}$ констант $b_1, \ldots, b_k$.

Ограничением булевой функции $f \in \mathcal{F}_n$ на множество $S \subseteq \mathbb{Z}_2^n$ называется функция $f|_S : S \to \mathbb{Z}_2$, такая, что $f|_S(x) = f(x)$ для всех $x \in S$.

Булева функция $f \in \mathcal{F}_n$ называется *уравновешенной* (или *сбалансированной*), если $\mathrm{wt}(f) = 2^{n-1}$. Уравновешенность также обобщают на ограничения булевых функций: функция f называется *уравновешенной на множестве* $D \subseteq \mathbb{Z}_2^n$, $|D|$ чётна, если она принимает значение 1 ровно на половине элементов множества D.

Представление булевой функции $f \in \mathcal{F}_n$ в виде

$$f(x_1, \ldots, x_n) = a_0 \oplus \bigoplus_{k=1}^{n} \bigoplus_{1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n} a_{i_1 \ldots i_k} x_{i_1} \ldots x_{i_k}, \text{ где } a_0, a_{i_1 \ldots i_k} \in \mathbb{Z}_2,$$

называется *алгебраической нормальной формой* (*АНФ*) или *полиномом Жегалкина*; $x_{i_1} \ldots x_{i_k}$ — *мономом степени* k; $a_{i_1 \ldots i_k}, a_0$ — *коэффициентами* при мономах. *Степенью* $\deg f$ функции f называется длина монома наибольшей степени с ненулевым коэффициентом (или $-\infty$, если все коэффициенты нулевые). Известно, что любая булева функция может быть представлена в виде АНФ, причём единственным способом.

Производной $D_\alpha f$ функции $f \in \mathcal{F}_n$ *по направлению* $\alpha \in \mathbb{Z}_2^n$ называется функция $f(x) \oplus f(x \oplus \alpha)$. Заметим, что если $\deg f > 0$, то $\deg D_\alpha f < \deg f$ для любого $\alpha \in \mathbb{Z}_2^n$.

Непустое множество $L \subseteq \mathbb{Z}_2^n$ называется *линейным подпространством* $\mathbb{Z}_2^n$, если для любых $a, b \in L$ верно, что $a \oplus b \in L$. Обозначим через $s \oplus D$, где $s \in \mathbb{Z}_2^n$ и $D \subseteq \mathbb{Z}_2^n$, *сдвиг* множества D, а именно $s \oplus D = \{s \oplus x : x \in D\}$. Множество $U \subseteq \mathbb{Z}_2^n$ называется *аффинным подпространством* $\mathbb{Z}_2^n$ (или просто *подпространством*), если оно является сдвигом некоторого линейного подпространства. *Размерностью* аффинного подпространства называется размерность соответствующего линейного подпространства. Размерность обозначается через $\dim U$. Отметим, что линейное подпространство также является аффинным подпространством. Далее в тексте будем часто опускать слово «аффинное», т. е. будем называть аффинное подпространство просто подпространством. Будем говорить, что L является подпространством U, если L и U являются подпространствами $\mathbb{Z}_2^n$ и $L \subseteq U$.

Аффинной булевой функцией от n переменных называется булева функция, степень которой не превосходит 1, или, другими словами, функция вида

$$\ell_{a,c}(x) = \langle a, x \rangle \oplus c \text{ для некоторых } a \in \mathbb{Z}_2^n, c \in \mathbb{Z}_2.$$

Через $\mathcal{A}_n$ обозначается множество всех аффинных булевых функций от n переменных.

Преобразованием Уолша — Адамара булевой функции $f \in \mathcal{F}_n$ называется функция $W_f : \mathbb{Z}_2^n \to \mathbb{Z}$, заданная равенством

$$W_f(y) = \sum_{x \in \mathbb{Z}_2^n} (-1)^{f(x) \oplus \langle x, y \rangle},$$

числа $W_f(y)$ называются *коэффициентами Уолша — Адамара*. Эти коэффициенты однозначно определяют функцию f. Для них справедливо *равенство Парсеваля*:

$$\sum_{y \in \mathbb{Z}_2^n} W_f^2(y) = 2^{2n}.$$

Для произвольной булевой функции $f \in \mathcal{F}_n$, линейного подпространства $L \subseteq \mathbb{Z}_2^n$ и $a, b \in \mathbb{Z}_2^n$ справедлива следующая формула:

$$\sum_{x \in a \oplus L} (-1)^{f(x) \oplus \langle b, x \rangle} = 2^{\dim L - n} (-1)^{\langle a, b \rangle} \sum_{y \in b \oplus L^\perp} W_f(y) (-1)^{\langle a, y \rangle}. \tag{1}$$

Бент-функцией называется булева функция от $2k$ переменных, все коэффициенты Уолша — Адамара которой по модулю равны 2^k. Множество всех бент-функций от $2k$ переменных обозначается через $\mathfrak{B}_{2k}$. Заметим, что для бент-функции $f \in \mathfrak{B}_{2k}$ справедливо

$$\mathrm{wt}(f),\ \mathrm{dist}(f, \ell_{y,c}) \in \{2^{2k-1} \pm 2^{k-1}\} \text{ для любых } y \in \mathbb{Z}_2^{2k}, c \in \mathbb{Z}_2.$$

С бент-функцией f связывают *дуальную* функцию $\tilde{f}$, определяемую равенством

$$(-1)^{\tilde{f}(y)} = \frac{1}{2^k} W_f(y) \text{ для всех } y \in \mathbb{Z}_2^{2k}.$$

Функция $\tilde{f}$ тоже является бент-функцией. Для бент-функции f формула (1) имеет более простой вид:

$$\sum_{x \in a \oplus L} (-1)^{f(x) \oplus \langle b, x \rangle} = 2^{\dim L - k} (-1)^{\langle a, b \rangle} \sum_{y \in b \oplus L^\perp} (-1)^{\tilde{f}(y) \oplus \langle a, y \rangle}, \tag{2}$$

где L — линейное подпространство $\mathbb{Z}_2^{2k}$; $a, b \in \mathbb{Z}_2^{2k}$.

Булевы функции $f, g \in \mathcal{F}_n$ называются *аффинно эквивалентными*, если существует невырожденная двоичная матрица A размера $n \times n$, вектор $b \in \mathbb{Z}_2^n$ и аффинная функция $\ell \in \mathcal{A}_n$, такие, что

$$f(x) = g(xA \oplus b) \oplus \ell(x) \text{ для всех } x \in \mathbb{Z}_2^n.$$

Булева функция называется *квадратичной*, если её степень равна 2. Для квадратичных функций справедлива *теорема Диксона*: любую квадратичную булеву функцию $f \in \mathcal{F}_n$ можно привести преобразованием вида $f(xA)$, где A — невырожденная двоичная матрица размера $n \times n$, к виду

$$x_1 x_2 \oplus x_3 x_4 \oplus \ldots \oplus x_{2t-1} x_{2t} \oplus \ell(x), \text{ где } \ell \in \mathcal{A}_n \text{ и } 1 \leqslant t \leqslant n/2.$$

Таким образом, любая квадратичная булева функция из $\mathcal{F}_n$ аффинно эквивалентна функции $g_t(x_1, \ldots, x_n) = x_1 x_2 \oplus x_3 x_4 \oplus \ldots \oplus x_{2t-1} x_{2t}$ для некоторого t, $1 \leqslant t \leqslant n/2$.

Определение 1. Булева функция $f \in \mathcal{F}_n$ *аффинна на подпространстве* L, если $f|_L = \ell_{a,c}|_L$, где $a \in \mathbb{Z}_2^n, c \in \mathbb{Z}_2$. Далее будем обозначать это как $f|_L(x) = \langle a, x \rangle \oplus c$.

В случае если $f|_L = c$, $c \in \mathbb{Z}_2$, будем говорить, что f *постоянна* на L.

Через Ind_D, $D \subseteq \mathbb{Z}_2^n$, обозначим булеву функцию от n переменных, принимающую значение 1 на всех элементах множества D (и только на них). Для бент-функции $f \in \mathfrak{B}_{2k}$ справедлива следующая конструкция. Пусть L — подпространство $\mathbb{Z}_2^{2k}$ размерности k и f аффинна на L. Тогда

$$f \oplus Ind_L \tag{3}$$

тоже является бент-функцией. Данная конструкция предложена К. Карле в работе [8].

Для $f, g \in \mathfrak{B}_{2k}$, $f \neq g$, справедливо $\operatorname{dist}(f, g) \geqslant 2^k$. В работе [3] доказан критерий расположения двух бент-функций на расстоянии 2^k.

Утверждение 1 [3]. Пусть $f \in \mathfrak{B}_{2k}$. Тогда все бент-функции на расстоянии 2^k от f имеют вид $f \oplus Ind_L$, где L — подпространство размерности k и f аффинна на L.

Более подробную информацию о бент-функциях можно найти в [9, 10].

2. Аффинность булевых функций на подпространстве

Рассмотрим существующие понятия, связанные с аффинностью булевой функции на подпространстве.

Гранью $\mathbb{Z}_2^n$ называется множество вида $\Gamma^{b_1,\ldots,b_k}_{i_1,\ldots,i_k} = \{x \in \mathbb{Z}_2^n : x_{i_1} = b_1, \ldots, x_{i_k} = b_k\}$, где $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$; $b_1, \ldots, b_k \in \mathbb{Z}_2$. Отметим, что грань является подпространством $\mathbb{Z}_2^n$.

Определение 2. Функция $f \in \mathcal{F}_n$ называется *k-аффинной*, если она аффинна на грани размерности $n - k$.

Уровнем аффинности f называется минимальное возможное k, такое, что f является k-аффинной. Эти определения ввели О. А. Логачёв, А. А. Сальников и В. В. Ященко в работе [11]. Аффинные функции обладают нулевым уровнем аффинности (и только они). Уровень аффинности не может превышать $n - 1$. В [12] доказано, что уровнем аффинности $n - 1$ обладают исключительно квадратичные функции, АНФ которых содержит все мономы степени 2. М. Л. Буряков в [13] доказал, что задача нахождения уровня аффиности булевой функции $f \in \mathcal{F}_n$, число мономов в АНФ которой не превосходит cn, где c — некоторая константа, является NP-трудной.

Обобщённым уровнем аффинности f называется минимальное возможное k, такое, что f аффинна на подпространстве размерности $n - k$. О. А. Логачёв в работе [14] доказал, что для почти всех булевых функций от n переменных обобщённый уровень аффинности лежит в интервале $[n - \log_2 n, n - \log_2 n + 1]$.

Определение 3. Функция $f \in \mathcal{F}_n$ называется *k-нормальной* (*k-слабо нормальной*), если она постоянна (аффинна) на некотором подпространстве размерности k.

Определение 4. Функция $f \in \mathcal{F}_n$ называется *нормальной* (*слабо нормальной*), если она $\lceil n/2 \rceil$-нормальна ($\lceil n/2 \rceil$-слабо нормальна).

Через $\lceil a \rceil$ обозначена целая часть сверху числа $a \in \mathbb{R}$.

Понятие нормальности предложено Х. Доббертином в работе [15], а затем обобщено П. Шарпин в [16]. Х. Доббертин ввёл его для функций от чётного числа переменных. Данное определение тесно связано с классом бент-функций. На тот момент вопрос о существовании бент-функций, не являющихся нормальными, оставался открытым. А. Канто, М. Даум, Х. Доббертин и Г. Леандр в [17] нашли примеры бент-функций от 10 переменных, которые не являются нормальными, и бент-функций от 14 переменных, которые не являются слабо нормальными. Авторы предложили также конструкцию, позволяющую по произвольной не нормальной (не слабо нормальной) бент-

функции от $2k$ переменных построить не нормальную (не слабо нормальную) бент-функцию от $2k+2$ переменных.

Определение 5. Булева функция $f \in \mathcal{F}_n$ задана в виде *линейного разветвления*, если существуют $k \in \mathbb{N}$, $1 \leqslant k \leqslant n$, функции $\Phi : \mathbb{Z}_2^{n-k} \to \mathbb{Z}_2^k$ и $\varphi \in \mathcal{F}_{n-k}$, такие, что

$$f(x, y) = \langle x, \Phi(y) \rangle \oplus \varphi(y)$$

для всех $x \in \mathbb{Z}_2^k$, $y \in \mathbb{Z}_2^{n-k}$.

Подробную информацию об этом представлении можно найти в [9]; см. также работу В. В. Ященко [18] о характеризации бент-функций в виде линейного разветвления.

3. Полностью аффинно расщепляемые булевы функции

Определение 6. Функция $f \in \mathcal{F}_n$ является *аффинно расщепляемой* по подпространству L, если функция f аффинна на каждом сдвиге L.

Определение 7. Функция $f \in \mathcal{F}_n$ называется *полностью аффинно расщепляемой* порядка k, $2 \leqslant k \leqslant n$, если она аффинна на некотором подпространстве $\mathbb{Z}_2^n$ размерности k и аффинно расщепляема по всем подпространствам размерности k, на которых она аффинна.

Порядки $k = 0$ и 1 рассматривать не имеет смысла, поскольку тогда бы все булевы функции удовлетворяли определению.

Тривиально доказывается следующее утверждение.

Утверждение 2. Пусть $f, g \in \mathcal{F}_n$ — аффинно эквивалентные булевы функции. Тогда f является полностью аффинно расщепляемой порядка k тогда и только тогда, когда g является полностью аффинно расщепляемой порядка k.

Все аффинные и квадратичные булевы функции обладают следующим свойством.

Утверждение 3. Пусть $f \in \mathcal{F}_n$ — аффинная или квадратичная. Тогда если f аффинна на подпространстве $L \subseteq \mathbb{Z}_2^n$, то f аффинна на каждом сдвиге L.

Доказательство. Пусть $a \in \mathbb{Z}_2^n$. Функция f аффинна на $a \oplus L$ тогда и только тогда, когда $f(x \oplus a)$ аффинна на L. Отметим, что $f(x \oplus a) = f(x) \oplus D_a f(x)$, при этом из условия утверждения следует, что $\deg D_a f \leqslant 1$. Следовательно, f аффинна на L тогда и только тогда, когда f аффинна на $a \oplus L$. ■

Докажем вспомогательные леммы об аффинности функции на подпространстве.

Лемма 1. Пусть $f \in \mathcal{F}_n$ аффинна на подпространстве $L \subseteq \mathbb{Z}_2^n$ ненулевой размерности. Тогда для некоторого подпространства $U \subset L$ и $a \in L$, таких, что $L = U \cup (a \oplus U)$, функция f постоянна и на U, и на $a \oplus U$.

Доказательство. Без ограничения общности можно считать, что L — линейное подпространство $\mathbb{Z}_2^n$. Тогда для любого $w \in \mathbb{Z}_2^n$ решением системы уравнений $\langle w, x \rangle = 0$, $x \in L$, является либо всё множество L, либо его подпространство размерности $\dim L - 1$. Лемма доказана. ■

Лемма 2. Пусть $f \in \mathcal{F}_n$, L — подпространство $\mathbb{Z}_2^n$ и f постоянна на L. Тогда f аффинна на подпространстве $L \cup (a \oplus L)$, $a \in \mathbb{Z}_2^n$, тогда и только тогда, когда f постоянна на $a \oplus L$.

Доказательство. Без ограничения общности можно считать, что L — линейное подпространство $\mathbb{Z}_2^n$.

Необходимость. Если f постоянна на $L \cup (a \oplus L)$, то утверждение очевидно. Пусть $f|_{L \cup (a \oplus L)}(x) = \langle w, x \rangle \oplus c$, $w \in \mathbb{Z}_2^n$, $c \in \mathbb{Z}_2$. Тогда $f|_{a \oplus L}(x) = \langle w, x \rangle \oplus c = f|_L(a \oplus x) \oplus \langle w, a \rangle$.

Достаточность. Пусть $f|_L = c_1$ и $f|_{a\oplus L} = c_2$, $c_1, c_2 \in \mathbb{Z}_2$. Если $c_1 = c_2$, то утверждение очевидно. Пусть $c_1 \neq c_2$, т. е. $c_2 = c_1 \oplus 1$. Рассмотрим $L^\perp$. Для некоторого $w \in L^\perp$ верно, что $\langle w, a\rangle = 1$, поскольку если $\langle w, a\rangle = 0$ для всех $w \in L^\perp$, то $a \in L^{\perp\perp} = L$, но $a \notin L$. Следовательно, $\langle w, x\rangle|_L = 0$ и $\langle w, x\rangle|_{a\oplus L} = 1$, отсюда $f|_{L\cup(a\oplus L)}(x) = \langle w, x\rangle \oplus c_1$. ■

Утверждение 4. Если булева функция является полностью аффинно расщепляемой порядка k, то она также полностью аффинно расщепляема порядка t для всех $2 \leqslant t \leqslant k$.

Для доказательства утверждения 4 достаточно воспользоваться следующей леммой.

Лемма 3. Пусть $f \in \mathcal{F}_n$ является полностью аффинно расщепляемой порядка k. Тогда если f аффинна на некотором линейном подпространстве U размерности $t < k$, то существует линейное подпространство L размерности k, такое, что $U \subseteq L$ и f аффинна на L.

Доказательство. Воспользуемся индукцией по размерности U. База индукции $\dim U = 0$ очевидно следует из условия леммы.

Предположим, что для всех линейных подпространств размерности t, $t \leqslant k-1$, утверждение леммы верно. Докажем, что оно верно и для U размерности $t+1$.

Представим U как $U' \cup (a \oplus U')$, где U' — подпространство U размерности t, $a \in U$. Тогда по предположению индукции существует L размерности k, $U' \subseteq L$ и f аффинна на L. Без ограничения общности можно считать, что $f|_L = 0$, поскольку прибавление аффинной функции не влияет на наличие или отсутствие аффинности. Тогда по лемме 2 верно, что $f|_{a\oplus U'} = c$, где $c \in \mathbb{Z}_2$. Поскольку по условию леммы f аффинна на $a \oplus L$, то по лемме 1 существует подпространство $a \oplus T$ размерности $k-1$, такое, что $a \oplus U' \subseteq a \oplus T \subset a \oplus L$ и $f|_{a\oplus T} = c$. А так как $f|_T = 0$, то по лемме 2 функция f аффинна на линейном подпространстве $T \cup (a \oplus T)$ размерности k, которое содержит U. ■

Далее докажем, что полностью аффинно расщепляемыми порядка 2 могут быть только аффинные и квадратичные булевы функции.

Лемма 4. Пусть $f \in \mathcal{F}_n$, $n > 2$, и $L = \{a, b, c, d\}$ — подпространство $\mathbb{Z}_2^n$ размерности 2. Тогда f аффинна на L тогда и только тогда, когда $f(a) \oplus f(b) \oplus f(c) \oplus f(d) = 0$.

Доказательство леммы очевидно.

Лемма 5. Пусть $f \in \mathcal{F}_n$, $n > 2$. Тогда существует подпространство размерности 2, на котором f аффинна.

Доказательство. Докажем, что f аффинна на некотором подпространстве размерности 2 при $n = 3$. Из этого будет следовать справедливость леммы, поскольку любая булева функция от большего числа переменных имеет подфункцию от трёх переменных.

В алгебраической нормальной форме f могут присутствовать четыре монома степеней 2 и 3: $x_1x_2x_3$, x_1x_2, x_1x_3 и x_2x_3. Рассмотрим два случая.

С л у ч а й 1. Моном $x_1x_2x_3$ не присутствует. Имеем два подслучая.

1) Не все мономы степени 2 присутствуют в АНФ. Тогда, очевидно, у присутствующих мономов есть общая переменная x_i, $1 \leqslant i \leqslant 3$. Следовательно, f_i^0 — аффинная.
2) Мономы x_1x_2, x_1x_3 и x_2x_3 присутствуют в АНФ. Заметим, что $x_1x_2 \oplus x_1x_3 \oplus \oplus x_2x_3 = x_1(x_2 \oplus x_3) \oplus x_2x_3$, поэтому f аффинна на подпространстве $D = \{(x_1, x_2, x_3) : x_2 \oplus x_3 = 1, x_1, x_2, x_3 \in \mathbb{Z}_2\}$ размерности 2.

С л у ч а й 2. Моном $x_1x_2x_3$ присутствует. Также имеем два подслучая.

1) В АНФ нет мономов степени 2. Тогда, очевидно, f_1^0 — аффинная.
2) В АНФ есть моном степени 2; без ограничения общности положим, что там присутствует x_1x_2. Тогда f_3^1 аффинна, поскольку у неё $x_1x_2x_3$ и x_1x_2 сократятся, а мономы x_1x_3 и x_2x_3 содержат x_3.

Лемма доказана. ■

Лемма 5 следует также из того, что уровнем аффинности $n-1$ обладают только квадратичные функции, АНФ которых содержит все мономы степени 2 (М. Л. Буряков, О. А. Логачёв, [12]).

Лемма 6. Пусть $f \in \mathcal{F}_n$ является полностью аффинно расщепляемой порядка 2. Тогда f либо аффинная, либо квадратичная.

Доказательство. Воспользуемся индукцией по числу переменных. Очевидно, что любая булева функция от двух и меньше переменных является либо аффинной, либо квадратичной. Предположим, что если $g \in \mathcal{F}_k$, $k < n$, является полностью аффинно расщепляемой порядка 2, то $\deg g \leqslant 2$. Докажем, что $\deg f \leqslant 2$.

Рассмотрим линейное подпространство

$$L = \{(\mathbf{0}, 0, 0), (\mathbf{0}, 0, 1), (\mathbf{0}, 1, 0), (\mathbf{0}, 1, 1)\} \subseteq \mathbb{Z}_2^n.$$

Тогда все сдвиги L можно представить следующим образом:

$$\{(\mathbf{x}, 0, 0), (\mathbf{x}, 0, 1), (\mathbf{x}, 1, 0), (\mathbf{x}, 1, 1)\},\ \mathbf{x} \in \mathbb{Z}_2^{n-2}.$$

Поскольку f является полностью аффинно расщепляемой порядка 2, то она либо аффинна на всех сдвигах L, либо не аффинна ни на одном из сдвигов. Поэтому по лемме 4 для некоторой константы $c \in \mathbb{Z}_2$ верно

$$\forall \mathbf{x} \in \mathbb{Z}_2^{n-2}\ (f(\mathbf{x}, 0, 0) \oplus f(\mathbf{x}, 0, 1) \oplus f(\mathbf{x}, 1, 0) \oplus f(\mathbf{x}, 1, 1) = c).$$

Разложим f по последним двум переменным:

$$\begin{gathered} f(\mathbf{x}, y, z) = (y \oplus 1)(z \oplus 1) f(\mathbf{x}, 0, 0) \oplus (y \oplus 1) z f(\mathbf{x}, 0, 1) \oplus y(z \oplus 1) f(\mathbf{x}, 1, 0) \oplus yz f(\mathbf{x}, 1, 1) = \\ = (f(\mathbf{x}, 0, 0) \oplus f(\mathbf{x}, 0, 1) \oplus f(\mathbf{x}, 1, 0) \oplus f(\mathbf{x}, 1, 1)) yz \oplus \\ \oplus (f(\mathbf{x}, 0, 0) \oplus f(\mathbf{x}, 1, 0)) y \oplus (f(\mathbf{x}, 0, 0) \oplus f(\mathbf{x}, 0, 1)) z \oplus f(\mathbf{x}, 0, 0). \end{gathered}$$

Пусть $f'(\mathbf{x}, y) = f(\mathbf{x}, y, 0)$ и $f''(\mathbf{x}, y) = f(\mathbf{x}, 0, y)$, т. е. это подфункции f, и $\alpha = (\mathbf{0}, 1) \in \in \mathbb{Z}_2^{n-1}$. Тогда

$$f(\mathbf{x}, y, z) = c \cdot yz \oplus y D_\alpha f'(x) \oplus z D_\alpha f''(x) \oplus f(\mathbf{x}, 0, 0). \tag{4}$$

Пусть h — любая из функций f', f'' или $f(\mathbf{x}, 0, 0)$. Если h от трёх и более переменных, то по лемме 5 она аффинна на некотором подпространстве размерности 2, при этом h — подфункция f, следовательно, она, как и f, является полностью аффинно расщепляемой порядка 2. Таким образом, по предположению индукции $\deg h \leqslant 2$.

Отсюда $\deg f(\mathbf{x}, 0, 0) \leqslant 2$, а $\deg D_\alpha f', \deg D_\alpha f'' \leqslant 1$. Исходя из равенства (4), получаем, что $\deg f \leqslant 2$. ■

Лемма 7. Бент-функция $f \in \mathfrak{B}_{2k}$ не может быть аффинна на подпространстве размерности больше k.

Доказательство. Пусть $f|_L(x) = \langle w, x\rangle \oplus c$, L — подпространство размерности $k+1$. Тогда бент-функция $f'(x) = f(x) \oplus \langle w, x\rangle \oplus c$ равна 0 на L. Так как размерность L больше k, существуют два различных подпространства U и $a \oplus U$, содержащиеся в L. Тогда $g = f' \oplus Ind_U \oplus Ind_{a\oplus U}$ тоже является бент-функцией по конструкции (3), при этом $\mathrm{wt}(g) = \mathrm{wt}(f') + 2^{k+1}$. Приходим к противоречию, поскольку вес бент-функции равен $2^{2k-1} \pm 2^{k-1}$. ■

Данную лемму также можно найти в [8].

Теорема 1. Пусть $f \in \mathcal{F}_n$. Справедливы следующие утверждения.

(i) Функция f является полностью аффинно расщепляемой порядка k, $2 \leqslant k \leqslant \lceil n/2 \rceil$, тогда и только тогда, когда f либо аффинная, либо квадратичная.

(ii) Функция f является полностью аффинно расщепляемой порядка k, $\lceil n/2 \rceil \leqslant k < n$, и не является полностью аффинно расщепляемой порядка $k+1$ тогда и только тогда, когда f аффинно эквивалентна функции

$$g_{n-k}(x_1, \ldots, x_n) = x_1x_2 \oplus x_3x_4 \oplus \ldots \oplus x_{2n-2k-1}x_{2n-2k}.$$

Доказательство. Заметим, что если f является полностью аффинно расщепляемой порядка k, то она либо аффинная, либо квадратичная: это следует из утверждения 4 и леммы 6.

Так как для аффинных и квадратичных булевых функций имеет место утверждение 3, для доказательства полной аффинной расщепляемости функции достаточно доказать существование подпространства соответствующей размерности, на котором функция аффинна.

Если f является аффинной, доказательство теоремы тривиально.

По теореме Диксона любая квадратичная функция аффинно эквивалентна функции $g_t(x_1, \ldots, x_n) = x_1x_2 \oplus x_3x_4 \oplus \ldots \oplus x_{2t-1}x_{2t}$ для некоторого t, $1 \leqslant t \leqslant n/2$. Таким образом, g_t аффинна на грани $x_2 = x_4 = \ldots = x_{2t} = 0$ размерности $n-t$, т. е. пункт (i) доказан. Для доказательства пункта (ii) достаточно воспользоваться тем, что функция $h(x_1, \ldots, x_{2n-2k}) = x_1x_2 \oplus x_3x_4 \oplus \ldots \oplus x_{2n-2k-1}x_{2n-2k}$ от $2n-2k$ переменных является бент-функцией и по лемме 7 не может быть аффинна на подпространстве размерности большей, чем $n-k$: тогда функция g не может быть аффинна на подпространстве размерности большей, чем $n-k+(n-(2n-2k)) = k$. ■

Таким образом, среди бент-функций полностью аффинно расщепляемыми являются только квадратичные бент-функции.

Отметим, что в работе [7] рассматривался частный случай полной аффинной расщепляемости. В ней доказано, что булева функция от n переменных является полностью аффинно расщепляемой порядка $\lceil n/2 \rceil$ тогда и только тогда, когда она либо аффинная, либо квадратичная.

4. Верхняя оценка числа бент-функций на расстоянии 2^k от произвольной бент-функции из $\mathfrak{B}_{2k}$

Докажем точную верхнюю оценку числа бент-функций на расстоянии 2^k от произвольной бент-функции $f \in \mathfrak{B}_{2k}$. Напомним, что число бент-функций на расстоянии 2^k от f равно числу подпространств $\mathbb{Z}_2^{2k}$ размерности k, на которых f аффинна (утверждение 1).

Поскольку любое подпространство размерности k, $k > 0$, можно представить как $L \cup (a \oplus L)$, где L — подпространство $\mathbb{Z}_2^n$ размерности $k-1$, следующее утверждение

даёт условие, при котором можно увеличить на 1 размерность подпространства, на котором булева функция аффинна.

Утверждение 5. Пусть $f \in \mathcal{F}_n$, L — подпространство $\mathbb{Z}_2^n$ и $f|_L(x) = \langle w, x\rangle \oplus c$ для некоторых $w \in \mathbb{Z}_2^n$ и $c \in \mathbb{Z}_2$. Тогда f аффинна на подпространстве $L \cup (a \oplus L)$, $a \in \mathbb{Z}_2^n$, тогда и только тогда, когда $f|_{a\oplus L}(x) = \langle w, x\rangle \oplus c'$ для некоторого $c' \in \mathbb{Z}_2$.

Доказательство. Рассмотрим функцию $f'(x) = f(x) \oplus \langle w, x\rangle \oplus c$. Очевидно, что $f'|_L = 0$. Следовательно, по лемме 2 функция f' аффинна на $L \cup (a \oplus L)$ тогда и только тогда, когда $f'|_{a\oplus L} = c'$ для некоторого $c' \in \mathbb{Z}_2$. ■

Далее оценим число способов, которыми можно, используя утверждение 5, увеличить на 1 размерность подпространства, на котором бент-функция аффинна. Для этого потребуется следующее понятие. Пусть $f \in \mathcal{F}_n$, $S \subseteq \mathbb{Z}_2^n$. *Неполным преобразованием Уолша* функции $f|_S$ называется отображение

$$W_{f_S}(y) = \sum_{x\in S} (-1)^{f(x)\oplus\langle y,x\rangle},\ y \in \mathbb{Z}_2^n.$$

Приведём аналог равенства Парсеваля для неполного преобразования Уолша:

$$\sum_{y\in\mathbb{Z}_2^n} W_{f_S}^2(y) = \sum_{y\in\mathbb{Z}_2^n}\sum_{u\in S}\sum_{v\in S} (-1)^{f(u)\oplus f(v)\oplus\langle u\oplus v,y\rangle} =$$
$$= \sum_{u\in S}\sum_{v\in S}(-1)^{f(u)\oplus f(v)} \sum_{y\in\mathbb{Z}_2^n}(-1)^{\langle u\oplus v,y\rangle} = \sum_{u\in S}(-1)^{f(u)\oplus f(u)}2^n = 2^n|S|.$$

Более подробную информацию о неполном преобразовании Уолша можно найти в монографии О. А. Логачёва, А. А. Сальникова, С. В. Смышляева, В. В. Ященко [9].

Лемма 8. Пусть f — бент-функция от $2k$ переменных, L — линейное подпространство $\mathbb{Z}_2^{2k}$ размерности t, $t \leqslant k$ и $a_1 \oplus L, \ldots, a_n \oplus L$ — различные сдвиги L. Пусть для некоторого $w \in \mathbb{Z}_2^{2k}$ верно, что

$$f|_{a_i\oplus L}(x) = \langle w, x\rangle \oplus c_i,\ c_i \in \mathbb{Z}_2 \ \text{ для всех } \ i = 1, \ldots, n.$$

Тогда $n \leqslant 2^{2k-2t}$. При этом в случае $n = 2^{2k-2t}$ функция $f(x) \oplus \langle w, x\rangle$ уравновешена на каждом $a \oplus L$, где $a \notin (a_1 \oplus L) \cup \ldots \cup (a_n \oplus L)$.

Доказательство. Известно, что для произвольных бент-функции f, линейного подпространства L и $a, w \in \mathbb{Z}_2^{2k}$ справедлива формула (см. (2) при $b = w$)

$$\sum_{x\in a\oplus L}(-1)^{f(x)\oplus\langle w,x\rangle} = 2^{\dim L-k}(-1)^{\langle a,w\rangle}\sum_{y\in w\oplus L^\perp}(-1)^{\tilde{f}(y)\oplus\langle a,y\rangle}. \tag{5}$$

Пусть $S = w \oplus L^\perp$. Рассмотрим неполное преобразование Уолша функции $\tilde{f}|_S$: $W_{\tilde{f}_S}(u) = \sum\limits_{y\in S}(-1)^{\tilde{f}(y)\oplus\langle u,y\rangle}$, $u \in \mathbb{Z}_2^{2k}$. Тогда, согласно равенству (5),

$$W_{\tilde{f}_S}(u) = 2^{k-t}(-1)^{\langle u,w\rangle}\sum_{x\in u\oplus L}(-1)^{f(x)\oplus\langle w,x\rangle}. \tag{6}$$

Пусть $V = (a_1 \oplus L) \cup \ldots \cup (a_n \oplus L)$. Из равенства (6) и условия леммы следует, что для всех $u \in V$ справедливо $|W_{\tilde{f}_S}(u)| = 2^{k-t}2^t = 2^k$. Так как для частичного преобразования Уолша функции $\tilde{f}|_S$ справедлив аналог равенства Парсеваля, а $|S| = 2^{2k-t}$ и $|V| = n2^t$, то

$$\sum_{u\in\mathbb{Z}_2^{2k}} W_{\tilde{f}_S}^2(u) = \sum_{u\in V} W_{\tilde{f}_S}^2(u) + \sum_{u\notin V} W_{\tilde{f}_S}^2(u) = n2^t2^{2k} + \sum_{u\notin V} W_{\tilde{f}_S}^2(u) = 2^{2k}2^{2k-t}.$$

Следовательно, $n \leqslant 2^{2k-2t}$. Если же $n = 2^{2k-2t}$, то $W_{\tilde{f}_S}(u) = 0$ при $u \notin V$. Отсюда по равенству (6) получаем, что $\sum\limits_{x \in u \oplus L} (-1)^{f(x) \oplus \langle w, x \rangle} = 0$ для $u \notin V$. ■

Сформулируем случай $n = 2^{2k-2t}$ из предыдущей леммы отдельно.

Утверждение 6. Пусть бент-функция $f \in \mathfrak{B}_{2k}$ постоянна на 2^{2k-2t} различных сдвигах подпространства $L \subseteq \mathbb{Z}_2^{2k}$ размерности t, $1 \leqslant t \leqslant k$. Тогда на всех других сдвигах L бент-функция f является уравновешенной.

Данный случай является обобщением утверждения, доказанного К. Карле.

Утверждение 7 [8]. Пусть бент-функция $f \in \mathfrak{B}_{2k}$ постоянна на некотором подпространстве L размерности k. Тогда f уравновешена на каждом сдвиге L, отличном от самого L.

Таким образом, утверждение 6 эквивалентно утверждению 7 в случае $t = k$. Используя идею утверждения 7, Х. Доббертин в работе [15] предложил конструкцию, порождающую нормальные бент-функции.

Докажем, что аффинное подпространство, на котором аффинна полностью аффинно расщепляемая бент-функция, можно «достроить» максимальным для бент-функции числом способов.

Лемма 9. Пусть $f \in \mathfrak{B}_{2k}$ и для некоторого линейного подпространства $L \subseteq \mathbb{Z}_2^{2k}$ размерности t, $t \leqslant k$, бент-функция f аффинна на каждом сдвиге L. Тогда $f(x) \oplus \langle w, x \rangle$ является константой ровно на 2^{2k-2t} различных сдвигах L для любого $w \in \mathbb{Z}_2^{2k}$.

Доказательство. Обозначим через S_w множество сдвигов L, на которых $f(x) \oplus \oplus \langle w, x \rangle$ является константой. Заметим, что если $f|_{a \oplus L}(x) = \langle w, x \rangle \oplus c$, то для любого $w' \in w \oplus L^{\perp}$ верно, что $f|_L(x) = \langle w', x \rangle \oplus \langle w \oplus w', a \rangle \oplus c$. Таким образом, $S_w = S_{w \oplus u}$ для $u \in L^{\perp}$. Поскольку f аффинна на каждом сдвиге L, а число различных сдвигов равно 2^{2k-t}, то должно быть справедливо

$$\frac{1}{2^{2k-t}} \sum_{w \in \mathbb{Z}_2^{2k}} |S_w| \geqslant 2^{2k-t},$$

при этом по лемме 8 $|S_w| \leqslant 2^{2k-2t}$. Следовательно, неравенство справедливо, только если $|S_w| = 2^{2k-2t}$ для всех $w \in \mathbb{Z}_2^{2k}$. ■

Докажем основную теорему.

Теорема 2. Пусть f — бент-функция от $2k$ переменных. Тогда число бент-функций на расстоянии 2^k от f не превосходит $2^k(2^1 + 1) \cdot \ldots \cdot (2^k + 1)$. При этом данная оценка достигается, только если f — квадратичная.

Доказательство. Обозначим через h произвольную квадратичную бент-функцию от $2k$ переменных. Определим следующее множество:

$$D^t(f) = \{a \oplus L : L \text{ — линейное подпространство } \mathbb{Z}_2^{2k} \text{ размерности } t, \\ a \in \mathbb{Z}_2^{2k} \text{ и } f \text{ аффинна на } a \oplus L\},\ 0 \leqslant t \leqslant k.$$

По утверждению 1 число бент-функций на расстоянии 2^k от f равно $|D^k(f)|$. Докажем, что $|D^k(f)| \leqslant |D^k(h)|$.

Воспользуемся индукцией по t, $0 \leqslant t \leqslant k$, и покажем, что $|D^t(f)| \leqslant |D^t(h)|$.

База индукции $t = 0$: очевидно, что $|D^0(f)| = |D^0(h)| = 2^{2k}$.

Пусть для $t < k$ верно, что $|D^t(f)| \leqslant |D^t(h)|$. Докажем, что $|D^{t+1}(f)| \leqslant |D^{t+1}(h)|$. Пусть $N_f(L) = \{U \in D^{t+1}(f) : L \subset U\}$, где $L \in D^t(f)$. Отметим, что любое $U \in N_f(L)$ имеет вид $U = L \cup (a \oplus L)$ для некоторого $a \in \mathbb{Z}_2^{2k}$. Тогда

$$|D^{t+1}(f)| = \frac{1}{2(2^{t+1}-1)} \sum_{L \in D^t(f)} |N_f(L)|,$$

поскольку в подпространстве U содержится ровно $2(2^{t+1}-1)$ различных подпространств размерности t. По утверждению 5 и леммам 8 и 9 для любых $L \in D^t(f)$ и $L' \in D^t(h)$ справделиво $|N_f(L)| \leqslant |N_h(L')| = 2^{2k-2t} - 1$. Отсюда $|D^{t+1}(f)| \leqslant |D^{t+1}(h)|$.

Таким образом, $|D^k(f)| \leqslant |D^k(h)|$. Поскольку $|N_h(L')| = 2^{2k-2\dim L'} - 1$, то

$$|D^k(h)| = 2^{2k} \prod_{t=0}^{k-1} \frac{2^{2k-2t}-1}{2(2^{t+1}-1)} = 2^k \prod_{t=1}^{k} \frac{2^{2t}-1}{2^t-1} = 2^k(2^1+1)\cdot\ldots\cdot(2^k+1).$$

Отметим, что значение $|D^k(h)|$ было подсчитано ранее в работе [4].

Докажем, что оценка достигается только на квадратичных бент-функциях. Пусть f не является квадратичной (из этого автоматически следует, что $k > 2$). Тогда по теореме 1 она не является полностью аффинно расщепляемой порядка k, т. е. f аффинна на подпространстве L размерности k и не аффинна на некотором его сдвиге (если f не аффинна ни на одном подпространстве размерности k, то $|D^k(f)| = 0$).

Без ограничения общности можем полагать, что L — линейное подпространство и $f|_L = 0$ (этого можно добиться за счёт преобразований вида $f(x \oplus a) \oplus \langle w, x\rangle \oplus c$). Из утверждения 6 следует, что на всех сдвигах L, отличных от L, функция f уравновешена.

Пусть L' — линейное подпространство L размерности $k-1$. Очевидно, что $f|_{L'} = 0$. Пусть $N_f(L') > 1$, т. е. функция f аффинна на $L' \cup (a \oplus L')$ для некоторого $a \notin L$. Тогда из леммы 2 следует, что $f|_{a \oplus L'} = c$ для некоторого $c \in \mathbb{Z}_2$. Но в силу уравновешенности f на $a \oplus L$ получаем, что $f|_{(a\oplus L)\setminus(a\oplus L')} = c \oplus 1$, и по лемме 2 функция f аффинна на $a \oplus L$.

Заметим, что если L' и L'' — различные линейные подпространства L размерности $k-1$, то f не может быть аффинна одновременно на $L' \cup (a \oplus L')$ и на $L'' \cup (a \oplus L'')$ в силу уравновешенности f на $a \oplus L$. Число различных L' равно $2^k - 1$. Число различных сдвигов L, не равных L, тоже равно $2^k - 1$. Поэтому если $N_f(L') > 1$ для всех L', то f аффинна на всех сдвигах L. Следовательно, $N_f(L') = 1$ для какого-то L', в то время как $N_h(U) = 3$ для любого $U \in D^{k-1}(h)$. ■

Заключение

Рассмотрим тривиальную верхнюю оценку числа бент-функций на расстоянии 2^k от произвольной бент-функции из $\mathfrak{B}_{2k}$.

Утверждение 8. Пусть $f \in \mathfrak{B}_{2k}$. Тогда число бент-функций на расстоянии 2^k от f не больше чем

$$2^k \frac{(2^{2k}-1)\cdot\ldots\cdot(2^{k+1}-1)}{(2^k-1)\cdot\ldots\cdot(2^1-1)}.$$

Это число аффинных подпространств $\mathbb{Z}_2^{2k}$ размерности k. Его можно оценить как

$$2^{k^2+k} < 2^k \frac{(2^{2k}-1)\cdot\ldots\cdot(2^{k+1}-1)}{(2^k-1)\cdot\ldots\cdot(2^1-1)} < 2^{k^2+2k}.$$

Таким образом, доказанная верхняя оценка близка к квадратному корню из тривиальной оценки.

ЛИТЕРАТУРА

1. *Rothaus O.* On bent functions // J. Combin. Theory. Ser. A. 1976. V. 20. No. 3. P. 300–305.
2. *Токарева Н. Н.* Бент-функции: результаты и приложения. Обзор работ // Прикладная дискретная математика. 2009. № 1. С. 15–37.
3. *Коломеец Н. А., Павлов А. В.* Свойство бент-функций, находящихся на минимальном расстоянии друг от друга // Прикладная дискретная математика. 2009. № 4. С. 5–20.
4. *Коломеец Н. А.* Перечисление бент-функций на минимальном расстоянии от квадратичной бент-функции // Дискретный анализ и исследование операций. 2012. Т. 19. № 1. С. 41–58.
5. *Потапов В. Н.* Спектр мощностей компонент корреляционно-иммунных функций, бент-функций, совершенных раскрасок и кодов // Проблемы передачи информации. 2012. Т. 48. № 1. С. 54–63.
6. *Tokareva N.* On the number of bent functions from iterative constructions: lower bounds and hypothesis // Adv. Math. Commun. 2011. V. 5. No. 4. P. 609–621.
7. *Коломеец Н. А.* Пороговое свойство квадратичных булевых функций // Дискретный анализ и исследование операций. 2014. Т. 21. № 2. С. 52–58.
8. *Carlet C.* Two new classes of bent functions // EUROCRYPT'93. LNCS. 1994. V. 765. P. 77–101.
9. *Логачёв О. А., Сальников А. А., Смышляев С. В., Ященко В. В.* Булевы функции в теории кодирования и криптологии. 2-е изд. М.: МЦНМО, 2012.
10. *Токарева Н. Н.* Нелинейные булевы функции: бент-функции и их обобщения. Saarbrucken: LAP LAMBERT Academic Publishing, 2011.
11. *Логачёв О. А., Сальников А. А., Ященко В. В.* Комбинирующие k-аффинные функции // Труды конф. «Математика и безопасность информационных технологий», Москва, 23–24 октября 2003 г. М.: МЦНМО, 2004. С. 176–178.
12. *Буряков М. Л., Логачёв О. А.* Об уровне аффинности булевых функций // Дискретная математика. 2005. Т. 17. № 4. С. 98–107.
13. *Буряков М. Л. О связи уровня аффинности с криптографическими параметрами булевых функций* // Дискретная математика. 2008. Т. 20. № 2. С. 3–14.
14. *Логачёв О. А.* О значениях уровня аффинности для почти всех булевых функций // Прикладная дискретная математика. 2010. № 3. С. 17–21.
15. *Dobbertin H.* Construction of bent functions and balanced Boolean functions with high nonlinearity // Fast Software Encryption Int. Workshop (Leuven, Belgium, December 14–16, 1994). LNCS. 1994. V. 1008. P. 61–74.
16. *Charpin P.* Normal Boolean functions // J. Complexity. 2004. V. 20. P. 245–265.
17. *Canteaut A., Daum M., Dobbertin H., and Leander G.* Finding nonnormal bent functions // Discrete Appl. Math. 2006. V. 154. No. 2. P. 202–218.
18. *Ященко В. В.* О критерии распространения для булевых функций и о бент-функциях // Проблемы передачи информации. 1997. Т. 33. № 1. С. 75–86.

2014 Теоретические основы прикладной дискретной математики № 3(25)

DOI 10.17223/20710410/25/4

УДК 621.391: 519.728

О ПОНЯТИИ РАВНОСИЛЬНОСТИ НЕДООПРЕДЕЛЁННЫХ АЛФАВИТОВ[1]

Л. А. Шоломов

Институт системного анализа РАН, г. Москва, Россия

E-mail: sholomov@isa.ru

Для недоопределённых алфавитов предложена формализация понятий: а) один алфавит сильнее другого и б) алфавиты равносильны. Рассмотрены несколько подходов к определению этих понятий. Функциональный подход основан на выразимости одного алфавита через другой, три остальных подхода — комбинаторный, вероятностный и алгоритмический — терминологически связаны с подходами Колмогорова к введению меры информации. Доказано, что эти подходы к сравнению алфавитов эквивалентны. Если алфавиты равносильны, то решение задачи оптимального сжатия для одного алфавита фактически обеспечивает решение этой задачи и для второго. Установлено, что соотношения (а) и (б) допускают проверку за полиномиальное время.

Ключевые слова: *недоопределённый алфавит, равносильные алфавиты, энтропия недоопределённых данных, сложность по Колмогорову.*

Введение

Работа имеет дело с недоопределёнными данными — последовательностями недоопределённых символов. Каждому недоопределённому символу соответствует некоторое множество основных (полностью определённых) символов, любым из которых он может быть замещен (доопределён). При оперировании с недоопределёнными данными часто бывает достаточно вместо самих данных иметь их доопределения. Такие более слабые требования к данным предоставляют дополнительные возможности, одной из которых является рассматриваемая в работе возможность нетривиальных равносильных преобразований недоопределённых алфавитов.

Обсуждаются вопросы, каким образом можно сравнивать недоопределённые алфавиты и заключать, что один из них сильнее другого либо что они равносильны. Представлены несколько подходов к введению соответствующих понятий. Первый — функциональный — основан на функциональной выразимости символов одного алфавита через символы другого. Следующие три подхода терминологически связаны с подходами к введению меры информации, описанными А. Н. Колмогоровым [1]. Это комбинаторный, вероятностный (статистический) и алгоритмический подходы. В работе доказано, что все эти подходы эквивалентны, т. е. приводят к одним и тем же соотношениям недоопределённых алфавитов по силе. Установлено, что эти соотношения допускают проверку за полиномиальное время. Равносильные преобразования недоопределённых данных изучались и раньше [2, 3], но речь шла не об алфавитах, а об источниках, порождающих недоопределённые символы с некоторыми вероятностями,

[1]Работа выполнена при поддержке ОНИТ РАН по проекту «Теоретические основы эффективного использования недоопределённой информации» программы «Интеллектуальные информационные технологии, системный анализ и автоматизация».

и в основу был положен информационный подход. Вытекающее из [2, 3] понятие равносильности по существу совпадает с введённым в данной работе.

Переход к равносильному алфавиту может оказаться полезным для ряда задач, имеющих дело с недоопределёнными данными. Одной из них является задача сжатия. В отличие от постановки этой задачи для полностью определённых данных, где кодирование должно обеспечить их полное восстановление [4], в случае недоопределённых данных требуется восстановить лишь некоторое доопределение. Если алфавиты равносильны, то решение задачи оптимального сжатия в одном из них обеспечивает решение аналогичной задачи и для другого алфавита. При этом переход к равносильному алфавиту иногда может облегчить решение исходной задачи сжатия. Ещё одна задача связана с двоичным представлением недоопределённых алфавитов. При двоичном представлении основным символам соответствуют двоичные слова некоторой длины s, а недоопределённым символам — недоопределённые двоичные слова длины s [3]. Возможна ситуация, когда для исходного недоопределённого алфавита такое представление не существует, но оно появляется при переходе к некоторому равносильному алфавиту. Рассмотрение задачи с точностью до равносильности позволяет также уменьшать длину s представлений. Задача наилучшей аппроксимации [3] недоопределённых алфавитов двоичными представлениями может быть поставлена лишь с точностью до равносильности.

Для анализа двух алфавитов на равносильность необходимо знать, как связаны символы одного алфавита с символами другого. Поэтому помимо самих алфавитов задаётся соответствие между их символами. Оно не предполагается взаимно однозначным, поскольку рассматриваются преобразования алфавитов, при которых несколько символов могут отображаться в один, и обратные преобразования, приводящие к возникновению нескольких образов одного символа. Отметим, что соотношение равносильности недоопределённых алфавитов, обладая рядом свойств отношения эквивалентности, не является эквивалентностью на множестве недоопределённых алфавитов, поскольку зависит также от введённого соответствия. В заключение эта зависимость устраняется и рассматривается отношение эквивалентности недоопределённых алфавитов.

1. Недоопределённые алфавиты

Задан конечный алфавит $A_0 = \{a_i : i \in M\}$ *основных* символов. Каждому непустому $T \subseteq M$ сопоставлен символ a_T, называемый *недоопределённым. Доопределением символа* a_T считается всякий основной символ a_i, $i \in T$. Символ a_M, доопределимый любым основным символом, называется *неопределённым* и обозначается $*$. Выделена система $\mathcal{T} \subseteq 2^M$ некоторых непустых подмножеств T множества M и ей соответствует *недоопределённый алфавит* $A = \{a_T : T \in \mathcal{T}\}$. Считаем, что для любого $i \in M$ найдётся $T \in \mathcal{T}$, для которого $i \in T$ (иначе i можно удалить из M). Символы a_T будем понимать также как множества доопределений $a_T = \{a_i : i \in T\}$ и применительно к ним использовать теоретико-множественные операции и отношения. Скажем, что символ a_T *чётче* символа $a_{T'}$ ($a_{T'}$ *размытее* a_T), если $a_T \subseteq a_{T'}$. Под *доопределением последовательности* $\mathbf{a} = a_{T_1} \dots a_{T_n}$ недоопределённых символов понимается любая последовательность основных символов, полученная из исходной заменой каждого символа каким-либо его доопределением, а под *частичным доопределением* (*размытием*) *последовательности* $\mathbf{a}$ — результат замены её символов более чёткими (размытыми) символами.

Пусть наряду с алфавитом A задан недоопределённый алфавит B, для которого основным алфавитом является $B_0 = \{b_j : j \in L\}$, а недоопределённые символы b_U соответствуют множествам U некоторой системы $\mathcal{U} \subseteq 2^L$. Считаем также, что задано соответствие $R_{AB} \subseteq A \times B$, область определения которого совпадает с A, область значений с B. В остальном соответствие произвольно, т. е. символы алфавита A могут иметь несколько образов, символы алфавита B — несколько прообразов. Наряду с записью $(a_T, b_U) \in R_{AB}$ будем использовать $a_T R_{AB} b_U$. Назовём

— алфавиты A и B с заданным для них соответствием R_{AB} *соответственными алфавитами*;
— символы a_T и b_U, такие, что $a_T R_{AB} b_U$, *соответственными символами*;
— последовательности $\mathbf{a} = a_{T_1} \dots a_{T_n}$ и $\mathbf{b} = b_{U_1} \dots b_{U_n}$, для которых $(\mathbf{a}, \mathbf{b}) \in R_{AB}^n$ (т. е. $a_{T_i} R_{AB} b_{U_i}$, $i = 1, \dots, n$), *соответственными последовательностями*.

Операции над соответствиями выполняются обычным образом, а именно: под *инверсией* соответствия R_{AB} понимается

$$R_{BA} = R_{AB}^{-1} = \{(b_U, a_T) : (a_T, b_U) \in R_{AB}\},$$

и если помимо A и B задан недоопределённый алфавит $C = \{c_V : V \in \mathcal{V}\}$, связанный с B соответствием R_{BC}, то *произведением* (*композицией*) соответствий R_{AB} и R_{BC} считается соответствие

$$R_{AB} \circ R_{BC} = \{(a_T, c_V) : \exists b_U \ (a_T R_{AB} b_U \wedge b_U R_{BC} c_V)\}.$$

2. Подходы к понятию равносильности недоопределённых алфавитов

Опишем несколько подходов к введению понятия равносильности для соответственных недоопределённых алфавитов. Первый из них — функциональный — основан на функциональной выразимости символов одного алфавита через символы другого. Следующие три подхода терминологически связаны с подходами к введению меры информации, представленными в работе А. Н. Колмогорова [1]. Это комбинаторный, вероятностный (статистический) и алгоритмический подходы.

2.1. Функциональный подход

Рассмотрим недоопределённые алфавиты A и B, связанные соответствием R_{AB}. Пусть A_0 и B_0 — ассоциированные с A и B основные алфавиты. Всякую функцию $F : A_0 \to B_0$ можно распространить на A, положив $F(a_T) = \{F(a_i) : a_i \in a_T\}$. Скажем, что *алфавит B функционально выразим через A*, если существует функция $F : A_0 \to B_0$, такая, что для всех пар $(a_T, b_U) \in R_{AB}$ имеет место $F(a_T) \subseteq b_U$. Последнее означает, что символ b_U может быть получен из a_T функциональным преобразованием F и размытием. Будем говорить, что алфавит A *функционально сильнее* B (B *функционально слабее* A), и записывать $A \succsim_f B$, если B функционально выразим через A. В случае $A \succsim_f B$ и $B \succsim_f A$ будем алфавиты A и B называть *функционально равносильными* и записывать $A \eqsim_f B$. Соотношение $A \eqsim_f B$ означает в развёрнутой записи, что для некоторых функций $F : A_0 \to B_0$ и $G : B_0 \to A_0$

$$a_T R_{AB} b_U \Rightarrow F(a_T) \subseteq b_U \wedge G(b_U) \subseteq a_T. \tag{1}$$

Равносильность алфавитов была определена через их соотношение по силе. Покажем теперь, как соотношение по силе может быть выражено через равносильность. Для соответственных алфавитов A и B введем алфавит AB, символы $a_T b_U$ которого ассоциированы с парами $(a_T, b_U) \in R_{AB}$, и определим соответствие $R_{AB,A} = \{(a_T b_U, a_T) : (a_T, b_U) \in R_{AB}\}$.

Лемма 1. Соотношение $A \succsim_f B$ выполнено тогда и только тогда, когда $AB \eqsim_f A$.

Доказательство. Если справедливо $A \succsim_f B$ и это соотношение выполняется с функцией $F : A_0 \to B_0$, то для получения соотношения $A \succsim_f AB$ достаточно взять в качестве $F' : A_0 \to (A_0 \times B_0)$ функцию $F'(a_i) = a_i F(a_i)$, а для соотношения $AB \succsim_f A$ — функцию $G' : (A_0 \times B_0) \to A_0$, где $G'(a_i b_j) = a_i$. В результате получаем $AB \eqsim_f A$.

Обратно, если имеет место $AB \eqsim_f A$ и соотношение $A \succsim_f AB$ установлено применением функции $F : A_0 \to (A_0 \times B_0)$, то функцию $F' : A_0 \to B_0$ для $A \succsim_f B$ можно получить, назначив $F'(a_i) = b_j$, где b_j определяется значением $F(a_i) = a_u b_j$. ■

Соотношения $A \succsim_f B$ и $A \eqsim_f B$ могут быть эквивалентно представлены в терминах соответственных последовательностей, а именно: $A \succsim_f B$ имеет место тогда и только тогда, когда существует такая функция $F : A_0 \to B_0$, что для всякой пары $\mathbf{a} = a_{T_1} \dots a_{T_n}$, $\mathbf{b} = b_{U_1} \dots b_{U_n}$ соответственных последовательностей и любого доопределения $\mathbf{a}^0 = a_{i_1} \dots a_{i_n}$ последовательности $\mathbf{a}$ последовательность $F(\mathbf{a}^0) = F(a_{i_1}) \dots F(a_{i_n})$ доопределяет $\mathbf{b}$. Для соотношения $A \eqsim_f B$ дополнительно требуется существование функции $G : B_0 \to A_0$, применение которой к любому доопределению $\mathbf{b}^0$ последовательности $\mathbf{b}$ даёт последовательность $G(\mathbf{b}^0)$, доопределяющую $\mathbf{a}$.

Другие понятия равносильности будут представлены в терминах соответственных последовательностей и по форме будут подобны лемме 1.

2.2. Комбинаторный подход

Для последовательности $\mathbf{a}$ в алфавите A обозначим через $\mathcal{K}(\mathbf{a})$ класс всех последовательностей $\mathbf{a}'$ в алфавите A, в которых каждый символ $a_T \in A$ встречается такое же, как в $\mathbf{a}$, число раз. Пусть $N(\mathbf{a})$ — минимальная мощность множества последовательностей в основном алфавите A_0, среди которых имеются доопределения всех последовательностей $\mathbf{a}'$ из $\mathcal{K}(\mathbf{a})$. Величина $\log N(\mathbf{a})$ называется *комбинаторной энтропией* класса $\mathcal{K}(\mathbf{a})$ [5] (всюду под $\log x$ понимается $\log_2 x$).

Пусть A и B — соответственные алфавиты. Пару последовательностей $\mathbf{a} = a_{T_1} \dots a_{T_n}$ и $\mathbf{b} = b_{U_1} \dots b_{U_n}$ одинаковой длины в алфавитах A и B будем воспринимать также как последовательность пар (a_{T_i}, b_{U_i}), $i = 1, \dots, n$. Обозначим через $\mathcal{K}(\mathbf{a}, \mathbf{b})$ класс всех пар последовательностей $(\mathbf{a}', \mathbf{b}')$, в которых каждая пара $(a_T, b_U) \in A \times B$ встречается столько же раз, сколько в $(\mathbf{a}, \mathbf{b})$. Отметим, что если пара $(\mathbf{a}, \mathbf{b})$ соответственна, то и каждая из пар $(\mathbf{a}', \mathbf{b}')$ соответственна. Через $N(\mathbf{a}, \mathbf{b})$ обозначим минимальную мощность множества пар $(\mathbf{u}^0, \mathbf{v}^0)$ полностью определённых последовательностей, среди которых имеются доопределения всех пар последовательностей из $\mathcal{K}(\mathbf{a}, \mathbf{b})$.

Будем говорить, что алфавит A *комбинаторно сильнее* алфавита B, и записывать $A \succsim_c B$, если для любых соответственных последовательностей $\mathbf{a}$ и $\mathbf{b}$ выполнено $N(\mathbf{a}, \mathbf{b}) = N(\mathbf{a})$. В случае $A \succsim_c B$ и $B \succsim_c A$ будем алфавиты A и B называть *комбинаторно равносильными* и записывать $A \eqsim_c B$. Это означает, что для любых соответственных последовательностей $\mathbf{a}$ и $\mathbf{b}$ выполнено $N(\mathbf{a}) = N(\mathbf{b}) = N(\mathbf{a}, \mathbf{b})$.

2.3. Статистический подход

Рассмотрим *недоопределённые источники* X в алфавите A, порождающие независимо символы $a_T \in A$ с некоторыми вероятностями p_T. Положим $P = (p_T, T \in \mathcal{T})$ и для источника будем использовать обозначение $X = (A, P)$. Задавшись набором $Q = (q_i, i \in M)$ вероятностей символов a_i основного алфавита A_0, введём функцию

$$\mathcal{H}(P, Q) = -\sum_{T \in \mathcal{T}} p_T \log \sum_{i \in T} q_i. \tag{2}$$

Величину

$$\mathcal{H}(P) = \min_Q \mathcal{H}(P, Q)$$

будем называть *энтропией* источника X и обозначать также $\mathcal{H}(X)$. Для недоопределённых источников она играет ту же роль, какую энтропия Шеннона играет для всюду определённых источников (подробнее см. в [5]).

Пусть имеются алфавиты A и B, связанные соответствием R_{AB}. *Источники* X и Y в алфавитах A и B, заданные совместным распределением $p(a_T, b_U)$, $a_T \in A$, $b_U \in B$, назовём *соответственными*, если $p(a_T, b_U) > 0$ только в случае $a_T R_{AB} b_U$.

Будем говорить, что алфавит A *статистически сильнее* алфавита B, и записывать $A \succsim_s B$, если для любых пар соответственных источников X и Y выполнено $\mathcal{H}(XY) = \mathcal{H}(X)$. В случае $A \succsim_s B$ и $B \succsim_s A$ будем алфавиты A и B называть *статистически равносильными* и записывать $A \eqsim_s B$. Это означает, что для любых соответственных источников X и Y выполнено $\mathcal{H}(X) = \mathcal{H}(Y) = \mathcal{H}(XY)$.

2.4. Алгоритмический подход

Приведём некоторые результаты о сложности по Колмогорову [1] и распространим их на случай недоопределённых последовательностей.

Рассматриваются алгоритмы $\varphi(\mathbf{p}) = \mathbf{x}$, переводящие слова в слова. Слово $\mathbf{p}$ предполагается двоичным и называется *программой*, $\mathbf{x}$ — слово в алфавите $A_0 = \{a_i : i \in M\}$. Сложность $K_\varphi(\mathbf{x})$ слова $\mathbf{x}$ по алгоритму φ измеряется минимальной длиной программы $\mathbf{p}$, для которой $\varphi(\mathbf{p}) = \mathbf{x}$, и равна ∞, если такого $\mathbf{p}$ нет. По теореме оптимальности Колмогорова [1] существует алгоритм ψ, такой, что для любого φ найдется константа $c = c_\varphi$, при которой для всех $\mathbf{x}$ выполнено $K_\varphi(\mathbf{x}) \leqslant K_\psi(\mathbf{x}) + c$. Алгоритм ψ с таким свойством называется *оптимальным*. Под *сложностью* $K(\mathbf{x})$ *слова* $\mathbf{x}$ понимается его сложность по любому фиксированному оптимальному алгоритму. При использовании разных оптимальных алгоритмов ψ и ψ' сложности $K(\mathbf{x})$ и $K'(\mathbf{x})$ связаны соотношением $K(\mathbf{x}) \approx K'(\mathbf{x})$, где $f \approx g$ означает, что разность $f - g$ ограничена. Под *сложностью* $K(\mathbf{a})$ *недоопределённой последовательности* $\mathbf{a}$ в алфавите A будем понимать минимальную из сложностей $K(\mathbf{a}^0)$ её доопределений $\mathbf{a}^0$. Эта величина также определена с точностью до $\approx$.

Пусть A и B — соответственные алфавиты. Будем говорить, что алфавит A *алгоритмически сильнее* алфавита B, и записывать $A \succsim_a B$, если для любых соответственных последовательностей $\mathbf{a}$ и $\mathbf{b}$ выполнено $K(\mathbf{ab}) \approx K(\mathbf{a})$. В случае $A \succsim_a B$ и $B \succsim_a A$ будем алфавиты A и B называть *алгоритмически равносильными* и записывать $A \eqsim_a B$. Алгоритмическая равносильность означает, что для любых соответственных последовательностей $\mathbf{a}$ и $\mathbf{b}$ выполнено $K(\mathbf{a}) \approx K(\mathbf{b}) \approx K(\mathbf{ab})$.

3. Доказательство эквивалентности подходов

Далее устанавливается, что все представленные выше подходы задают для недоопределённых алфавитов одинаковые соотношения по силе и, как следствие, одинаковые понятия равносильности.

Лемма 2. Из $A \succsim_f B$ следуют соотношения $A \succsim_c B$ и $A \succsim_a B$.

Доказательство. Пусть справедливо $A \succsim_f B$ и это соотношение выполняется с функцией $F : A_0 \to B_0$. Рассмотрим соответственные последовательности $\mathbf{a}$ и $\mathbf{b}$ в алфавитах A и B.

Если имеется доопределяющее множество для класса $\mathcal{K}(\mathbf{a})$, то, заменив в этом множестве каждую последовательность $\mathbf{u}^0$ парой $(\mathbf{u}^0, F(\mathbf{u}^0))$, получим доопределяющее множество для $\mathcal{K}(\mathbf{a}, \mathbf{b})$. Обратно, из всякого доопределяющего множества для $\mathcal{K}(\mathbf{a}, \mathbf{b})$,

взяв в каждой его паре $(\mathbf{u}^0, \mathbf{v}^0)$ лишь $\mathbf{u}^0$, можно образовать доопределяющее множество для $\mathcal{K}(\mathbf{a})$. Отсюда следует равенство $N(\mathbf{a}, \mathbf{b}) = N(\mathbf{a})$, приводящее к $A \succsim_c B$.

Пусть значение $K(\mathbf{a}) = K_\psi(\mathbf{a})$ достигается на программе $\mathbf{p}$, т. е. совпадает с её длиной $l(\mathbf{p})$. Можно также рассматривать $\mathbf{p}$ как программу алгоритма φ, который сначала находит для $\mathbf{a}$ доопределение $\mathbf{a}^0 = \psi(\mathbf{p})$, а затем по $\mathbf{a}^0$ строит конкатенацию $\mathbf{a}^0 F(\mathbf{a}^0)$, доопределяющую $\mathbf{ab}$. Это даёт

$$K(\mathbf{ab}) \leqslant K_\varphi(\mathbf{ab}) + c_\varphi \leqslant l(\mathbf{p}) + c_\varphi \leqslant K(\mathbf{a}) + c_\varphi.$$

Аналогично, рассмотрев программу, на которой достигается $K(\mathbf{ab})$, и считая её программой алгоритма θ, который находит доопределение для $\mathbf{ab}$, а затем выдаёт его половину, доопределяющую $\mathbf{a}$, приходим к неравенству $K(\mathbf{a}) \leqslant K(\mathbf{ab}) + c_\theta$. В результате получаем соотношение $K(\mathbf{ab}) \approx K(\mathbf{a})$, приводящее к $A \succsim_a B$. ■

Лемма 3. Из $A \succsim_c B$ следует $A \succsim_s B$.

Доказательство. Рассмотрим некоторое кодирование последовательностей $\mathbf{a}$ длины n, порождаемых недоопределённым источником X в алфавите A [5]. Пусть $p(\mathbf{a})$ — вероятность порождения $\mathbf{a}$ источником X; $l_\mathbf{a}$ — длина кода для $\mathbf{a}$. В [5] доказано, что если кодирование разделимо[2], то

$$\sum_{\mathbf{a}\in A^n} p(\mathbf{a}) l_\mathbf{a} \geqslant n\mathcal{H}(X). \tag{3}$$

Там же указано разделимое кодирование, для которого $l_\mathbf{a} \leqslant \log N(\mathbf{a}) + c_1 \log n$ и

$$\sum_{\mathbf{a}\in A^n} p(\mathbf{a}) l_\mathbf{a} \leqslant \sum_{\mathbf{a}\in A^n} p(\mathbf{a}) \log N(\mathbf{a}) + c_1 \log n \leqslant n\mathcal{H}(X) + c_2 \log n.$$

Объединяя эти факты, получаем

$$n\mathcal{H}(X) - c_1 \log n \leqslant \sum_{\mathbf{a}\in A^n} p(\mathbf{a}) \log N(\mathbf{a}) \leqslant n\mathcal{H}(X) + c_2 \log n. \tag{4}$$

Пусть имеет место $A \succsim_c B$ и заданы соответственные источники X и Y. Применим к XY левое неравенство из (4), затем, подставив $N(\mathbf{a}, \mathbf{b}) = N(\mathbf{a})$, воспользуемся для X правой частью (4):

$$\begin{aligned} n\mathcal{H}(XY) - c_3 \log n \leqslant \sum_{(\mathbf{a},\mathbf{b})\in R^n_{AB}} p(\mathbf{a},\mathbf{b}) \log N(\mathbf{a},\mathbf{b}) = \sum_{(\mathbf{a},\mathbf{b})\in R^n_{AB}} p(\mathbf{a},\mathbf{b}) \log N(\mathbf{a}) = \\ = \sum_{\mathbf{a}\in A^n} p(\mathbf{a}) \log N(\mathbf{a}) \leqslant n\mathcal{H}(X) + c_2 \log n. \end{aligned}$$

Разделив обе части на n и перейдя к пределу при $n \to \infty$, получим $\mathcal{H}(XY) \leqslant \mathcal{H}(X)$. Обратное соотношение $\mathcal{H}(XY) \geqslant \mathcal{H}(X)$ справедливо всегда [5].

Равенство $\mathcal{H}(XY) = \mathcal{H}(X)$ означает $A \succsim_s B$. ■

Лемма 4. Из $A \succsim_a B$ следует $A \succsim_s B$.

Доказательство. Если k — натуральное число, а $\sigma_1\sigma_2\ldots\sigma_r$, $r \leqslant \log k + 1$, — его двоичная запись, начинающаяся с 1, то через $\tilde{k}$ будем обозначать двоичное слово $\sigma_1\sigma_1\ldots\sigma_r\sigma_r 01$. Для него $l(\tilde{k}) \leqslant 2\log k + 4$.

[2]Кодирование разделимо, если последовательность произвольно приписанных друг к другу кодовых слов однозначно разбивается на кодовые слова.

Пусть $\mathbf{p_a}$ — программа построения доопределения последовательности $\mathbf{a} \in A^n$ оптимальным алгоритмом, на которой достигается $K(\mathbf{a})$. Если кодами последовательностей $\mathbf{a}$, порождаемых источником X, считать $\mathbf{p_a}$, кодирование может оказаться неразделимым. Чтобы превратить его в разделимое, в качестве кодов будем использовать слова $\tilde{l}(\mathbf{p_a})\mathbf{p_a}$. Их длина удовлетворяет оценке $l_{\mathbf{a}} \leqslant K(\mathbf{a}) + c_4 \log n$. Подставив её в (3), получаем

$$\sum_{\mathbf{a}\in A^n} p(\mathbf{a})K(\mathbf{a}) \geqslant n\mathcal{H}(X) - c_4 \log n. \tag{5}$$

Пусть X и Y — соответственные источники в алфавитах A и B. Аналог для XY неравенства (5) и учёт соотношения $K(\mathbf{ab}) \approx K(\mathbf{a})$ дают

$$\begin{aligned} n\mathcal{H}(XY) - c_5 \log n \leqslant \sum_{(\mathbf{a},\mathbf{b})\in R^n_{AB}} p(\mathbf{a},\mathbf{b})K(\mathbf{ab}) = \\ = \sum_{(\mathbf{a},\mathbf{b})\in R^n_{AB}} p(\mathbf{a},\mathbf{b})K(\mathbf{a}) + c_6 = \sum_{\mathbf{a}\in A^n} p(\mathbf{a})K(\mathbf{a}) + c_6. \end{aligned} \tag{6}$$

Оценим $K(\mathbf{a})$. Пусть $M = \{0, 1, \ldots, m-1\}$, $\mathcal{T} = \{T_1, \ldots, T_s\}$ и символы $a_{T_1}, \ldots, a_{T_s}$ входят в последовательность $\mathbf{a}$ соответственно $u_1, \ldots, u_s$ раз. В работе [6] описан градиентный (жадный) алгоритм построения доопределяющего множества для $\mathcal{K}(\mathbf{a})$, согласно которому некоторым образом находится набор натуральных параметров $(v_0, v_1, \ldots, v_{k-1})$ и доопределяющее множество образуется последовательностями, в которых символы a_i, $i = 0, 1, \ldots, k-1$, встречаются v_i раз. Оно строится последовательно. На каждом шаге добавляется последовательность, которая доопределяет наибольшее число последовательностей класса $\mathcal{K}(\mathbf{a})$, не получивших доопределений на предыдущих шагах, и расположена лексикографически раньше других последовательностей, обладающих этим свойством. С ней связывается номер шага, на котором она включена в множество. В [6] доказано, что мощность $\hat{N}$ построенного множества удовлетворяет оценке $\log \hat{N} \leqslant \log N(\mathbf{a}) + c_7 \log n$.

В качестве программы нахождения доопределения для $\mathbf{a}$ возьмём

$$\mathbf{p_a} = \tilde{n}\tilde{m}\tilde{s}\tilde{u}_1 \ldots \tilde{u}_s\tilde{v}_0 \ldots \tilde{v}_{m-1}\mu,$$

где μ — двоичная запись номера последовательности, доопределяющей $\mathbf{a}$. Она позволяет однозначно указать параметры класса $\mathcal{K}(\mathbf{a})$ и параметры $(v_0, v_1, \ldots, v_{k-1})$ класса, из которого берутся доопределения, а затем применением градиентной процедуры вплоть до шага, двоичной записью которого является μ, найти доопределение для $\mathbf{a}$. Имеем

$$K(\mathbf{a}) \leqslant l(\mathbf{p_a}) + c_8 \leqslant \log \hat{N} + c_9 \log n \leqslant \log N(\mathbf{a}) + c_{10} \log n.$$

Подставив эту оценку в (6), приходим к ситуации, имевшей место в предыдущей лемме, и завершаем доказательство, как там. ■

В дальнейшем понадобится следующий факт из [5].

Утверждение 1. Набор вероятностей Q минимизирует функцию $\mathcal{H}(P, Q)$ из (2) тогда и только тогда, когда при каждом j, $j \in M$, выполнено

$$\sum_{T:\, j\in T} \frac{p_T}{\sum\limits_{k\in T} q_k} \leqslant 1, \tag{7}$$

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

Скажем, что символ a_i *мажорирует* в алфавите A символ a_j, если для всякого $a_T \in A$ принадлежность $a_j \in a_T$ влечёт $a_i \in a_T$. Отношение мажорирования транзитивно и рефлексивно.

Лемма 5. Пусть в недоопределённом алфавите A отсутствуют мажорируемые символы и a_i — произвольный фиксированный символ из A_0. Тогда найдётся набор вероятностей $P = (p_T,\, T \in \mathcal{T})$ со строго положительными p_T, для которого компонента q_i всякого набора Q, минимизирующего функцию $\mathcal{H}(P, Q)$, строго положительна.

Доказательство. Рассмотрим произвольный символ $a_i \in A_0$. Введём обозначения $\mathcal{T}' = \{T : T \in \mathcal{T}, i \in T\}$, $\mathcal{T}'' = \mathcal{T} \setminus \mathcal{T}'$, $u = |\mathcal{T}'|$, $v = |\mathcal{T}''|$. Зададимся параметрами $p > 0$ и $\varepsilon > 0$, удовлетворяющими условию

$$pu + \varepsilon v = 1, \tag{8}$$

и назначим набор вероятностей $P = (p_T,\, T \in \mathcal{T})$, положив

$$p_T = \begin{cases} p, & T \in \mathcal{T}', \\ \varepsilon, & T \in \mathcal{T}''. \end{cases}$$

Покажем, что при подходящем выборе параметров p и ε он удовлетворяет условиям леммы.

Пусть набор Q минимизирует функцию $\mathcal{H}(P, Q)$. Для него справедливо утверждение 1. Рассмотрим произвольное $j \in M$, $j \neq i$. Дальше сумму из левой части (7) будем обозначать $S(T \mid j \in T)$. Будем рассматривать также суммы $S(T \mid \theta(T))$ более общего вида, где $\theta(T)$ — условия на множества T, по которым ведётся суммирование.

а) Если $S(T \mid j \in T) < 1$, то $q_j = 0$ по утверждению 1.

б) Пусть $S(T \mid j \in T) = 1$. Запишем это равенство в виде

$$S(T \mid T \in \mathcal{T}', j \in T) + S(T \mid T \in \mathcal{T}'', j \in T) = 1. \tag{9}$$

Поскольку a_j не мажорируется символом a_i, найдётся $T' \in \mathcal{T}'$, для которого $j \notin T'$. Тогда

$$S(T \mid T \in \mathcal{T}', T \neq T') + \frac{p_{T'}}{\sum\limits_{k \in T'} q_k} = S(T \mid i \in T) \leqslant 1.$$

Отсюда, принимая во внимание соотношения $\sum\limits_{k \in T'} q_k \leqslant 1$ и $p_{T'} = p$, получаем

$$S(T \mid T \in \mathcal{T}', T \neq T') \leqslant 1 - \frac{p_{T'}}{\sum\limits_{k \in T'} q_k} \leqslant 1 - p_{T'} = 1 - p.$$

С учётом этого находим

$$S(T \mid T \in \mathcal{T}', j \in T) \leqslant S(T \mid T \in \mathcal{T}', T \neq T') \leqslant 1 - p.$$

Это неравенство и (9) дают $1 - p + S(T \mid T \in \mathcal{T}'', j \in T) \geqslant 1$, что приводит к

$$p \leqslant \sum_{T:\, T \in \mathcal{T}'', j \in T} \frac{p_T}{\sum\limits_{k \in T} q_k} \leqslant \sum_{T:\, T \in \mathcal{T}'', j \in T} \frac{p_T}{q_j}.$$

Подставив сюда $|\mathcal{T}''| = v$ и $p_T = \varepsilon$ для $T \in \mathcal{T}''$, получаем $\dfrac{\varepsilon v}{q_j} \geqslant p$, т. е. $q_j \leqslant \dfrac{\varepsilon v}{p}$. Учитывая пункт (а), заключаем, что эта оценка справедлива для всех $j \neq i$. Поэтому

$$q_i = 1 - \sum_{j,\, j \neq i} q_j \geqslant 1 - \frac{\varepsilon v(u+v)}{p}.$$

Отсюда и из (8) следует, что при достаточно малом ε выполнено $q_i > 0$. ■

Пусть задан недоопределённый источник $X = (A, P)$ и символ $a_j \in A_0$ мажорируется в A символом a_i, $a_i \neq a_j$. Введём операцию *исключения мажорируемого символа* a_j из алфавита A и из источника X, при выполнении которой каждый символ $a_T \in A$ заменяется символом[3] $a_{T \setminus j}$ (поскольку a_j мажорируется некоторым a_i, множество T' непусто). В результате операции получается алфавит $A' = \{a_{T'} : T' = T \setminus j,\ a_T \in A\}$, для которого основным алфавитом является $A'_0 = A_0 \setminus a_j$. Источник $X = (A, P)$ преобразуется в $X' = (A', P')$, где набор P' образован вероятностями $p'_{T'} = p_{T'} + p_{T' \cup j}$. Здесь $p_{T'}$ и $p_{T' \cup j}$ — вероятности символов $a_{T'}$ и $a_{T' \cup j}$ источника X; при отсутствии какого-либо из этих символов в алфавите A соответствующая вероятность считается равной нулю.

Если наряду с X задан источник Y с алфавитом B, то при устранении из A мажорируемого символа a_j соответствие R_{AB} переходит в

$$R_{A'B} = \{(a_{T'}, b_U) : \exists T (T \setminus j = T' \wedge (a_T, b_U) \in R_{AB})\}. \tag{10}$$

Совместное распределение P_{XY} источников X и Y образовано вероятностями $p_{TU} = p(a_T, b_U)$ для пар $(a_T, b_U) \in R_{AB}$, а совместное распределение $P_{X'Y}$ источников X' и Y — вероятностями $p'_{T'U} = p_{T'U} + p_{T' \cup j, U}$, где $(a_{T'}, b_U) \in R_{A'B}$. Здесь вероятности $p_{T'U}$ и $p_{T' \cup j, U}$ берутся из P_{XY}, а если какой-либо из них в P_{XY} нет, она полагается равной нулю. В случае соответственных источников X и Y источники X' и Y также соответственны.

Лемма 6. Если источник X' образован из X удалением мажорируемого символа, то $\mathcal{H}(X') = \mathcal{H}(X)$ и $\mathcal{H}(X'Y) = \mathcal{H}(XY)$.

Доказательство. Докажем равенство $\mathcal{H}(X'Y) = \mathcal{H}(XY)$. Будем считать для определённости, что исключаемым из A символом является a_1, и он мажорируется символом a_2. Если в $(a_T, b_U) \in R_{AB}$ имеется символ (a_1, b_u), то там имеется и (a_2, b_u). Поэтому (a_2, b_u) мажорирует (a_1, b_u).

Функция $\mathcal{H}(P_{XY}, Q)$ для произведения XY имеет вид

$$\mathcal{H}(P_{XY}, Q) = - \sum_{(T,U) : a_T R_{AB} b_U} p_{TU} \log \sum_{i \in T, j \in U} q_{ij}. \tag{11}$$

Легко понять, что функция $\mathcal{H}(P_{X'Y}, Q')$ для $X'Y$ может быть получена из неё подстановкой нуля вместо всех q_{1j}. Поскольку минимум при отсутствии ограничений на Q не превосходит минимума по наборам Q с условием $q_{1j} = 0$, выполнено $\mathcal{H}(X'Y) \geqslant \mathcal{H}(XY)$.

Рассмотрим набор Q, на котором достигается минимум в (11) т. е. значение $\mathcal{H}(X, Y)$. Образуем из Q набор Q', удалив все компоненты q_{1j} и заменив компоненты q_{2j} на $q_{1j} + q_{2j}$. Так как символ (a_2, b_j) мажорирует (a_1, b_j), значение каждой из сумм $\sum\limits_{i \in T, j \in U} q_{ij}$

[3] Правильнее было бы писать $a_{T \setminus \{j\}}$, но в целях простоты записей мы не различаем одноэлементные множества и элементы.

из (11) при отбрасывании q_{1j} и подстановке $q_{1j} + q_{2j}$ вместо q_{2j} не уменьшится. Это даёт

$$\mathcal{H}(XY) = \mathcal{H}(P_{XY}, Q) \geqslant \mathcal{H}(P_{X'Y}, Q') \geqslant \mathcal{H}(X'Y).$$

С учётом предшествующего неравенства получаем $\mathcal{H}(X'Y) = \mathcal{H}(XY)$.

Равенство $\mathcal{H}(X') = \mathcal{H}(X)$ доказывается аналогично. ■

Обозначим через $\hat{X} = (\hat{A}, \hat{P})$ источник, полученный из $X = (A, P)$ последовательным удалением мажорируемых символов, пока они есть.

Следствие 1. Имеют место равенства $\mathcal{H}(\hat{X}) = \mathcal{H}(X)$ и $\mathcal{H}(\hat{X}Y) = \mathcal{H}(XY)$.

Лемма 7. Из $A \succsim_s B$ следует $A \succsim_f B$.

Доказательство. Рассмотрим недоопределённые алфавиты A и B, удовлетворяющие условию $A \succsim_s B$.

а) Сначала будем полагать, что в алфавите A нет мажорируемых символов.

Пусть a_i — произвольный символ из A_0. В соответствии с леммой 5 возьмём набор $P = (p_T,\, T \in \mathcal{T})$ положительных вероятностей, для которого в любом наборе Q, минимизирующем $\mathcal{H}(P, Q)$, компонента q_i положительна. Рассмотрим источники X и Y в алфавитах A и B, заданные совместным распределением p_{TU}, $T \in \mathcal{T}$, $U \in \mathcal{U}$, удовлетворяющим условиям

$$p_{TU} > 0 \Leftrightarrow a_T R_{AB} b_U, \quad \sum_U p_{TU} = p_T,$$

где p_T — компонента выбранного набора P. Источники X и Y соответственны, поэтому $\mathcal{H}(XY) = \mathcal{H}(X)$.

Пусть $\mathcal{H}(XY)$ достигается на наборе $Q = (q_{uj}, u \in M, j \in L)$. Введём величины $q_u^0 = \sum\limits_j q_{uj}$ и положим $Q^0 = (q_u^0,\, u \in M)$. Принимая во внимание равенство $\sum\limits_u q_u^0 = \sum\limits_{u,j} q_{uj} = 1$, находим

$$\begin{aligned} \mathcal{H}(XY) &= -\sum_{T,U} p_{TU} \log \sum_{u \in T, j \in U} q_{uj} \geqslant -\sum_{T,U} p_{TU} \log \sum_{u \in T} q_u^0 = \\ &= -\sum_T p_T \log \sum_{u \in T} q_u^0 = \mathcal{H}(P, Q^0) \geqslant \mathcal{H}(X). \end{aligned} \tag{12}$$

Значения левой и правой частей совпадают, поэтому все неравенства могут быть заменены равенствами. Одно из них имеет вид $\mathcal{H}(P, Q^0) = \mathcal{H}(X)$ и потому по выбору P выполнено $q_i^0 > 0$. Следовательно, при некотором j_i имеет место $q_{ij_i} > 0$.

Рассмотрим произвольные T и U, такие, что $i \in T$ и $(a_T, b_U) \in R_{AB}$. Для них $p_{TU} > 0$. Воспользуемся равенством

$$p_{TU} \log \sum_{u \in T, j \in U} q_{uj} = p_{TU} \log \sum_{u \in T} q_u^0,$$

возникшим в (12) при замене неравенств равенствами. В силу определения q_i^0 из него вытекает

$$\sum_{j \in U} q_{uj} = q_i^0 = \sum_j q_{uj}.$$

Так как q_{ij_i} положительно, $j_i \in U$. Это означает, что для заданного $a_i \in A_0$ имеется символ b_{j_i}, такой, что если $a_i \in a_T$ и $a_T R_{AB} b_U$, то $b_{j_i} \in b_U$.

Применим такие рассуждения ко всем символам $a_i \in A_0$ и для каждого из них найдем символ b_{j_i} с указанным свойством. Введём функцию $F : A_0 \to B_0$, положив $F(a_i) = b_{j_i}$. Так определённая функция F удовлетворяет условию

$$a_T R_{AB} b_U \Rightarrow f(a_T) \subseteq b_U. \tag{13}$$

б) Пусть теперь алфавит A произволен и все пары соответственных источников X и Y в алфавитах A и B удовлетворяют условию $\mathcal{H}(XY) = \mathcal{H}(X)$.

Рассмотрим одну из соответственных пар X и Y. Путём последовательного удаления из X всех мажорируемых символов построим источник $\hat{X} = (\hat{A}, \hat{P})$. Пусть его основным алфавитом является $\hat{A}_0 = \{a_i : i \in \hat{M}\}$, тогда $\hat{A} = \{a_{\hat{T}} = a_{T \cap \hat{M}} : a_T \in A\}$. По следствию 1 выполнено $\mathcal{H}(\hat{X}Y) = \mathcal{H}(XY)$ и $\mathcal{H}(\hat{X}) = \mathcal{H}(X)$. Поэтому $\mathcal{H}(\hat{X}Y) = \mathcal{H}(\hat{X})$. Легко видеть, что, произвольно варьируя в источниках X вероятности символов алфавита A, можно в качестве $\hat{A}$ получить все возможные источники в алфавите $\hat{A}$. Поэтому к алфавитам $\hat{A}$ и B применим результат пункта (а), в соответствии с которым существует функция $\hat{F} : \hat{A}_0 \to B_0$, удовлетворяющая условию $a_T R_{\hat{A}B} b_U \Rightarrow \hat{F}(a_{\hat{T}}) \subseteq b_U$. На её основе образуем функцию $F : A_0 \to B_0$, положив $F(a_i) = \hat{F}(a_i)$ для $a_i \in \hat{A}_0$ и назначив для $a_i \in A_0 \setminus \hat{A}_0$ значение $F(a_i)$ равным $\hat{F}(a_u)$, где a_u — символ из $\hat{A}_0$, мажорирующий a_i. Функция F удовлетворяет условию (13), ибо для $a_i \in a_T \setminus a_{\hat{T}}$ в $a_{\hat{T}}$ имеется мажорирующий a_i символ a_u и $F(a_i) = \hat{F}(a_u) \in b_U$. ■

Объединяя результаты лемм 2, 3, 4 и 7, получаем следующий факт.

Теорема 1. Введённые соотношения недоопределённых алфавитов по силе эквивалентны, т. е.

$$A \gtrsim_f B \Leftrightarrow A \gtrsim_c B \Leftrightarrow A \gtrsim_s B \Leftrightarrow A \gtrsim_a B.$$

Введённые понятия равносильности недоопределённых алфавитов эквивалентны, т. е.

$$A \eqsim_f B \Leftrightarrow A \eqsim_c B \Leftrightarrow A \eqsim_s B \Leftrightarrow A \eqsim_a B.$$

С учётом теоремы дальше будем применять записи $A \gtrsim B$ и $A \eqsim B$ без уточнения, в каком смысле они понимаются.

4. Некоторые операции над алфавитами. Приведение

Функциональный подход к введению соотношений $A \gtrsim B$ и $A \eqsim B$ более удобен и конструктивен, чем другие подходы, поскольку имеет дело непосредственно с алфавитами A и B, а не с соответственными последовательностями в этих алфавитах. Дальше будем базироваться на функциональном подходе.

Напомним, что соотношение $A \gtrsim B$ означает существование функции $F : A_0 \to B_0$, для которой $a_T R_{AB} b_U \Rightarrow F(a_T) \subseteq b_U$, а соотношение $A \eqsim B$ — одновременное выполнение соотношений $A \gtrsim B$ и $B \gtrsim A$.

Пусть $A = \{a_T : T \in \mathcal{T}\}$, $B = \{b_U : U \in \mathcal{U}\}$ и $C = \{c_V : V \in \mathcal{V}\}$ — недоопределённые алфавиты, R_{AB}, R_{BC} и R_{AC} — заданные для них соответствия.

Лемма 8. Пусть соответствия удовлетворяют условию

$$R_{AC} \subseteq R_{AB} \circ R_{BC}. \tag{14}$$

Тогда

$$\begin{array}{rcl} A \gtrsim B, B \gtrsim C & \Rightarrow & A \gtrsim C, \\ A \eqsim B, B \eqsim C & \Rightarrow & A \eqsim C. \end{array}$$

Доказательство.

1. Пусть выполнены $A \gtrsim B$ и $B \gtrsim C$ и эти соотношения устанавливаются с использованием функций $F : A_0 \to B_0$ и $G : B_0 \to C_0$. Рассмотрим произвольную пару (a_T, c_V), такую, что $a_T R_{AC} c_V$. Из (14) следует, что при некотором b_U имеют место $a_T R_{AB} b_U$ и $b_U R_{BC} c_V$. Тогда $F(a_T) \subseteq b_U$, а потому $G(F(a_T)) \subseteq G(b_U) \subseteq c_V$ и в качестве функции $A_0 \to C_0$ в соотношении $A \gtrsim C$ может быть взята $G(F)$.

2. В случае $A \eqsim B$ и $B \eqsim C$ справедливы, в частности, соотношения $A \gtrsim B$ и $B \gtrsim C$. Согласно п. 1 доказательства, из них следует $A \gtrsim C$. Кроме того, включение (14) может быть эквивалентно переписано в виде $R_{CA} \subseteq R_{CB} \circ R_{BA}$. Из него и соотношений $C \gtrsim B$ и $B \gtrsim A$, вытекающих из $B \eqsim C$ и $A \eqsim B$, в силу п. 1 следует $C \gtrsim A$. В результате получаем $A \eqsim C$. ■

Согласно лемме 8, соотношения $\gtrsim$ и $\eqsim$ транзитивны при условии (14).

Дальше будет встречаться ситуация, когда алфавит A преобразуется в B последовательно: $A = A^{(0)} \to A^{(1)} \to \ldots \to A^{(s)} = B$. При этом алфавиты A и B связаны соответствием $R_{AB} = R_{A^{(0)}A^{(s)}}$ и для каждого шага i, $i = 1, \ldots, s$, имеется соответствие $R_{A^{(i-1)}A^{(i)}}$. Произведение соответствий $R_{A^{(0)}A^{(1)}} \circ R_{A^{(1)}A^{(2)}} \circ \ldots \circ R_{A^{(s-1)}A^{(s)}}$ понимается в обычном смысле:

$$a^{(0)}(R_{A^{(0)}A^{(1)}} \circ R_{A^{(1)}A^{(2)}} \circ \ldots \circ R_{A^{(s-1)}A^{(s)}})a^{(s)} \Leftrightarrow$$
$$\Leftrightarrow \exists a^{(1)} \ldots \exists a^{(s-1)}(a^{(0)} R_{A^{(0)}A^{(1)}} a^{(1)} \wedge \ldots \wedge a^{(s-1)} R_{A^{(s-1)}A^{(s)}} a^{(s)}).$$

Аналогично лемме 8 доказывается следующее её обобщение.

Лемма 9. Пусть соответствия удовлетворяют условию

$$R_{A^{(0)}A^{(s)}} \subseteq R_{A^{(0)}A^{(1)}} \circ R_{A^{(1)}A^{(2)}} \circ \ldots \circ R_{A^{(s-1)}A^{(s)}}. \tag{15}$$

Тогда

$$A^{(0)} \gtrsim A^{(1)}, A^{(1)} \gtrsim A^{(2)}, \ldots, A^{(s-1)} \gtrsim A^{(s)} \Rightarrow A^{(0)} \gtrsim A^{(s)},$$
$$A^{(0)} \eqsim A^{(1)}, A^{(1)} \eqsim A^{(2)}, \ldots, A^{(s-1)} \eqsim A^{(s)} \Rightarrow A^{(0)} \eqsim A^{(s)}.$$

Пусть заданы недоопределённый алфавит A и функция $F : A_0 \to A_0$. Скажем, что алфавит A' получен из A *функциональным преобразованием* F, если $A' = \{F(a_T) : a_T \in A\}$, $R_{AA'} = \{(a_T, F(a_T)) : a_T \in A\}$.

Пусть недоопределённый алфавит A не содержит символа a_T, совпадающего с a_j. Тогда выполнима операция *исключения символа* a_j, в результате которой возникает алфавит $A' = \{a_{T\setminus j} : a_T \in A\}$, связанный с A соответствием $R_{AA'} = \{(a_T, a_{T\setminus j}) : a_T \in A\}$. Отметим что операция исключения мажорируемого символа a_j выполнима всегда, ибо всякий символ a_T, содержащий a_j, содержит и мажорирующий его символ a_i.

Лемма 10. Если A' получен из A посредством а) функционального преобразования, б) исключения символа, в) исключения мажорируемого символа, то а) $A \gtrsim A'$, б) $A' \gtrsim A$, в) $A' \eqsim A$.

Доказательство. Пункт (а) фактически следует из определения, пункт (б) — из того, что всякое доопределение символа $a_{T\setminus j}$ доопределяет a_T. Рассмотрим пункт (в). Пусть удаляемый символ a_j мажорируется символом a_i. Операцию удаления a_j можно трактовать как функциональное преобразование F, при котором $F(a_j) = a_i$ и $F(a_u) = a_u$ для $u \neq j$. Поэтому, согласно (а), выполнено $A \gtrsim A'$. С другой стороны, это операция удаления символа, и в соответствии с (б) имеет место $A' \gtrsim A$. В итоге получаем $A' \eqsim A$. ■

Рассмотрим теперь случай последовательного исключения из алфавита A мажорируемых символов. Обозначим через J множество индексов j исключенных символов a_j. Результат исключения даёт алфавит $A_J = \{a_{T\setminus J} : a_T \in A\}$, связанный с A соответствием $R_{AA_J} = \{(a_T, a_{T\setminus J}) : a_T \in A\}$.

Лемма 11. Алфавит, полученный последовательным исключением мажорируемых символов, равносилен исходному.

Доказательство. Применим индукцию по процедуре исключения. Утверждение леммы для одноэлементного множества J вытекает из пункта (в) леммы 10. Предположим, что утверждение справедливо для множества J, и пусть из алфавита A_J исключается символ $a_{j'}$. Положим $J' = J \cup j'$ и рассмотрим произвольную пару $(a_T, a_{T\setminus J'}) \in R_{AA_{J'}}$. Символ $a_{T\setminus J'}$ возник из $a_{T\setminus J} \in A_J$ в результате исключения $a_{j'}$, а потому $(a_T, a_{T\setminus J}) \in R_{AA_J}$ и $(a_{T\setminus J}, a_{T\setminus J'}) \in R_{A_JA_{J'}}$. Это означает $(a_T, a_{T\setminus J'}) \in R_{AA_J} \circ R_{A_JA_{J'}}$ и приводит к включению $R_{AA_{J'}} \subseteq R_{AA_J} \circ R_{A_JA_{J'}}$, совпадающему для этих соответствий с условием (14). Из соотношений $A \approx A_J$ и $A_J \approx A_{J'}$, первое из которых выполнено по предположению индукции, а второе — по лемме 10, в силу леммы 8 вытекает $A \approx A_{J'}$. ■

Недоопределённый алфавит, у которого отсутствуют мажорируемые символы, называется *приведённым*. Последовательно устраняя в произвольном порядке из алфавита A мажорируемые символы (пока это возможно), придём к некоторому приведённому алфавиту $\hat{A}$. Если $\hat{M} \subseteq M$ — множество индексов неустранённых символов, то $\hat{A} = \{a_{T\cap\hat{M}} : a_T \in A\}$, $R_{A\hat{A}} = \{(a_T, a_{T\cap\hat{M}}) : a_T \in A\}$. По лемме 11 $\hat{A} \approx A$.

Устраняя из A мажорируемые символы в другом порядке, можно получить другой приведённый алфавит $\check{A}$, найти соответствие $R_{A\check{A}}$, затем соответствие $R_{\hat{A}\check{A}} = R_{\hat{A}A} \circ \circ R_{A\check{A}}$, где $R_{\hat{A}A} = R^{-1}_{A\hat{A}}$.

Соответственные алфавиты A и B назовем *изоморфными*, если существует биекция $\pi : A_0 \to B_0$, такая, что соответствие

$$R_{\pi(A)B} = \{(\pi(a_T), b_U) : (a_T, b_U) \in R_{AB}\}$$

является диагональю. Здесь $\pi(a_T) = \{\pi(a_i) : i \in T\}$.

Лемма 12. Изоморфные алфавиты равносильны.

Доказательство. В качестве функции F в (1) может быть взята биекция π, а в качестве G — её обращение π^{-1}. Поскольку $R_{\pi(A)B}$ — диагональ, для $(a_T, b_U) \in R_{AB}$ выполнено $\pi(a_T) = b_U$ и $\pi^{-1}(b_U) = a_T$. ■

Следующее утверждение показывает, что приведённый алфавит единственен с точностью до переименования основных символов.

Теорема 2. Все приведённые алфавиты, образованные из заданного алфавита A, изоморфны.

Доказательство. Символ a_j алфавита A назовем *строго мажорируемым*, если в A существует такой символ a_i, что a_i мажорирует a_j и a_j не мажорирует a_i. Поскольку отношение мажорирования транзитивно, символы, не являющиеся строго мажорируемыми, разбиваются на классы эквивалентности, состоящие из символов a_u, индексы u которых входят в одну и ту же совокупность множеств T, $T \in \mathcal{T}$. В результате построения приведённого алфавита будут исключены все строго мажорируемые символы, а в каждом классе эквивалентности нестрого мажорируемых символов остаётся один. Если $\hat{A}$ и $\check{A}$ — различные приведённые алфавиты и им соответствуют множества неисключённых символов $\hat{A}_0$ и $\check{A}_0$, то можно задать биекцию $\pi : \hat{A}_0 \to \check{A}_0$, сопоставив

символу $a_{\hat{T}} \in \hat{A}_0$ единственный символ $a_{\check{T}} \in \check{A}_0$ из того же класса эквивалентности. В приведённых алфавитах $\hat{A}$ и $\check{A}$ символу $a_T \in A$ соответствуют символы $a_{\hat{T}} = a_T \cap \hat{A}_0$ и $a_{\check{T}} = a_T \cap \check{A}_0$, такие, что $a_{\check{T}} = \pi(a_{\hat{T}})$. Поэтому соответствие $R_{\pi(\hat{A})\check{A}}$, образованное парами $(\pi(a_{\hat{T}}), a_{\check{T}})$, является диагональю. ■

В качестве равносильных преобразований мы использовали удаления мажорируемых символов. Поскольку соотношение равносильности симметрично, в результате добавления мажорируемых символов также возникают равносильные алфавиты.

Если требуется решить некоторую прикладную задачу, то переход от исходного алфавита к равносильному может упростить её решение либо даже сделать нерешаемую задачу решаемой. Приведём два примера. Первый относится к задаче сжатия недоопределённых данных и использует сокращение основного алфавита за счёт удаления мажорируемых символов, второй относится к задаче (двоичного) разложения недоопределённого алфавита и сопровождается увеличением основного алфавита за счёт введения дополнительных мажорируемых символов.

Пример 1. Задача сжатия недоопределённых данных ставится как задача такого кодирования недоопределённых последовательностей, которое обеспечиват возможность восстановления какого-либо их доопределения [5]. Пусть имеются равносильные алфавиты A и B и соответственные последовательности $\mathbf{a} = a_{T_1} \dots a_{T_n}$ и $\mathbf{b} = b_{U_1} \dots b_{U_n}$ в этих алфавитах. Всякое кодирование последовательности $\mathbf{a}$ может рассматриваться также как кодирование последовательности $\mathbf{b}$. Действительно, применив функцию F к доопределению $\mathbf{a}^0 = a_{i_1} \dots a_{i_n}$ последовательности $\mathbf{a}$, найденному по её коду, получим последовательность $F(\mathbf{a}^0) = F(a_{i_1}) \dots F(a_{i_n})$, которая в силу условий $F(a_{T_i}) \subseteq b_{U_i}$ доопределяет $\mathbf{b}$. Подобным же образом код для $\mathbf{b}$ может рассматриваться как код для $\mathbf{a}$. Если кодирование в одном из алфавитов оптимально (минимизирует среднюю длину кода), то оно оптимально и во втором.

В настоящее время не известно каких-либо эффективных методов кодирования недоопределённых данных. Для них применяется либо метод случайного кодирования, либо жадный (градиентный) алгоритм (первый не конструктивен, второй не эффективен). Не исключена возможность, что в результате приведения заданного недоопределённого алфавита возникнет всюду определённый алфавит. Тогда известные эффективные методы кодирования всюду определённых данных [7] обеспечат эффективное кодирование последовательностей в исходном недоопределённом алфавите. Приведём простейший пример этой ситуации. Пусть основным алфавитом является $A_0 = \{a_0, a_1, a_2, a_3\}$, а недоопределённым — $A = \{a_1, a_{23}, a_{02}\}$. Символы a_0 и a_3 мажорируются символом a_2. В результате исключения мажорируемых символов придём к всюду определённому алфавиту $\hat{A} = \{a_1, a_2\}$, оптимальное кодирование в котором обеспечит оптимальное кодирование в исходном недоопределённом алфавите.

Пример 2. Задача (двоичного) разложения недоопределённого алфавита состоит в том [3], чтобы каждому символу a_i основного алфавита A_0 сопоставить двоичное слово $\lambda_i = \lambda_i(1) \dots \lambda_i(s)$ некоторой длины s, а каждому недоопределённому символу $a_T \in A$ — двоичное слово $\lambda_T = \lambda_T(1) \dots \lambda_T(s)$ длины s в алфавите $\{0, 1, *\}$ так, чтобы множество доопределений слова λ_T совпало с $\{\lambda_i : i \in T\}$. Может оказаться, что такое разложение невыполнимо для заданного алфавита, но становится возможным при переходе к равносильному алфавиту. Проиллюстрируем это.

Пусть $A_0 = \{a_0, a_1, a_2, a_3, a_4\}$ и требуется разложить недоопределённый алфавит $A = \{a_{01}, a_{12}, a_{23}, a_{34}, a_{40}\}$. Предположим, что разложение существует и задаётся словами λ_i и $\lambda_{is(i)}$, $i = 0, \dots, 4$, $s(i) = (i+1) \bmod 5$. Слово $\lambda_{is(i)}$, имеющее два доопределения,

содержит единственный символ $*$, а потому λ_i и $\lambda_{s(i)}$ различаются в одной позиции. Последовательность $\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4\lambda_0$ позволяет получить λ_0 из λ_0 за 5 шагов, на каждом из которых изменяется один символ, но для нечётного числа шагов это невозможно. Полученное противоречие показывает, что рассматриваемый алфавит неразложим.

Приведём разложение равносильного алфавита, найденное методом работы [3]. Основной алфавит разложения образован словами $\lambda_0 = 100$, $\lambda_1 = 110$, $\lambda_2 = 011$, $\lambda_3 = 001$, $\lambda_4 = 000$, $\lambda_5 = 010$, $\lambda_6 = 111$, а недоопределённый алфавит — словами $\lambda_{01} = 1{*}0$, $\lambda_{12} = {*}1{*}$, $\lambda_{23} = 0{*}1$, $\lambda_{34} = 00{*}$, $\lambda_{40} = {*}00$. Все недоопределённые слова $\lambda_{is(i)}$, исключая λ_{12}, имеют нужные доопределения λ_i и $\lambda_{s(i)}$, а у слова λ_{12} имеются, помимо λ_1 и λ_2, доопределения λ_5 и λ_6. Но лишние доопределения λ_5 и λ_6 мажорируются каждым из слов λ_1 и λ_2 и могут быть удалены с сохранением равносильности.

5. Распознавание равносильности алфавитов

Выше рассмотрены равносильные преобразования заданного алфавита A, теперь поставим вопрос о равносильности двух исходно заданных соответственных алфавитов $A = \{a_T : T \in \mathcal{T} \subseteq 2^M\}$ и $B = \{b_U : U \in \mathcal{U} \subseteq 2^L\}$. Путём исключения мажорируемых символов в алфавитах A и B перейдём к приведённым алфавитам $\hat{A}$ и $\hat{B}$. Они единственны с точностью до изоморфизма и равносильны исходным алфавитам. Приведённые алфавиты имеют вид $\hat{A} = \{a_{T\cap\hat{M}} : a_T \in A\}$ и $\hat{B} = \{b_{U\cap\hat{L}} : b_U \in B\}$, где $\hat{M}$ и $\hat{L}$ — множества индексов неисключённых символов в алфавитах A и B. Приведённые алфавиты связаны соответствием $R_{\hat{A}\hat{B}} = \{(a_{T\cap\hat{M}}, b_{U\cap\hat{L}}) : a_T R_{AB} b_U\}$.

Теорема 3. Соответственные алфавиты A и B равносильны тогда и только тогда, когда построенные по ним приведённые алфавиты $\hat{A}$ и $\hat{B}$ изоморфны.

Доказательство.

1. Пусть приведённые алфавиты $\hat{A}$ и $\hat{B}$ изоморфны. В силу лемм 11 и 12 имеют место равносильности $A \approx \hat{A}$, $\hat{A} \approx \hat{B}$ и $\hat{B} \approx B$. Рассмотрим произвольную пару $(a_T, b_U) \in R_{AB}$. Из $a_T R_{A\hat{A}} a_{T\cap\hat{M}}$, $a_{T\cap\hat{M}} R_{\hat{A}\hat{B}} b_{U\cap\hat{L}}$ и $b_{U\cap\hat{L}} R_{\hat{B}B} b_U$ следует включение $R_{AB} \subseteq R_{A\hat{A}} \circ R_{\hat{A}\hat{B}} \circ R_{\hat{B}B}$, играющее роль (15). По лемме 9 заключаем, что $A \approx B$.

2. Пусть имеет место равносильность $A \approx B$. Наряду с ней по лемме 11 справедливы равносильности $\hat{A} \approx A$ и $B \approx \hat{B}$. Пары из $R_{\hat{A}\hat{B}}$ имеют вид $(a_{T\cap\hat{M}}, b_{U\cap\hat{L}})$, где $(a_T, b_U) \in R_{AB}$. С учётом $a_{T\cap\hat{M}} R_{\hat{A}A} a_T$ и $b_U R_{B\hat{B}} b_{U\cap\hat{L}}$ заключаем, что выполнено включение $R_{\hat{A}\hat{B}} \subseteq R_{\hat{A}A} \circ R_{AB} \circ R_{B\hat{B}}$, играющее роль (15). Воспользовавшись леммой 9, приходим к равносильности $\hat{A} \approx \hat{B}$. Из неё следует существование функций $F : \hat{A}_0 \to \hat{B}_0$ и $G : \hat{B}_0 \to \hat{A}_0$, таких, что для $(a_{T\cap\hat{M}}, b_{U\cap\hat{U}}) \in R_{\hat{A}\hat{B}}$ выполнено $F(a_{T\cap\hat{M}}) \subseteq b_{U\cap\hat{L}}$ и $G(b_{U\cap\hat{L}}) \subseteq a_{T\cap\hat{M}}$.

Пусть символ $a_i \in \hat{A}_0$ произволен, $F(a_i) = b_j$, $G(b_j) = a_u$. Покажем, что $a_u = a_i$.

Возьмём любой символ $a_{T\cap\hat{M}} \in \hat{A}$, содержащий a_i, и рассмотрим произвольную пару $(a_{T\cap\hat{M}}, b_{U\cap\hat{L}}) \in R_{\hat{A}\hat{B}}$. Из $F(a_{T\cap\hat{M}}) \subseteq b_{U\cap\hat{L}}$ следует $b_j \in b_{U\cap\hat{L}}$, и в силу $G(b_{U\cap\hat{L}}) \subseteq a_{T\cap\hat{M}}$ справедливо $a_u \in a_{T\cap\hat{M}}$. Так как символ $a_{T\cap\hat{M}}$, содержащий a_i, произволен, символ a_u мажорирует a_i и обязан совпасть с a_i, поскольку приведённый алфавит $\hat{A}$ не содержит отличных от a_i символов, мажорирующих a_i. Одновременно установлено, что

$$F(a_i) = b_j \Rightarrow G(b_j) = a_i. \tag{16}$$

Функция F инъективна, ибо в силу (16) равенства $F(a_i) = b_j$ и $F(a_{i'}) = b_j$ влекут $a_i = a_{i'}$. Аналогично (16) можно доказать, что $G(b_j) = a_u \Rightarrow F(a_u) = b_j$. Отсюда вытекает, что F сюръективна, поскольку всякий символ $b_j \in \hat{B}_0$ может быть получен как $F(a_u)$, где $a_u = G(b_j)$. Таким образом, F биективна, а из (16) следует, что $G = F^{-1}$.

Для произвольной пары $(a_{T\cap\hat{M}}, b_{U\cap\hat{L}}) \in R_{\hat{A}\hat{B}}$ выполнено $F(a_{T\cap\hat{M}}) \subseteq b_{U\cap\hat{L}}$ и $F^{-1}(b_{U\cap\hat{L}}) = G(b_{U\cap\hat{L}}) \subseteq a_{T\cap\hat{M}}$. Применив к последнему соотношению функцию F, получаем $b_{U\cap\hat{L}} \subseteq F(a_{T\cap\hat{M}})$, что приводит к $F(a_{T\cap\hat{M}}) = b_{U\cap\hat{L}}$. Это означает, что соответствие $R_{F(\hat{A})\hat{B}}$, образованное парами $(F(a_{T\cap\hat{M}}), b_{U\cap\hat{L}})$, является диагональю и алфавиты $\hat{A}$ и $\hat{B}$ изоморфны. ■

Как обычно [8], *эффективными* будем считать алгоритмы, время работы которых ограничено полиномом от размера исходных данных.

Теорема 4. Для соответственных алфавитов A и B существуют эффективные алгоритмы проверки соотношений $A \gtrsim B$ и $A \approx B$.

Доказательство. Достаточно рассмотреть соотношение $A \approx B$, поскольку $A \gtrsim B$ сводится к нему применением леммы 1. В силу теоремы 3 равносильность алфавитов A и B имеет место тогда и только тогда, когда приведённые алфавиты $\hat{A}$ и $\hat{B}$ изоморфны.

По A, B и соответствию R_{AB} построим (эффективно) $\hat{A}$, $\hat{B}$ и $R_{\hat{A}\hat{B}}$. Произвольно занумеруем $(a_{\hat{T}_s}, b_{\hat{U}_s})$, $s = 1, 2, \ldots, N$, все пары $(a_{\hat{T}}, b_{\hat{U}})$ соответствия $R_{\hat{A}\hat{B}}$. Множества $\hat{M} = \bigcup\limits_s \hat{T}_s$ и $\hat{L} = \bigcup\limits_s \hat{U}_s$ образованы индексами всех символов основных алфавитов $\hat{A}_0$ и $\hat{B}_0$. Для $i \in \hat{M}$ введём набор

$$\eta_i = (\eta_{i1}, \eta_{i2}, \ldots, \eta_{iN}),$$

где η_{is} равны 1 и 0 в случаях $i \in \hat{T}_s$ и $i \notin \hat{T}_s$. Поскольку в $\hat{A}_0$ мажорируемых символов нет, все наборы η_i различны. Аналогично с каждым $j \in \hat{L}$ свяжем набор

$$\xi_j = (\xi_{j1}, \xi_{j2}, \ldots, \xi_{jN}),$$

определяемый принадлежностью j к множествам $\hat{U}_s$. Все наборы ξ_j также различны.

Легко видеть, что $\hat{A}$ и $\hat{B}$ изоморфны тогда и только тогда, когда мощности множеств $\hat{M}$ и $\hat{L}$ совпадают и для каждого $i \in \hat{M}$ имеется (единственное) $j \in \hat{L}$, при котором $\eta_i = \xi_j$. Если для этих i и j положить $b_j = \pi(a_i)$, получим биекцию $\pi : \hat{M} \to \hat{L}$, участвующую в определении изоморфизма. Трудоёмкость описанной процедуры полиномиальна. ■

Из доказательства извлекается полиномиальный способ построения функций F и G, присутствующих в определении равносильности (1). В качестве значения функции $F : A_0 \to B_0$ для $a_j \in A_0$ можно взять $F(a_j) = \pi(a_i)$, где a_i — произвольный символ из $\hat{A}_0$, мажорирующий в A символ a_j. Функция G строится аналогично.

До сих пор рассматривалась равносильность соответственных алфавитов. Обсудим теперь понятие равносильности недоопределённых алфавитов без заданного для них соответствия.

Будем использовать запись $A \sim B$ для обозначения того, что для недоопределённых алфавитов A и B существует соответствие R_{AB}, при котором $A \approx B$. В отличие от соотношения $A \approx B$, которое нельзя рассматривать как отношение на множестве алфавитов, ибо оно зависит также от соответствия R_{AB}, соотношение $A \sim B$ представляет собой отношение.

Утверждение 2. Отношение $A \sim B$ на множестве недоопределённых алфавитов является эквивалентностью.

Доказательство. Очевидно, что это отношение рефлексивно и симметрично. Докажем его транзитивность. Если имеют место равносильности $A \sim B$, $B \sim C$ и соотношения $A \approx B$, $B \approx C$ справедливы при соответствиях R_{AB}, R_{BC}, то по лемме 8 при $R_{AC} = R_{AB} \circ R_{BC}$ выполнено $A \approx C$, а потому $A \sim C$. ■

Алфавиты A и B, для которых $A \sim B$, будем называть *эквивалентными*. Следствием теоремы 3 является следующий факт.

Утверждение 3. Алфавиты A и B эквивалентны тогда и только тогда, когда алфавиты $\hat{A}$ и $\hat{B}$, полученные их приведением, изоморфны.

Алфавит A' называется *минимальным* для A, если $A' \sim A$ и A' имеет наименьшую мощность $|A'|$ и наименьшую мощность $|A'_0|$ основного алфавита среди всех алфавитов, эквивалентных A.

Утверждение 4. Алфавит $\hat{A}$, полученный из A приведением, минимален для A, и, таким образом, задача построения минимального алфавита решается эффективно.

Доказательство. Действительно, если имеется алфавит B, эквивалентный A и имеющий меньшую, чем у $\hat{A}$, мощность либо мощность основного алфавита, то это же будет справедливо для алфавита $\hat{B}$, полученного приведением B. В этом случае алфавит $\hat{B}$ не изоморфен $\hat{A}$, и по предыдущей лемме алфавит B не может быть эквивалентен A. ■

Преобразование недоопределённого алфавита называется *эквивалентным*, если в применении к любому алфавиту оно даёт эквивалентный алфавит. Одним из эквивалентных преобразований является исключение мажорируемого символа (лемма 10). Поскольку отношение эквивалентности симметрично, эквивалентным является и обратное преобразование — добавление мажорируемого символа. Более подробно эта операция состоит в следующем. К основному алфавиту A_0 добавляется новый символ a_s. Выбирается какой-либо символ $a_i \in A_0$, затем некоторым символам $a_T \in A$, таким, что $T \ni i$, сопоставляются и добавляются в алфавит A символы $a_{T \cup s}$. При добавлении $a_{T \cup s}$ символ a_T может быть оставлен в алфавите либо удалён из него. Эквивалентным преобразованием является также операция переименования символов. При её выполнении символы основного алфавита переименовываются некоторым образом (без отождествления) и символы $a_T = \{a_i : i \in T\}$ алфавита A заменяются на $\{\pi(a_i) : i \in T\}$, где $\pi(a_i)$ — результат переименования символа a_i. Система эквивалентных преобразований называется *полной*, если для любых двух эквивалентных алфавитов существует последовательность преобразований из этой системы, переводящая один алфавит в другой.

Утверждение 5. Операции исключения мажорируемого символа, добавления мажорируемого символа и переименования символов образуют полную систему эквивалентных преобразований недоопределённых алфавитов.

Доказательство. Действительно, если $A \sim B$, то приведённые алфавиты $\hat{A}$ и $\hat{B}$ изоморфны (утверждение 3). Поэтому из алфавита A можно устранением мажорируемых символов получить $\hat{A}$, переименованием символов преобразовать его в $\hat{B}$, а затем добавлением мажорируемых символов перейти к B. ■

Задача распознавания равносильности алфавитов решается эффективно (теорема 4), а решение задачи распознавания эквивалентности алфавитов связано с трудностями, поскольку к её частному случаю, когда символы недоопределённого алфавита имеют по два доопределения, сводится задача об изоморфизме графов [8], являющая-

ся одной из наиболее известных комбинаторных задач, безуспешные попытки решения которой продолжаются в течение нескольких десятков лет.

ЛИТЕРАТУРА

1. *Колмогоров А. Н.* Три подхода к определению понятия «количество информации» // Проблемы передачи информации. 1965. Т. 1. Вып. 1. С. 3–11.
2. *Шоломов Л. А.* Преобразование нечетких данных с сохранением информационных свойств // Дискретный анализ и исследование операций. 2005. Сер. 1. Т. 12. № 3. С. 85–104.
3. *Шоломов Л. А.* Разложение недоопределенных данных // Дискретный анализ и исследование операций. 2012. Т. 19. № 6. С. 72–98.
4. *Галлагер Р.* Теория информации и надежная связь. М.: Сов. радио, 1974. 720 с.
5. *Шоломов Л. А.* Элементы теории недоопредёленной информации // Прикладная дискретная математика. Приложение. 2009. № 2. С. 18–42.
6. *Шоломов Л. А.* О функционалах, характеризующих сложность систем недоопределенных булевых функций // Проблемы кибернетики. Вып. 19. М.: Наука, 1967. С. 123–139.
7. *Потапов В. Н.* Обзор методов неискажающего кодирования дискретных источников // Дискретный анализ и исследование операций. 1999. Сер. 1. Т. 6. № 4. С. 49–91.
8. *Гэри М., Джонсон Д.* Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982. 416 с.

ПРИКЛАДНАЯ ТЕОРИЯ ГРАФОВ

DOI 10.17223/20710410/25/5

УДК 519.1

АТТРАКТОРЫ В КОНЕЧНЫХ ДИНАМИЧЕСКИХ СИСТЕМАХ ДВОИЧНЫХ ВЕКТОРОВ, АССОЦИИРОВАННЫХ С ОРИЕНТАЦИЯМИ ПАЛЬМ

А. В. Жаркова

Саратовский государственный университет им. Н. Г. Чернышевского, г. Саратов, Россия

E-mail: VAnastasiyaV@gmail.com

Описываются аттракторы в конечных динамических системах двоичных векторов, ассоциированных с ориентациями пальм; определяется свойство принадлежности состояния аттрактору. Состояниями динамической системы являются все возможные ориентации данной пальмы, а эволюционная функция переориентирует все дуги, входящие в стоки.

Ключевые слова: *аттрактор, двоичный вектор, конечная динамическая система, пальма, сверхстройное (звездообразное) дерево.*

Введение

Графовые модели, в которых отказы процессоров интерпретируются как удаление соответствующих вершин, а отказы сетевых каналов — как удаление дуг, занимают важное место в задачах, связанных с отказоустойчивостью компьютерных сетей. Здесь можно выделить следующие три основные конструкции, получившие и самостоятельное значение в теории графов: минимальное расширение графа [1, 2], T-неприводимое расширение графа [3], бесконтурный граф с заданной структурой источников и стоков [4]. В модели [4] в качестве механизма восстановления работоспособности сети предлагается так называемая SER-динамика бесконтурных связных ориентированных графов. Это позволяет использовать при изучении модельных графов идеи и методы теории конечных динамических систем, и, в частности, динамических систем двоичных векторов (см., например, [5, 6]) — когда имеется естественная двоичная кодировка графов рассматриваемого класса. В указанных работах по отказоустойчивости графовых систем основные результаты получены для систем, в основе которых лежат цепи, циклы и частные типы деревьев. К числу деревьев, для которых найдено описание как минимальных, так и T-неприводимых расширений, относятся пальмы [2, 3]. Дерево называется *пальмой*, если оно является объединением цепей, имеющих общую концевую вершину, причём все эти цепи, за исключением, быть может, одной, имеют длину 1. Пальма является частным случаем *сверхстройного* (*звездообразного*) *дерева* (дерево, в котором в точности одна вершина имеет степень больше 2). В настоящей работе пальмы изучаются с точки зрения динамического подхода к отказоустойчивости графовых систем.

Под *конечной динамической системой* понимается пара $(S,\ \delta)$, где S — конечное непустое множество, элементы которого называются *состояниями системы*; $\delta : S \to S$ — отображение множества состояний в себя, называемое *эволюционной*

функцией системы. Таким образом, каждой конечной динамической системе сопоставляется карта, представляющая собой орграф с множеством вершин S и дугами, проведёнными из каждой вершины $s \in S$ в вершину $\delta(s)$. Компоненты связности графа, задающего динамическую систему, называются её *бассейнами*. Получается, что каждый бассейн представляет собой контур с входящими в него деревьями. Контуры, в свою очередь, называются предельными циклами, или *аттракторами*.

Основными проблемами теории конечных динамических систем являются задачи отыскания эволюционных параметров без проведения динамики. К их числу относятся *ветвление* (количество непосредственных предшественников данного состояния), свойство *недостижимости* состояния (состояние имеет нулевое ветвление), свойство принадлежности состояния аттрактору и описание аттракторов системы (их количество, вид и длина). Автором составлены программы для ЭВМ, позволяющие вычислять различные параметры динамических систем двоичных векторов, ассоциированных с некоторыми типами графов, в частности [7], и описаны аттракторы конечных динамических систем двоичных векторов, ассоциированных с ориентациями таких типов графов, как цепи и циклы [8 – 10].

В данной работе описываются аттракторы в конечных динамических системах двоичных векторов, ассоциированных с ориентациями пальм; определяется свойство принадлежности состояния аттрактору в таких системах.

1. Описание динамической системы

Пусть пальма p образована объединением цепей p_0, p_1, …, p_c, имеющих общую концевую вершину. Будем считать, что p_0 имеет среди этих цепей максимальную длину $s \geqslant 1$. Назовём p_0 *стволом пальмы* p, цепи p_1, p_2, …, p_c, имеющие длину 1, — её *листьями*, а их совокупность — *кроной*. Будем говорить, что p является пальмой типа (s, c). Пальма с точностью до изоморфизма определятся своим типом. При $c = 1$ пальма вырождается в цепь (см., например, [6, 8]), поэтому далее не будем рассматривать этот случай, считая $c > 1$.

Пусть имеется пальма p типа (s, c), $s + c = n$. Перенумеруем рёбра пальмы p, как показано на рис. 1.

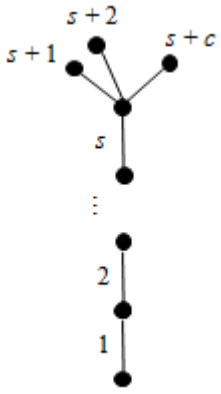


Рис. 1. Нумерация рёбер пальмы

Придадим каждому ребру пальмы p произвольную ориентацию и сопоставим полученному ориентированному графу n-мерный двоичный вектор $v(p)$, полагая его i-ю компоненту равной 1, если i-е ребро пальмы p ориентировано от корня (начальной вершины ствола), и 0 — в противном случае. Теперь можно последовательно выписать

получившуюся последовательность из нулей и единиц: $v = v_1 \dots v_s.v_{s+1} \dots v_{s+c}$, где v_i, $0 < i \leqslant s+c$, принимает значение 0 или 1 в зависимости от ориентации i-го ребра пальмы. Таким образом, каждой ориентации пальмы сопоставляется n-мерный двоичный вектор, причём $n = s+c$. В свою очередь, каждый такой вектор $v = v_1 \dots v_s.v_{s+1} \dots v_{s+c}$ однозначно определяет некоторую ориентацию пальмы $p(v)$ типа (s, c). Таким образом, между множеством P_{s+c}, $s > 0$, $c > 1$, всевозможных ориентированных пальм типа (s, c) и множеством B^{s+c}, $s > 0$, $c > 1$, всех двоичных векторов размерности $n = s + c$ устанавливается взаимно однозначное соответствие. В дальнейшем ориентации пальмы для простоты также будем называть пальмами, часть $v_1 \dots v_s$ вектора v — *стволом вектора* v, а $v_{s+1} \dots v_{s+c}$ — его *кроной*.

Опишем конечную динамическую систему ориентаций (s, c)-пальмы p на языке двоичных векторов. Пусть состоянием динамической системы в данный момент времени является вектор $v = v_1 \dots v_s v_{s+1} \dots v_{s+c} \in B^{s+c}$. Тогда в следующий момент времени она окажется в состоянии $\gamma(v) = v'$, полученном путём одновременного применения следующих правил:

I. Если $v_1 = 0$, то $v'_1 = 1$.

II. Если $v_i = 1$ и $v_{i+1} = 0$ для некоторого $0 < i < s$, то $v'_i = 0$ и $v'_{i+1} = 1$.

III. Если $v_i = 1$ для некоторого $s < i \leqslant s + c$, то $v'_i = 0$.

IV. Если $v_s = 1$ и $v_i = 0$ для всех $s < i \leqslant s + c$, то $v'_s = 0$ и $v'_i = 1$ для всех $s < i \leqslant s + c$.

V. Других отличий между v и $\gamma(v)$ нет.

Например, на рис. 2 показана эволюция вектора 011.11 в динамической системе (B^{3+2}, γ).

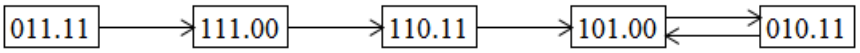


Рис. 2. Эволюция состояния 011.11 в динамической системе (B^{3+2}, γ)

Пусть теперь имеется n-рёберная (s, c)-пальма. На языке ориентаций пальм эволюция динамической системы вводится следующим образом: если дана некоторая ориентированная пальма $p \in P_{s+c}$, то её динамическим образом $\gamma(p)$ является пальма, получаемая из p одновременным превращением всех стоков в источники. Напомним, что *стоком* в ориентированном графе называется вершина с нулевой степенью исхода, а *источником* — вершина с нулевой степенью захода. Это частный случай динамики бесконтурных связных графов, введённой в [4]. Преобразования ориентаций пальм в динамической системе (P_{s+c}, γ), $s > 0$, $c > 1$, соответствуют эволюционным преобразованиям соотносимых им двоичных векторов в динамической системе (B^{s+c}, γ), $s > 0$, $c > 1$, и обратно, а именно $v(\gamma(p)) = \gamma(v(p))$ [11]. Таким образом, динамические системы (B^{s+c}, γ) и (P_{s+c}, γ), $s > 0$, $c > 1$, изоморфны. На рис. 3 и 4 изображены карты изоморфных динамических систем (B^{1+2}, γ) и (P_{1+2}, γ).

2. Недостижимые состояния динамической системы (B^{s+c}, γ)

В работе [12] рассмотрена динамическая система (Γ_G, α), где через Γ_G обозначено множество всех возможных ориентаций данного графа G, а эволюционная функция α задаётся следующим образом: если дан некоторый орграф $G \in \Gamma_G$, то его динамическим образом $\alpha(G)$ является орграф, полученный из G одновременной переориентацией всех дуг, входящих в стоки; других отличий между G и $\alpha(G)$ нет.

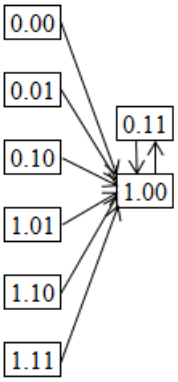


Рис. 3. Карта динамической системы (B^{1+2}, γ)

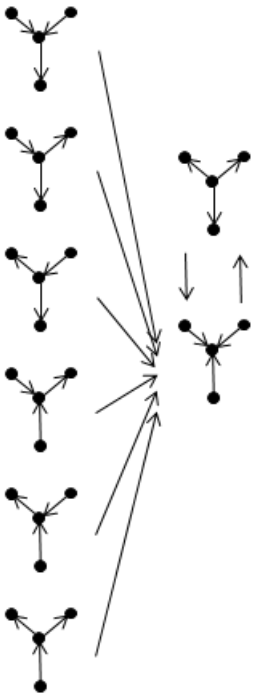


Рис. 4. Карта динамической системы (P_{1+2}, γ)

Множество источников ориентированного графа назовем *допустимым*, если из него в каждый сток этого графа есть дуга.

Теорема 1 [12]. Состояние s динамической системы (Γ_G, α) недостижимо тогда и только тогда, когда в орграфе G, представляющем состояние s, есть по крайней мере один сток и при этом нет ни одного допустимого множества источников, или, другими словами, когда существует хотя бы один сток в G, не смежный с источниками.

Рассматриваемая в данной работе динамическая система (P_{s+c}, γ) является частным случаем системы (Γ_G, α), поэтому теорема 1 применима и к ней. Из теоремы 1 можно выразить свойство недостижимости состояний динамической системы (B^{s+c}, γ), $s > 0$, $c > 1$, на языке двоичных векторов.

Следствие 1. Состояние $v = v_1 \dots v_s . v_{s+1} \dots v_{s+c}$ динамической системы (B^{s+c}, γ), $s > 0$, $c > 1$, недостижимо тогда и только тогда, когда выполняется хотя бы одно из следующих условий:

1) $v_1v_2 = 00$;

2) $v_iv_{i+1}v_{i+2}v_{i+3} = 1100$ для некоторого $0 < i < s - 1$;

3) среди последних c компонент имеются различные;

4) $v_s = v_{s+1} = \ldots = v_{s+c} = 1$.

Доказательство. Необходимость. Пусть состояние $v = v_1 \ldots v_s.v_{s+1} \ldots v_{s+c}$ динамической системы (B^{s+c}, γ), $s > 0$, $c > 1$, является недостижимым. Покажем, что для него выполняется хотя бы одно из условий 1–4.

Так как состояние v является недостижимым, то по теореме 1 в соответствующей ему ориентации пальмы $p(v)$ существует хотя бы один сток, не смежный с источниками. Вершины пальмы обозначим за $p_v^1, p_v^2, \ldots, p_v^{s+c+1}$ (нумерация от корня). Пусть, не теряя общности, в $p(v)$ существует единственный сток, не смежный с источниками; рассмотрим ситуации в зависимости от расположения данного стока в пальме.

I. Данный сток находится в вершине p_v^1 пальмы $p(v)$, тогда $v_1v_2 = 00$. Обратим внимание, что при $s = 1$ данное условие также выполняется, так как за счёт $v_2 = 0$ вершина p_v^2 не является источником.

II. Данный сток находится в вершине p_v^2 пальмы $p(v)$. Такая ситуация невозможна, так как p_v^1 в таком случае является источником, а значит, сток в вершине p_v^2 смежен с источником.

III. Данный сток находится в одной из вершин p_v^j, $2 < j < s$, тогда $v_iv_{i+1}v_{i+2}v_{i+3} = 1100$ для $i = j - 2$.

IV. Данный сток находится в вершине p_v^s, тогда $v_{s-2}v_{s-1}v_s = 110$ и среди компонент $v_{s+1}, v_{s+2}, \ldots, v_{s+c}$ обязательно есть нуль.

V. Данный сток находится в вершине p_v^{s+1}. Такая ситуация невозможна, так как в таком случае $v_{s+1} = v_{s+2} = \ldots = v_{s+c} = 0$, а значит, сток в вершине p_v^{s+1} смежен с источниками.

VI. Данный сток находится в вершине p_v^j для некоторого j, $s + 1 < j \leqslant s + c + 1$.

а) При $v_s = 0$ среди компонент v_k, $k \neq j, s + 1 < k \leqslant s + c + 1$, обязательно есть нуль.

б) При $v_s = 1$ дополнительных условий не требуется, так как p_v^{s+1} не является источником.

Таким образом, для недостижимого состояния $v = v_1 \ldots v_s.v_{s+1} \ldots v_{s+c}$ динамической системы (B^{s+c}, γ), $s > 0$, $c > 1$, выполняется одно из следующих условий:

1) $v_1v_2 = 00$ (за счёт пункта I);

2) $v_iv_{i+1}v_{i+2}v_{i+3} = 1100$ для некоторого i, $0 < i < s - 1$ (за счёт пунктов III и IV; для пункта IV заметим, что если $v_{s+1} = 1$, то среди компонент $v_{s+2}, v_{s+3}, \ldots, v_{s+c}$ обязательно есть нуль, то есть имеем ситуацию, когда среди последних c компонент имеются различные);

3) среди последних c компонент имеются различные (за счёт пункта VI; заметим, что в данном случае не учтена ситуация, когда $v_s = v_{s+1} = \ldots = v_{s+c} = 1$);

4) $v_s = v_{s+1} = \ldots = v_{s+c} = 1$ (за счёт пункта VI, *б*).

Достаточность. Пусть для состояния $v = v_1 \ldots v_s.v_{s+1} \ldots v_{s+c}$ динамической системы (B^{s+c}, γ), $s > 0$, $c > 1$, выполняется хотя бы одно из условий 1–4. Для такого состояния v очевидно наличие в $p(v)$ стока, не смежного с источниками; тогда по теореме 1 данное состояние является недостижимым. ■

3. Аттракторы динамической системы (B^{s+c}, γ)

Через $p_{\mathrm{st}}(v)$ обозначим количество пар совпадающих соседних компонент в стволе вектора v; назовём данную величину *плотностью ствола вектора* v. По анало-

гии через $p_{\mathrm{cr}}(v)$ обозначим *плотность кроны вектора* v. Под *блоком* будем понимать максимальное по включению множество подряд стоящих нулей (0-блок) или единиц (1-блок) в количестве > 1. *Длина блока* — число нулей (единиц) в блоке, уменьшенное на 1. Обозначим через $p^0_{\mathrm{st}}(v)$ ($p^0_{\mathrm{cr}}(v)$), $p^1_{\mathrm{st}}(v)$ ($p^1_{\mathrm{cr}}(v)$) суммы длин 0-блоков и 1-блоков соответственно в стволе (кроне) вектора v.

Введём необходимые обозначения:

— x^k — в состоянии компонента x повторяется $k \geqslant 0$ раз;
— $(x)^k$ — в состоянии совокупность компонент x повторяется $k \geqslant 0$ раз;
— $\{01\}^k$ — произвольный набор из 0 и 1 размера $k \geqslant 0$;
— $\{0[x]1\}^k$ — произвольный набор из 0 и 1 размера $k > 0$, содержащий в себе обязательный элемент x;
— $v \to \gamma(v)$ — один шаг выполнения динамики системы.

Теорема 2. Динамическая система (B^{s+c}, γ), $s > 0$, $c > 1$, имеет единственный бассейн и аттрактор, представляющий собой двухэлементный контур, образуемый состояниями $(01)^{(s-1)/2}0.1^c$ и $(10)^{(s-1)/2}1.0^c$ при нечётном s и состояниями $(01)^{s/2}.0^c$ и $(10)^{s/2}.1^c$ при чётном s.

Доказательство. Рассмотрим состояния динамической системы (B^{s+c}, γ), $n = s + c$, $s > 0$, $c > 1$, в зависимости от наличия и количества 0- и 1-блоков в стволе соответствующего вектора v.

I. Рассмотрим состояния, в стволе которых нет ни 0-, ни 1-блоков, то есть $p_{\mathrm{st}}(v) = 0$.

1) Пусть s — нечётное. Тогда имеем

а) $(01)^{(s-1)/2}0.1^c \to (10)^{(s-1)/2}1.0^c$, при этом $(10)^{(s-1)/2}1.0^c \to (01)^{(s-1)/2}0.1^c$. Таким образом, данные состояния образуют аттрактор длины 2;

б) $(01)^{(s-1)/2}0.\{0[0]1\}^c \to (10)^{(s-1)/2}1.0^c$, то есть приходит в состояние из п. I $(1,a)$, которое принадлежит аттрактору;

в) $(10)^{(s-1)/2}1.\{0[1]1\}^c \xrightarrow{1} (01)^{(s-1)/2}1.0^c \xrightarrow{2} (10)^{(s-3)/2}110.1^c \xrightarrow{3} \cdots \xrightarrow{s-2} 011(01)^{(s-3)/2}.0^c \xrightarrow{s-1} 110(10)^{(s-3)/2}.1^c \xrightarrow{s} (10)^{(s-1)/2}1.0^c$, то есть в итоге приходит в состояние из п. I $(1,a)$, которое принадлежит аттрактору.

2) Пусть s — чётное. Тогда имеем

а) $(01)^{s/2}.0^c \to (10)^{s/2}.1^c$, при этом $(10)^{s/2}.1^c \to (01)^{s/2}.0^c$. Таким образом, данные состояния образуют аттрактор длины 2;

б) $(01)^{s/2}.\{0[1]1\}^c \xrightarrow{1} 1(01)^{(s-2)/2}1.0^c \xrightarrow{2} (01)^{(s-2)/2}10.1^c \xrightarrow{3} \cdots \xrightarrow{s-2} 01(10)^{(s-2)/2}.1^c \xrightarrow{s-1} 11(01)^{(s-2)/2}.0^c \xrightarrow{s} (10)^{s/2}.1^c$, то есть в итоге приходит в состояние из п. I $(2,a)$, которое принадлежит аттрактору;

в) $(10)^{s/2}.\{0[0]1\}^c \to (01)^{s/2}.0^c$, то есть приходит в состояние из п. I $(2,a)$, которое принадлежит аттрактору.

Теперь рассмотрим состояния, в стволе которых присутствуют 0- или 1-блоки. Исключим пока из рассмотрения состояния вида $01^{s-1}.\{0[1]1\}^c$ и $\{01\}^t101^{s-t-2}.\{0[1]1\}^c$, где $s > 2$, $c > 1$, $t \geqslant 0$, и рассмотрим их в самом конце доказательства (п. VII). Согласно следствию 1, данные состояния являются недостижимыми, значит, не могут получиться из какого-либо другого состояния в процессе эволюции, поэтому данное ограничение не повлияет на дальнейшее рассмотрение пп. II–VI.

II. Рассмотрим состояния, в стволе которых нет 0-блоков и есть хотя бы один 1-блок, т. е. $p^0_{\mathrm{st}}(v) = 0$, $p^1_{\mathrm{st}}(v) > 0$. Рассмотрим эволюцию таких состояний в зависимости от количества 1-блоков.

1) Вектор имеет единственный 1-блок. Рассмотрим эволюцию таких состояний в зависимости от расположения 1-блока:

а) $p^1_{\mathrm{st}}(v) = s - 1$;

а′) $p^0_{\mathrm{cr}}(v) = c - 1$. Тогда эволюция состояния выглядит следующим образом: $1^s.0^c \stackrel{1}{\to} 1^{s-1}0.1^c \stackrel{2}{\to} 1^{s-2}01.0^c \stackrel{3}{\to} \cdots$;

а′1) s — нечётное: $\cdots \stackrel{s-2}{\to} 11(01)^{(s-3)/2}0.1^c \stackrel{s-1}{\to} (10)^{(s-1)/2}1.0^c$, то есть в итоге приходит в состояние из п. I $(1,a)$, которое принадлежит аттрактору;

а′2) s — чётное: $\cdots \stackrel{s-2}{\to} 11(01)^{(s-2)/2}.0^c \stackrel{s-1}{\to} (10)^{s/2}.1^c$, то есть в итоге приходит в состояние из п. I $(2,a)$, которое принадлежит аттрактору.

Заметим, что длина 1-блока в стволе на каждом очередном шаге уменьшается на единицу за счёт поглощения самой левой компоненты, и за ним следуют чередующиеся нули и единицы (самая левая компонента 1-блока на каждом очередном шаге переходит в нуль по правилу II или IV, тем самым длина блока уменьшается на единицу), что продолжается до тех пор, пока все компоненты блока, кроме последней, не поглотятся. Тем самым состояние в итоге приходит в состояние, в стволе которого нет ни 0-, ни 1-блоков, то есть в состояние из п. I;

а″) $p^0_{\mathrm{cr}}(v) < c - 1$. Тогда имеем $1^s.\{0[1]1\}^c \to 1^s.0^c$, то есть данное состояние переходит в состояние из п. II $(1,a,a')$;

б) вектор v содержит единственный 1-блок в стволе и при этом $p^1_{\mathrm{st}}(v) < s - 1$.

При эволюции в стволе 1-блок на каждом очередном шаге смещается влево на одну компоненту (с каждым очередным шагом эволюции нуль, стоящий перед 1-блоком, переходит в единицу по правилу I или II, а последняя единица 1-блока переходит в нуль по правилу II или IV, тем самым длина 1-блока сохраняется), пока не встаёт в начало состояния, затем длина 1-блока на каждом очередном шаге уменьшается на единицу за счёт поглощения самой левой компоненты, и за ним следуют чередующиеся нули и единицы, что продолжается до тех пор, пока все его компоненты, кроме последней, не поглотятся (см. п. I $(1,a)$). Тем самым состояние в итоге приходит в состояние, в стволе которого нет ни 0-, ни 1-блоков, то есть в состояние из п. I.

2) Вектор имеет несколько 1-блоков в стволе.

Из п. II (1) имеем: при эволюции в стволе каждый 1-блок на очередном шаге смещается влево на одну компоненту, пока не встанет в начало вектора, затем длина 1-блока на каждом очередном шаге уменьшается на единицу за счёт поглощения самой левой компоненты, и за ним следуют чередующиеся нули и единицы, что продолжается до тех пор, пока все компоненты 1-блока не поглотятся. Таким образом, в данном случае последними поглотятся все компоненты, кроме последней, самого правого 1-блока, то есть такое состояние при эволюции в итоге приходит в состояние, в стволе которого нет ни 0-, ни 1-блоков, то есть в состояние из п. I.

III. Рассмотрим состояния, в стволе которых нет 1-блоков и есть хотя бы один 0-блок, то есть $p^0_{\mathrm{st}}(v) > 0$, $p^1_{\mathrm{st}}(v) = 0$. Рассмотрим эволюцию таких состояний в зависимости от количества 0-блоков в стволе.

1) Вектор имеет единственный 0-блок в стволе. Рассмотрим эволюцию таких состояний в зависимости от расположения 0-блока в стволе.

а) $p^0_{\mathrm{st}}(v) = s - 1$. Тогда эволюция состояния выглядит следующим образом: $0^s.\{01\}^c \stackrel{1}{\to} 10^{s-1}.0^c \stackrel{2}{\to} 010^{s-2}.0^c \stackrel{3}{\to} \cdots$;

а1) s — нечётное: $\cdots \stackrel{s-2}{\to} (10)^{(s-1)/2}0.0^c \stackrel{s-1}{\to} (01)^{(s-1)/2}0.0^c$, то есть приходит в состояние из п. I;

а2) s — чётное: $\cdots \stackrel{s-2}{\to} (01)^{(s-2)/2}00.0^c \stackrel{s-1}{\to} (10)^{s/2}.0^c$, то есть приходит в состояние из п. I.

Заметим, что в стволе длина 0-блока на каждом очередном шаге эволюции уменьшается на единицу за счёт поглощения самой правой компоненты, и перед ним идут чередующиеся нули и единицы (на каждом очередном шаге эволюции самая первая компонента 0-блока переходит в единицу по правилу I или II, тем самым длина 0-блока уменьшается на единицу), что продолжается до тех пор, пока все его компоненты, кроме первой, не поглотятся. Тем самым состояние в итоге приходит в состояние, в стволе которого нет ни 0-, ни 1-блоков, то есть в состояние из п. I;

б) вектор содержит единственный 0-блок в стволе и при этом $p^0_{\mathrm{st}}(v) < s - 1$.

При эволюции в стволе 0-блок на каждом очередном шаге смещается вправо на одну компоненту (с каждым очередным шагом эволюции первый ноль 0-блока переходит в единицу по правилу I или II; единица, стоящая после 0-блока, переходит в ноль по правилу II или IV, тем самым длина 0-блока сохраняется), пока не встаёт в конец состояния, затем длина 0-блока на каждом очередном шаге уменьшается на единицу за счёт поглощения самой правой компоненты, и перед ним идут чередующиеся нули и единицы, что продолжается до тех пор, пока все его компоненты, кроме первой, не поглотятся (см. п. III $(1,a)$). Тем самым состояние в итоге приходит в состояние, в стволе которого нет ни 0-, ни 1-блоков, то есть в состояние из п. I.

2) Вектор имеет несколько 0-блоков в стволе.

Из п. III (1) имеем: при эволюции в стволе каждый 0-блок на очередном шаге смещается вправо на одну компоненту, пока не достигает конца ствола вектора, затем длина 0-блока на каждом очередном шаге уменьшается на единицу за счёт поглощения самой правой компоненты, и перед ним следуют чередующиеся нули и единицы, что продолжается до тех пор, пока все компоненты 0-блока не поглотятся. Таким образом, в данном случае последними поглотятся все компоненты, кроме первой, самого левого 0-блока, то есть такое состояние при эволюции в итоге приходит в состояние, в стволе которого нет ни 0-, ни 1-блоков, то есть в состояние из п. I.

IV. Вектор содержит в себе 1-блоки, после которых идут 0-блоки.

Из предыдущих пунктов имеем, что в данном случае на каждом шаге эволюции одновременно 0- и 1-блоки начнут движение вправо и влево соответственно, пока не достигнут своих концов вектора и не начнут поглощаться, в результате чего останется вектор, не содержащий в своём стволе ни 0-, ни 1-блоков, то есть данное состояние на очередном шаге эволюции приходит в состояние из п. I.

V. В стволе вектора присутствуют 0-блоки, после которых идут 1-блоки.

1. Рассмотрим сначала эволюцию данных состояний, имеющих в своем стволе по одному 0- и 1-блоку, в зависимости от их значений $p^0_{\mathrm{st}}(v)$ и $p^1_{\mathrm{st}}(v)$.

Если между 0- и 1-блоком стоят чередующиеся нули и единицы, то на каждом шаге эволюции в стволе одновременно 0-блок начнёт движение вправо на одну компоненту, 1-блок — влево за счёт поглощения компонент между блоками (первая компонента 0-блока при эволюции заменится на единицу по правилу I или IV; единица, следующая за 0-блоком, при эволюции заменится на нуль по правилу II, таким образом длина 0-блока сохранится, аналогичное происходит и с 1-блоком), пока они не окажутся стоящими рядом. Рассмотрим эволюцию уже с этого шага в зависимости от сумм длин 0- и 1-блоков в стволе.

а) Состояния, для которых $p^0_{\mathrm{st}}(v) < p^1_{\mathrm{st}}(v)$.

Когда в стволе 0-блок и 1-блок оказываются стоящими рядом, то с каждым следующим шагом эволюции их длины начинают уменьшаться у каждого на единицу за счёт поглощения компонент друг друга (у 0-блока первая компонента перейдёт в единицу по правилу эволюции I или II, а у 1-блока последняя компонента перейдёт в нуль по

правилу эволюции II или IV, тем самым длины 0- и 1-блоков уменьшаются на единицу), пока 0-блок полностью не поглотится. В результате в стволе будет присутствовать 1-блок и чередующиеся нули и единицы, дальнейшая эволюция такого состояния описана в п. II (1).

б) Состояния, для которых $p^0_{\mathrm{st}} = p^1_{\mathrm{st}}$.

Из рассуждений п. V $(1,a)$ получаем, что блоки будут поглощать друг друга с каждым следующим шагом эволюции, пока от них не останется по одной компоненте, причем это уже будет вектор, ствол которого не содержит ни 0-, ни 1-блоков, а про его эволюцию сказано в п. I.

в) Состояния, для которых $p^0_{\mathrm{st}}(v) > p^1_{\mathrm{st}}(v)$.

Ситуация аналогична п. V $(1,a)$, только в итоге в стволе состояния остаётся один 0-блок и чередующиеся нули и единицы, дальнейшая эволюция описана в п. III (1).

2) Рассмотрим теперь состояния, включающие в себя несколько 0-блоков, после которых идут 1-блоки, также в зависимости от значений $p^0_{\mathrm{st}}(v)$ и $p^1_{\mathrm{st}}(v)$.

а) Состояния, для которых $p^0_{\mathrm{st}}(v) < p^1_{\mathrm{st}}(v)$.

Из п. V (1) получаем, что в стволе таких состояний на каждом шаге эволюции 0-блоки начинают двигаться вправо, 1-блоки — влево за счёт поглощения компонент, стоящих между блоками; когда они оказываются стоящими рядом, то блоки начинают уменьшаться на единицу каждый за счёт поглощения компонент друг друга, пока один из блоков полностью не поглотится, после чего опять продолжаются сдвиги блоков навстречу друг другу, пока не поглотится самый последний 0-блок (кроме его первой компоненты). Таким образом, в состоянии останутся только 1-блоки и чередующиеся нули и единицы, дальнейшая эволюция описана в п. II.

б) Состояния, для которых $p^0_{\mathrm{st}}(v) = p^1_{\mathrm{st}}(v)$.

Из рассуждений пп. V (1) и V $(2,a)$ получаем, что при эволюции 0-блоки и 1-блоки начнут движение навстречу друг другу, пока не окажутся рядом, а затем их длины начнут уменьшаться на единицу с каждым следующим шагом эволюции, пока не поглотится один из блоков, после чего продолжится аналогичное движение, пока рядом не окажутся последние оставшиеся 0-блок и 1-блок, длины которых начнут уменьшаться на единицу с каждым следующим шагом эволюции, пока от них не останется по одной компоненте, причём это уже будет вектор, ствол которого не содержит ни 0-, ни 1-блоков, эволюция которого описана в п. I.

в) Состояния, для которых $p^0_{\mathrm{st}}(v) > p^1_{\mathrm{st}}(v)$.

Ситуация аналогична п. V $(2,a)$, только в состоянии в итоге остаются одни 0-блоки и чередующиеся нули и единицы, дальнейшая эволюция состояния описана в п. III.

VI. Ствол вектора содержит 0-блоки и 1-блоки в произвольном порядке.

Из рассуждений предыдущих пунктов получаем, что при эволюции такого состояния в его стволе 0-блоки будут сдвигаться вправо, при этом если 0-блок встречается с 1-блоком, то его длина уменьшается с очередным шагом эволюции, то есть если есть подряд стоящие 0-блоки, то они или все поглотятся, если следующие за ними подряд стоящие 1-блоки в сумме имеют равную или большую длину, или поглотят сами следующие за ними подряд стоящие 1-блоки и продолжат сдвиг вправо, встречая очередные 1-блоки, если сумма длин 0-блоков больше суммы длин 1-блоков. Для рассмотрения 1-блоков ситуация получается аналогичная. В итоге получим состояние, имеющее только 0-блоки, только 1-блоки, или состояние, не имеющее ни 0-, ни 1-блоков, чьи эволюции описаны в пп. I–III.

VII. Теперь рассмотрим состояния вида $01^{s-1}.\{0[1]1\}^c$ и $\{01\}^t 101^{s-t-2}.\{0[1]1\}^c$, где $s > 2$, $c > 1$, $t \geqslant 0$.

1) $01^{s-1}.\{0[1]1\}^c \to 1^s.0^c$; таким образом, данное состояние при эволюции переходит в состояние из п. II $(1,a)$.

2) $\{01\}^t 101^{s-t-2}.\{0[1]1\}^c \to \{01\}^t 01^{s-t-1}.0^c$; таким образом, далее нужно рассматривать получившееся состояние, эволюция которого описана в пп. II, V или VI.

Рассмотрев все возможные ситуации, заключаем, что динамическая система (B^{s+c},γ), $s>0$, $c>1$, имеет единственный бассейн и аттрактор, представляющий собой двухэлементный контур, образуемый состояниями $(01)^{(s-1)/2}0.1^c$ и $(10)^{(s-1)/2}1.0^c$ при нечётном s и состояниями $(01)^{s/2}.0^c$ и $(10)^{s/2}.1^c$ при чётном s. ■

Следствие 2. В динамической системе $(B^{s+c},\ \gamma)$, $s>0$, $c>1$, состояния $(01)^{(s-1)/2}0.1^c$ и $(10)^{(s-1)/2}1.0^c$ при нечётном s и состояния $(01)^{s/2}.0^c$ и $(10)^{s/2}.1^c$ при чётном s, и только они принадлежат аттрактору.

Заключение

В работе рассмотрены конечные динамические системы двоичных векторов, ассоциированных с ориентациями пальм; показано свойство недостижимости состояния данной системы на языке двоичных векторов; описаны аттракторы систем, их количество, вид и длина; определено свойство принадлежности состояния аттрактору.

ЛИТЕРАТУРА

1. *Hayes J. P.* A graph model for fault-tolerant computing system // IEEE Trans. Comput. 1976. V. C25. No. 9. P. 875–884.
2. *Абросимов М. Б.* Графовые модели отказоустойчивости. Саратов: Изд-во Сарат. ун-та, 2012. 192 с.
3. *Курносова С. Г.* Т-неприводимые расширения для некоторых классов графов // Теоретические проблемы информатики и её приложений. 2004. Вып. 6. С. 113–125.
4. *Barbosa V. C.* An Atlas of Edge-Reversal Dynamics. Boca Raton: Chapman&Hall/CRC, 2001. 385 p.
5. *Colon-Reyes O., Laubenbacher R., and Pareigis B.* Boolean monomial dynamical systems // Ann. Combinator. 2004. V. 8. P. 425–439.
6. *Салий В. Н.* Об одном классе конечных динамических систем // Вестник Томского государственного университета. Приложение. 2005. № 14. С. 23–26.
7. *Власова А. В.* Исследование эволюционных параметров в динамических системах двоичных векторов // Свидетельство о государственной регистрации программы для ЭВМ № 2009614409, выданное Роспатентом. Зарегистрировано в Реестре программ для ЭВМ 20 августа 2009 г.
8. *Власова А. В.* Аттракторы в динамической системе (B,δ) двоичных векторов // Компьютерные науки и информационные технологии: материалы науч. конф. Саратов: Изд-во Сарат. ун-та, 2010. С. 35–41.
9. *Власова А. В.* Аттракторы динамических систем, ассоциированных с циклами // Прикладная дискретная математика. 2011. № 2 (12). С. 90–95.
10. *Жаркова А. В.* Количество аттракторов в динамических системах, ассоциированных с циклами // Матем. заметки. 2014. Т. 95. Вып. 4. С. 529–537.
11. *Власова А. В.* Динамические системы, определяемые пальмами // Компьютерные науки и информационные технологии: материалы Междунар. науч. конф. Саратов: Изд-во Сарат. ун-та, 2009. С. 57–60.
12. *Жаркова А. В.* О ветвлении и непосредственных предшественниках состояний в конечной динамической системе всех возможных ориентаций графа // Прикладная дискретная математика. Приложение. 2013. № 6. С. 76–78.

DOI 10.17223/20710410/25/6

УДК 519.6

ЛОКАЛЬНАЯ ПРИМИТИВНОСТЬ ГРАФОВ И НЕОТРИЦАТЕЛЬНЫХ МАТРИЦ

С. Н. Кяжин*, В. М. Фомичев*,**

**Национальный исследовательский ядерный университет «МИФИ», г. Москва, Россия*

*** Финансовый университет при Правительстве Российской Федерации, г. Москва, Россия*

E-mail: s.kyazhin@kaf42.ru, fomichev@nm.ru

Для ряда объектов, моделируемых неотрицательными матрицами (графами), важные свойства достигаются тогда, когда положительны их подматрицы (подграфы являются полными). В связи с этим в данной работе известные понятия примитивности и экспонента матрицы (графа) обобщаются до понятий локальной примитивности, квазипримитивности и локальных экспонентов матрицы (графа). Получены условия локальной примитивности, субпримитивности и квазипримитивности орграфа. Установлена связь экспонента матрицы (орграфа) с локальными экспонентами.

Ключевые слова: *экспонент, локальный экспонент, локальный субэкспонент, локальный квазиэкспонент, примитивная матрица, локальная примитивность.*

Введение

В ряде прикладных задач для изучения коммуникативных свойств системы объектов представляет интерес определение экспонента примитивной неотрицательной 0,1-матрицы или системы 0,1-матриц, кодирующих связи между объектами системы. Экспонентам неотрицательных матриц и систем матриц посвящено множество трудов. Обзор результатов в этом направлении, полученных до 2012 г., можно найти в [1].

Вместе с тем в некоторых приложениях важные свойства объектов, моделируемых неотрицательными матрицами, достигаются тогда, когда положительной является не вся матрица, а лишь некоторая её часть, например подматрица, получаемая вычеркиванием некоторых строк и столбцов. Такая ситуация имеет место при изучении, в частности, композиций преобразований векторного пространства, составляющих полугруппу или группу преобразований состояний генератора гаммы: во многих случаях достаточно, чтобы от всех битов начального состояния генератора существенно зависела лишь выделенная часть битов промежуточных состояний. В связи с этим в рамках матрично-графового подхода к изучению коммуникативных свойств объектов в данной работе в порядке обобщения понятий примитивности и экспонента введены и исследованы понятия локальной примитивности, субпримитивности и квазипримитивности, а также локальных экспонентов, субэкспонентов и квазиэкспонентов неотрицательных матриц и графов. Начальные результаты в этом направлении представлены в [2].

Основные обозначения:

$\mathbb{N}$ — множество натуральных чисел;

$N_n = \{1, \ldots, n\}$, где $n \in \mathbb{N}$;

Ω_n — множество всех непустых подмножеств множества N_n;

$M_0(n \times m)$ — множество 0,1-матриц размера $n \times m$;

$M_0(n) = M_0(n \times m)$ при $n = m$;

$A = (a_{i,j})$, $B = (b_{i,j})$, где $A, B \in M_0(n)$;
$A^t = (a_{i,j}^{(t)})$, $t \in \mathbb{N}$;
$\nu(A)$ — носитель неотрицательной матрицы A;
$S(n)$ — группа всех подстановочных матриц порядка n.

1. Локальная примитивность матриц

Пусть $n, m \in \mathbb{N}$, $A \in M_0(n \times m)$, $I = \{i_1, \ldots, i_k\}$, $J = \{j_1, \ldots, j_r\}$, $\varnothing \neq I \subseteq N_n$, $\varnothing \neq J \subseteq N_m$. Рассмотрим подматрицу $A(I \times J)$ размера $k \times r$, $0 < k \leqslant n$, $0 < r \leqslant m$, полученную из A вычёркиванием строк с номерами $i \neq i_1, \ldots, i_k$ и столбцов с номерами $j \neq j_1, \ldots, j_r$. Матрицу $A(I \times J)$ при $I = J$ обозначим $A(J^2)$, при $I = N_n$ обозначим $A(*J)$ и при $J = N_m$ — $A(I*)$.

Матрица A называется $I \times J$-положительной (J^2-положительной при $I = J$, $*J$-положительной при $I = N_n$, $I*$-положительной при $J = N_m$), если положительна матрица $A(I \times J)$ (матрица $A(J^2)$, матрица $A(*J)$, матрица $A(I*)$). Множество $I \times J$-положительных (J^2-положительных, $*J$-положительных, $I*$-положительных) матриц обозначим $M_+(I \times J)$ ($M_+(J^2)$, $M_+(*J)$, $M_+(I*)$).

Матрица A называется:

— s-положительной, если не содержит нулевых строк;
— c-положительной, если не содержит нулевых столбцов;
— sc-положительной, если не содержит нулевых строк и столбцов.

Матрица A называется $I \times J$-s-положительной (J^2-s-положительной при $I = J$, $*J$-s-положительной при $I = N_n$, $I*$-s-положительной при $J = N_m$), если s-положительной является матрица $A(I \times J)$ (матрица $A(J^2)$, матрица $A(*J)$, матрица $A(I*)$). Множество s-положительных ($I \times J$-s-положительных, J^2-s-положительных, $*J$-s-положительных, $I*$-s-положительных) матриц обозначим $Q_s(n \times m)$ ($Q_s(I \times J)$, $Q_s(J^2)$, $Q_s(*J)$, $Q_s(I*)$). Аналогично определяются $I \times J$-c-положительные и $I \times J$-sc-положительные матрицы. Соответствующие множества матриц обозначим $Q_c(n \times m)$, $Q_c(I \times J)$, $Q_c(J^2)$, $Q_c(*J)$, $Q_c(I*)$ и $Q_{sc}(n \times m)$, $Q_{sc}(I \times J)$, $Q_{sc}(J^2)$, $Q_{sc}(*J)$, $Q_{sc}(I*)$.

Далее считаем $n = m > 1$, $A \in M_0(n)$. Рассмотрим $M_0(n)$ как кольцо относительно операций сложения и умножения (обозначаемых $\pm$ и $\circ$ соответственно), где $A \pm B = \nu(A + B)$, $A \circ B = \nu(AB)$, то есть $A \pm B$ и $A \circ B$ суть 0,1-матрицы, полученные из матриц $A + B$ и AB соответственно заменой положительных элементов единицами.

Рассмотрим свойства типа $I \times J$-положительности для степеней квадратной матрицы. Матрица A называется $I \times J$-примитивной (J^2-примитивной при $I = J$, $*J$-примитивной при $I = N_n$, $I*$-примитивной при $J = N_n$), если существует натуральное число γ, такое, что матрица $A^t(I \times J)$ (матрица $A^t(J^2)$, матрица $A^t(*J)$, матрица $A^t(I*)$) положительна при любом $t \geqslant \gamma$. Наименьшее такое число γ назовём $I \times J$-экспонентом (J^2-экспонентом при $I = J$, $*J$-экспонентом при $I = N_n$, $I*$-экспонентом при $J = N_n$) матрицы A, обозначим $I \times J$-exp$A$ (J^2-exp$A$, $*J$-exp$A$, $I*$-exp$A$). Множество примитивных ($I \times J$-примитивных, J^2-примитивных, $*J$-примитивных, $I*$-примитивных) матриц обозначим $P(n)$ ($P(I \times J)$, $P(J^2)$, $P(*J)$, $P(I*)$).

Обозначим $S(J)$ подгруппу группы $S(n)$, определяемую условием: если $i \notin J$, то в любой матрице из $S(J)$ единица в i-й строке расположена на главной диагонали. Иначе говоря, при умножении матрицы из подгруппы $S(J)$ на любой вектор координаты вектора с номерами $i \notin J$ остаются неизменными, а остальные координаты могут быть переставлены. Группа $S(J)$ изоморфна группе $S(|J|)$. Заметим, что подстановоч-

ные матрицы не являются $I \times J$-примитивными, если хотя бы одно из множеств I, J имеет порядок больше единицы.

Отметим некоторые алгебраические свойства рассматриваемых множеств матриц.

Утверждение 1. При любых допустимых множествах I, J выполнено:

а) если $A \in Q_s(I^2) \cup Q_c(J^2)$ и γ — наименьшее натуральное число, при котором $A^\gamma(I \times J) > 0$, то $A \in P(I \times J)$ и $I \times J\text{-exp}A = \gamma$;

б) $Q_s(J^2)$ и $Q_c(J^2)$ — мультипликативные моноиды, содержащие подгруппу $S(J)$;

в) $P(J^2)$ — наследственное подмножество множества $Q_{sc}(n)$.

Доказательство.

а) Если $A^t(I \times J) > 0$, то $a_{i,j}^{(t)} > 0$ при любых $i \in I$ и $j \in J$, $t \geqslant 1$. По правилу умножения матриц

$$a_{i,j}^{(t+1)} = a_{i,1}a_{1,j}^{(t)} + \ldots + a_{i,n}a_{n,j}^{(t)}; \tag{1}$$

$$a_{i,j}^{(t+1)} = a_{i,1}^{(t)}a_{1,j} + \ldots + a_{i,n}^{(t)}a_{n,j}. \tag{2}$$

Если $A \in Q_s(I^2)$ и $i \in I$, то во множестве $\{a_{i,r} : r \in I\}$ содержится положительное число, тогда $a_{i,j}^{(t+1)} > 0$ в соответствии с (1). Если $A \in Q_c(J^2)$ и $j \in J$, то во множестве $\{a_{r,j} : r \in J\}$ содержится положительное число, тогда $a_{i,j}^{(t+1)} > 0$ в соответствии с (2). Следовательно, в обоих случаях $A^{t+1}(I \times J) > 0$. Отсюда если $A^\gamma(I \times J) > 0$, то $A^t(I \times J) > 0$ при любом $t \geqslant \gamma$.

б) Единичная матрица является J^2-s-положительной при любом допустимом множестве J, то есть достаточно показать, что множество $Q_s(J^2)$ замкнуто относительно умножения. Пусть $C = (c_{i,j}) = AB$, где A и B суть J^2-s-положительные матрицы, тогда $a_{i,l(i)} > 0$ и $b_{l(i),j} > 0$ при любом $i \in J$ и при некоторых $l(i), j \in J$. Тогда

$$c_{i,j} = a_{i,1}b_{1,j} + \ldots + a_{i,n}b_{n,j} \geqslant a_{i,l(i)}b_{l(i),j} > 0,$$

то есть i-я строка матрицы C ненулевая при всех $i \in J$. Значит, $C \in Q_s(J^2)$.

Для множества $Q_c(J^2)$ доказательство аналогичное.

Включение $S(J) \subseteq Q_s(J^2) \cap Q_c(J^2)$ следует из определений данных множеств.

в) Если матрица содержит нулевую строку (столбец), то и любая её степень содержит нулевую строку (столбец), отсюда $P(J^2) \subseteq Q_{sc}(n)$. Любая степень J^2-примитивной матрицы также J^2-примитивная, значит, подмножество $P(J^2)$ — наследственное. ■

Замечание 1. Множество $Q_{sc}(I \times J)$ не замкнуто относительно умножения. Например, при $n = 3$, $I = \{1, 2, 3\}$, $J = \{1, 2\}$:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \in Q_{sc}(I \times J), \qquad A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \notin Q_{sc}(I \times J).$$

Рассмотрим частичные порядки на множестве $M_0(n)$ и других множествах.

Напомним, что функция $f : X \to L$, где X — частично упорядоченное множество, L — линейно упорядоченное множество, называется изотонной (антиизотонной), если для любых $x, x' \in X$ из отношения $x \leqslant x'$ следует, что $f(x) \leqslant f(x')$ ($f(x) \geqslant f(x')$).

Для $I, J, I', J' \subseteq N_n$ положим: $(I, J) \leqslant (I', J') \Leftrightarrow I \subseteq I'$ и $J \subseteq J'$. Данное бинарное отношение $\leqslant$ рефлексивно, антисимметрично и транзитивно, следовательно, является отношением частичного порядка на множестве Ω_n^2.

Для $A, B \in M_0(n)$ положим: $A \leqslant B \Leftrightarrow a_{i,j} \leqslant b_{i,j}$ для всех $i, j = 1, \ldots, n$. Если при этом существуют такие i и j, что $a_{i,j} < b_{i,j}$, то $A < B$. Бинарное отношение $\leqslant$ обладает

свойствами рефлексивности, транзитивности и антисимметричности, следовательно, является отношением частичного порядка на множестве $M_0(n)$.

Выполнены следующие монотонные свойства.

Утверждение 2.

а) $I \times J$-exp$A$ есть изотонная функция $\Omega_n^2 \to \mathbb{N}$ при любой матрице $A \in M_0(n)$;

б) $I \times J$-exp$A$ есть антиизотонная функция $M_0(n) \to \mathbb{N}$ при любых допустимых I, J;

в) $(A(J^2))^t \leqslant A^t(J^2)$ при любой матрице $A \in M_0(n)$ и любом допустимом J, $t \geqslant 1$, вследствие этого J^2-exp$A \leqslant \exp A(J^2)$.

Доказательство.

а) Следует из того, что если $(I, J) \leqslant (I', J')$ и $A(I' \times J') > 0$, то $A(I \times J) > 0$.

б) Индукция по степени t матрицы. Если $B \leqslant A$, то $B^t \leqslant A^t$ при $t \geqslant 1$. Действительно, при $t = 1$ неравенство выполнено. Пусть оно выполнено при t. Тогда $A^{t+1} = AA^t = (B + C)(B^t + C')$, где $C, C' \in M_0(n)$. Тогда

$$A^{t+1} = B^{t+1} + B',$$

где $B' = CB^t + BC' + CC' \in M_0(n)$. Отсюда $B^{t+1}(I \times J) \leqslant A^{t+1}(I \times J)$ при любых $I, J \subseteq N_n$ и $t \geqslant 0$. Следовательно, $I \times J$-exp$A \leqslant I \times J$-exp$B$.

в) Индукция по степени t матрицы. Пусть $(A(J^2))^t = (c_{i,j}^{(t)})$, $t = 1, 2, \ldots$ При $t = 1$ неравенство выполнено. Пусть оно выполнено для t, где $t \geqslant 1$. Докажем неравенство для $t + 1$. По правилу умножения матриц для любых $i, j \in N_n$

$$a_{i,j}^{(t+1)} = \sum_{k=1}^{n} a_{i,k}^{(t)} a_{k,j}, \quad c_{i,j}^{(t+1)} = \sum_{k \in J} c_{i,k}^{(t)} c_{k,j}.$$

По предположению $c_{i,k}^{(s)} \leqslant a_{i,k}^{(s)}$ для любого $s \leqslant t$ и для любых $i, j \in J$. Тогда для $i, j \in J$

$$c_{i,j}^{(t+1)} = \sum_{k \in J} c_{i,k}^{(t)} c_{k,j} \leqslant \sum_{k \in J} a_{i,k}^{(t)} a_{k,j} \leqslant \sum_{k=1}^{n} a_{i,k}^{(t)} a_{k,j} = a_{i,j}^{(t+1)}.$$

Индукция доказана. ■

Для матрицы A обозначим $M_{A,J} = \{B \in M_0(n) : B(J^2) \geqslant A(J^2)\}$, где $\varnothing \neq J \subseteq N_n$.

Утверждение 3. Если $A(J^2) \leqslant A^2(J^2)$, то $M_{A,J}$ — полугруппа; если при этом $B(J^2) \leqslant AB(J^2)$ $(B(J^2) \leqslant BA(J^2))$, то $M_{B,J}$ — правый (левый) идеал полугруппы $M_{A,J}$.

Доказательство. Пусть $C, D \in M_{A,J}$. Тогда $C(J^2) = A(J^2) + C'(J^2)$, $D(J^2) = A(J^2) + D'(J^2)$, где C', D' — неотрицательные матрицы. Отсюда

$$CD(J^2) = ((A + C')(A + D'))(J^2) = (A^2 + AD' + C'A + C'D')(J^2) \geqslant A^2(J^2) \geqslant A(J^2).$$

Следовательно, $CD \in M_{A,J}$, то есть $M_{A,J}$ — полугруппа.

Пусть $K \in M_{B,J}$, то есть $K(J^2) \geqslant B(J^2)$. Так как $C(J^2) = A(J^2) + C'(J^2)$, где $C' \in M_0(n)$, то получим

$$CK(J^2) = ((A + C')K)(J^2) = (AK + C'K)(J^2) \geqslant AK(J^2) \geqslant AB(J^2) \geqslant B(J^2).$$

Следовательно, $CK \in M_{B,J}$ и $M_{B,J}$ — левый идеал полугруппы $M_{A,J}$. Для правого идеала доказательство аналогичное. ■

Следствие 1. Если $A(J^2)$ — единичная матрица, то $P(J^2)$ — идеал моноида $M_{A,J}$.

2. Локальная субпримитивность матриц и графов

Пусть A — матрица смежности вершин n-вершинного графа Γ. По определению матрица A и граф Γ одновременно примитивны или не примитивны. Для графа Γ обозначим через D_t матрицу достижимости вершин за t шагов, $t = 1, 2, \ldots$:

$$D_t = \sum_{i=1}^{t} A^i.$$

Известно, что $D_n > 0$, если граф Γ сильносвязный. В соответствии с [3, разд. 10.2], наименьшее натуральное число t, при котором $D_t > 0$, называется субэкспонентом матрицы A, обозначается $\mathrm{sbxp}A$, при этом матрица A называется субпримитивной.

Критерий субпримитивности таков: матрица A субпримитивная, если и только если сильно связен граф Γ, при этом $\mathrm{sbxp}A$ равен диаметру графа Γ. Заметим, что понятие субэкспонента распространено [3, разд. 10.2] на систему матриц таким образом, что понятия субэкспонента системы матриц и диаметра соответствующей системы графов существенно различаются.

Дадим иное обобщение субэкспонента. Матрица A называется $I \times J$-субпримитивной (J^2-субпримитивной при $I = J$, $*J$-субпримитивной при $I = N_n$, $I*$-субпримитивной при $J = N_n$), если для некоторого натурального числа γ матрица $D_\gamma(I \times J)$ (матрица $D_\gamma(J^2)$, матрица $D_\gamma(*J)$, матрица $D_\gamma(I*)$) положительна. Наименьшее такое γ назовём $I \times J$-субэкспонентом (J^2-субэкспонентом при $I = J$, $*J$-субэкспонентом при $I = N_n$, $I*$-субэкспонентом при $J = N_n$) матрицы A и обозначим $I \times J$-$\mathrm{sbxp}A$ (J^2-$\mathrm{sbxp}A$, $*J$-$\mathrm{sbxp}A$, $I*$-$\mathrm{sbxp}A$). Заметим, что если $D_\gamma(I \times J) > 0$, то $D_t(I \times J) > 0$ при всех $t \geqslant \gamma$. Множество субпримитивных ($I \times J$-субпримитивных, J^2-субпримитивных, $*J$-субпримитивных, $I*$-субпримитивных) матриц обозначим $\Sigma(n)$ ($\Sigma(I \times J)$, $\Sigma(J^2)$, $\Sigma(*J)$, $\Sigma(I*)$).

Из определений следует, что при любых допустимых множествах I, J

$$P(n) \subseteq \Sigma(n) \cap P(I \times J) \subseteq P(I \times J) \subseteq \Sigma(I \times J).$$

Следовательно, для любой матрицы $A \in M_0(n)$

$$I \times J\text{-}\mathrm{sbxp}A \leqslant I \times J\text{-}\exp A \leqslant \max\{I \times J\text{-}\exp A, \mathrm{sbxp}A\} \leqslant \exp A.$$

Утверждение 4. Если матрицы $A, B \in M_0(n)$ сопряжены в $S(J)$, то они одновременно обладают или не обладают свойством ξ, где ξ — любое из свойств: J^2-примитивность, J^2-субпримитивность, J^2-s-положительность, J^2-c-положительность, J^2-sc-положительность.

Доказательство. Если матрицы A и B сопряжены в $S(J)$, то найдётся перестановочная матрица $T \in S(J)$, такая, что $A = T^{-1}BT$, отсюда $A^t = T^{-1}B^tT$ при любом $t \in \mathbb{N}$. Следовательно, при любом натуральном t матрицы A^t и B^t отличаются лишь перестановкой строк и столбцов с номерами $i \in J$, то есть обладают одинаковыми указанными свойствами. ■

Обозначим: l_{ij} — длина кратчайшего пути от вершины i до вершины j в графе Γ; $d(I, J) = \max\limits_{(i,j) \in I \times J} l_{ij}$ — расстояние между наиболее удалёнными вершинами множеств I и J. При $I = J = N_n$ величина $d(I, J)$ совпадает с диаметром графа Γ.

Теорема 1. Матрица A является $I \times J$-субпримитивной, если и только если в графе Γ существует путь из любой вершины $i \in I$ в любую вершину $j \in J$, при этом $I \times J$-$\mathrm{sbxp}A = d(I, J)$.

Доказательство. Согласно [4, с. 143], $a_{i,j}^{(t)} > 0$, если и только если в орграфе Γ имеется путь из i в j длины t. Отсюда матрица $D_t(I \times J) > 0$, если и только если в графе Γ для любых $i \in I$ и $j \in J$ существует путь из i в j длины не более t. Следовательно, $I \times J\text{-sbxp}A = d(I, J)$. ■

3. Локальная примитивность ориентированных графов

Граф Γ называется $I \times J$-примитивным, если матрица A является $I \times J$-примитивной, при этом соответствующие $I \times J$-экспоненты матрицы A и графа Γ равны. Получим условия $I \times J$-примитивности на теоретико-графовом языке.

Обозначим при любых допустимых I, J:

$\bar{J}$ — множество $N_n \setminus J$;

$\gamma_{I\times J}$ — величину $I \times J\text{-exp}\,\Gamma$ для $I \times J$-примитивного графа Γ, где $\gamma_{i\times j} = \gamma_{I\times J}$ при $I = \{i\}$, $J = \{j\}$;

γ_{J2} — величину $J^2\text{-exp}\,\Gamma$ для J^2-примитивного графа Γ;

γ_{*J} — величину $*J\text{-exp}\,\Gamma$ для $*J$-примитивного графа Γ;

γ_{I*} — величину $I*\text{-exp}\,\Gamma$ для $I*$-примитивного графа Γ;

$\rho(I, J)$ — величину $\min\limits_{(i,j)\in I\times J} l_{ij}$, то есть расстояние в графе Γ от множества I до множества J ($\rho(I, J) = 0$ при $I \cap J \neq \varnothing$, $\rho(i, J) = \rho(I, J)$ при $I = \{i\}$, $\rho(I, j) = \rho(I, J)$ при $J = \{j\}$);

$\theta(I, J)$ — величину $\max\limits_{i\in I} \rho(i, J)$, то есть расстояние достижимости из любой вершины множества I некоторой вершины множества J ($\theta(i, J) = \rho(i, J) = \theta(I, J)$ при $I = \{i\}$);

$\tau(I, J)$ — величину $\max\limits_{j\in J} \rho(I, j)$, то есть расстояние достижимости из некоторой вершины множества I любой вершины множества J ($\tau(I, j) = \rho(I, j) = \tau(I, J)$ при $J = \{j\}$).

Обозначим также в графе Γ:

$L(i, j)$ — множество длин всех простых путей из i в j;

$L(\Gamma)$ — множество длин всех простых циклов;

U — множество вершин графа $\tilde{U}$, где $\tilde{U}$ — часть графа Γ (цикл, подграф, ...).

Пусть $Y \subseteq N_n$. Путь в Γ из i в j, проходящий через некоторые вершины непустого множества Y, назовем Y-путём из i в j. Сильносвязный подграф $\tilde{U}$ орграфа Γ назовём i, j-связывающим, если в Γ существует U-путь из i в j. В частности, сильносвязный орграф есть i, j-связывающий орграф при любых i, j.

Пусть $L(\Gamma) = \{l_1, \ldots, l_m\}$, где $l_1 < \ldots < l_m$. Заметим, что $(l_1, \ldots, l_m) = 1$, если и только если орграф Γ примитивный [5, с. 226].

Утверждение 5. Если орграф Γ является $I \times J$-примитивным, то орграф Γ имеет i, j-связывающий подграф для любой пары $(i, j) \in I \times J$.

Доказательство. Если орграф Γ является $I \times J$-примитивным, то матрица A является $I \times J$-примитивной, тогда $A^t(I \times J) > 0$ при любом $t \geqslant \gamma$, где $\gamma \in \mathbb{N}$. Отсюда в Γ имеется путь w длины t из i в j для любой пары $(i, j) \in I \times J$ и любого $t \geqslant \gamma$. Число вершин в подграфе не более n, значит, при $t > n$ путь w содержит цикл $\tilde{C}$. Следовательно, подграф $\Gamma(C)$ с множеством вершин C является сильносвязным и i, j-связывающим подграфом. ■

Утверждение 6. Связный циклический орграф Γ является $I \times J$-примитивным, если и только если Γ является $i \times j$-примитивным для любой пары вершин $(i, j) \in I \times J$, при этом $\gamma_{I\times J} = \max\limits_{(i,j)\in I\times J} \gamma_{i\times j}$.

В силу утверждений 5 и 6 достаточно получить условия $i \times j$-примитивности для связного циклического орграфа Γ.

Для множества взаимно простых натуральных чисел $\{a_1, \ldots, a_m\}$ обозначим $g(a_1, \ldots, a_m)$ число Фробениуса, то есть наибольшее натуральное число, не принадлежащее аддитивной полугруппе, порождённой множеством $\{a_1, \ldots, a_m\}$.

Сильносвязный орграф Γ с n вершинами называется r-дольным с блоками V_0, V_1, $\ldots$, V_{r-1}, где $r > 1$, если $V_0, V_1, \ldots, V_{r-1}$ — блоки разбиения множества вершин орграфа Γ, такие, что если (i, j) — дуга и $i \in V_t$, то $j \in V_{(t+1) \bmod r}$, $t = 0, 1, \ldots, r - 1$.

Множество всех путей орграфа Γ образует частичный моноид [6] относительно операции конкатенации, обозначаемой точкой. Данная операция определена на паре путей (u, v), если и только если конечная вершина пути u совпадает с начальной вершиной пути v. Если u — путь с начальной вершиной i и конечной вершиной a, v — путь с начальной вершиной a и конечной вершиной j, то $w = u \cdot v$ есть путь с начальной вершиной i и конечной вершиной j. При этом $\operatorname{len} w = \operatorname{len} u + \operatorname{len} v$, где $\operatorname{len} w$ — длина пути w, то есть число дуг в нём.

Обозначим в графе Γ: $[i, j]$ — путь из i в j; $\langle i, j\rangle$ — кратчайший путь из i в j. Часть цикла $\tilde{C}$ в орграфе Γ, являющаяся путём из i в j, где $i, j \in C$, обозначим $[i, j]_C$. При обходе цикла $\tilde{C}$ выделим его вершину a как начальную, цикл $\tilde{C}$ в этом случае обозначим $\tilde{C}(a)$. Для целого неотрицательного e через $e\tilde{C}(a)$ обозначим цикл, составленный из e-кратно пройденного цикла $\tilde{C}(a)$, где $0\tilde{C}(a)$ — пустой путь.

Лемма 1. Пусть Γ — сильносвязный n-вершинный орграф, $L(\Gamma) = \{l_1, \ldots, l_m\}$ и $(l_1, \ldots, l_m) = d > 1$. Тогда Γ является d-дольным орграфом с блоками $V_0, V_1, \ldots, V_{d-1}$, и при $s = 0, \ldots, d - 1$ для любых вершин $i, j \in V_s$ существует путь $[i, j]$ длины t, где t — любое число, кратное d, и $t \geqslant f_n(l_1, \ldots, l_m)$, где

$$f_n(l_1, \ldots, l_m) = dg\left(\frac{l_1}{d}, \ldots, \frac{l_m}{d}\right) + d + n(m + 1) - 1 - \sum_{k=1}^{m} l_k. \tag{3}$$

Доказательство. Не ограничивая общности, положим, что $i, j \in V_0$. Заметим, что путь $[i, j]$ существует, так как орграф Γ сильносвязный. Пусть цикл $\tilde{C}_k$ имеет длину l_k, $k = 1, \ldots, m$. Путь $[i, j]$ представим в виде

$$[i, j] = [i, a_1] \cdot (e_1\tilde{C}_1(a_1)) \cdot [a_1, a_2] \cdot (e_2\tilde{C}_2(a_2)) \cdot \ldots \cdot [a_{m-1}, a_m] \cdot (e_m\tilde{C}_m(a_m)) \cdot \langle a_m, j\rangle,$$

где $e_1, \ldots, e_m$ — целые неотрицательные числа; a_1 — вершина цикла $\tilde{C}_1$, ближайшая к вершине i; a_k — вершина цикла $\tilde{C}_k$, ближайшая к вершине a_{k-1}, $k = 2, \ldots, m$. Тогда

$$\operatorname{len}[i, j] = e_1l_1 + \ldots + e_ml_m + \operatorname{len}[i, a_1] + \sum_{k=2}^{m} \operatorname{len}[a_{k-1}, a_k] + \operatorname{len}\langle a_m, j\rangle. \tag{4}$$

По условию $(l_1, \ldots, l_m) = d > 1$. Тогда, в соответствии с [5, с. 390], орграф Γ является d-дольным; пусть $V_0, V_1, \ldots, V_{d-1}$ — его блоки. Так как $i, j \in V_0$, то $\operatorname{len}[i, j]$ кратна d.

Вместе с тем $(l_1/d, \ldots, l_m/d) = 1$. Тогда любое натуральное число, большее $g(l_1/d, \ldots, l_m/d)$, может быть представлено линейной комбинацией чисел $l_1/d, \ldots, l_m/d$. Отсюда любое натуральное число, кратное d и большее $dg(l_1/d, \ldots, l_m/d)$, может быть

представлено линейной комбинацией чисел $l_1, \ldots, l_m$. Значит, подбирая коэффициенты $e_1, \ldots, e_m$, можно получить сумму $e_1 l_1 + \ldots + e_m l_m$, равную любому числу t, которое кратно d и превышает $dg(l_1/d, \ldots, l_m/d)$. Так как $\operatorname{len}\langle a_m, j\rangle \leqslant n-1$ и расстояние от любой вершины до цикла $\tilde{C}_k$ не превышает $n - l_k$, $k = 1, \ldots, m$, из (4) следует, что, подбирая коэффициенты $e_1, \ldots, e_m$, можно получить $\operatorname{len}[i, j]$, равную любому числу t, которое кратно d и не меньше $f_n(l_1, \ldots, l_m)$. ■

При натуральном d множество натуральных чисел M назовём d-полным, если M содержит полную систему вычетов по модулю d. Наименьшим d-трансверсалом d-полного множества M назовём подмножество (обозначаемое $M(d)$), состоящее из наименьших чисел множества M, образующих полную систему вычетов по модулю d.

Пусть $R = \{r_1, r_2, \ldots\}$ и $R' = \{r'_1, r'_2, \ldots\}$ — множества натуральных чисел. Под суммой множеств (обозначается $R + R'$) понимаем множество натуральных чисел

$$R + R' = \{r_i + r'_j : i, j = 1, 2, \ldots\}.$$

Теорема 2. Связный циклический орграф Γ является $i \times j$-примитивным, если в графе Γ выполнено хотя бы одно из условий:

а) имеются примитивные i, j-связывающие подграфы $\tilde{U}_1, \ldots, \tilde{U}_k$, $k \geqslant 1$; тогда

$$\gamma_{i\times j} \leqslant \min_{1\leqslant r\leqslant k}\{\rho(i, U_r) + \exp\tilde{U}_r + \rho(U_r, j)\};$$

б) имеется i, j-связывающий подграф $\tilde{V}$, $L(\tilde{V}) = \{l_1, \ldots, l_m\}$, $(l_1, \ldots, l_m) = d > 1$, и для некоторых вершин μ, ν графа $\tilde{V}$ множество $M_{(i,\mu)+(\nu,j)}$ является d-полным, где $M_{(i,\mu)+(\nu,j)} = L(i, \mu) + L(\nu, j)$; тогда

$$\gamma_{i\times j} \leqslant dg\left(\frac{l_1}{d}, \ldots, \frac{l_m}{d}\right) + 2d + n(m+1) + q_{i,j} - 2 - \sum_{k=1}^{m} l_k,$$

где $q_{i,j}$ — наибольшее число d-трансверсала $M_{(i,\mu)+(\nu,j)}(d)$.

Доказательство.

а) Если $\tilde{U}$ — примитивный подграф графа Γ, то для любого $t \geqslant \exp\tilde{U}$ существует путь в $\tilde{U}$ длины t из любой вершины в любую. По условию множество U достижимо из вершины i не более чем за $\rho(i, U)$ шагов, и из множества U достижима вершина j не более чем за $\rho(U, j)$ шагов. Тогда в Γ существует путь длины t из i в j при любом $t \geqslant \exp\tilde{U} + \rho(i, U) + \rho(U, j)$. Значит, граф Γ является $i \times j$-примитивным и неравенство для локального экспонента $\gamma_{i\times j}$ выполнено.

б) По условию орграф $\tilde{V}$ является d-дольным. Пусть $V_0, V_1, \ldots, V_{d-1}$ — его блоки, $\mu \in V_s$, $\nu \in V_h$, где $s, h \in \{0, \ldots, d-1\}$. Рассмотрим систему путей $[i, j]_p$, $p = 0, \ldots, d-1$:

$$[i, j]_p = [i, \mu]_p \cdot [\mu, \mu'] \cdot [\mu', \nu] \cdot [\nu, j]_p,$$

где μ' — ближайшая к вершине ν вершина из блока V_s; пути $[i, \mu]_p$ и $[\nu, j]_p$ являются простыми и $\operatorname{len}[i, \mu]_p + \operatorname{len}[\nu, j]_p \equiv p \pmod d$ — по условию для любого $p = 0, \ldots, d-1$ в Γ такая пара путей имеется. Тогда при $p = 0, \ldots, d-1$

$$\operatorname{len}[i, j]_p = \operatorname{len}[\mu', \nu] + (\operatorname{len}[i, \mu]_p + \operatorname{len}[\nu, j]_p) + \operatorname{len}[\mu, \mu']. \tag{5}$$

По определению вершины μ' имеем: $\operatorname{len}[\mu', \nu] \leqslant d-1$. Так как наибольшее число d-трансверсала множества $L(i, \mu) + L(\nu, j)$ равно $q_{i,j}$, то $\operatorname{len}[i, \mu]_p + \operatorname{len}[\nu, j]_p \leqslant q_{i,j}$ при

любом p. По лемме 1 можно построить путь $[\mu,\mu']$, длина которого равна любому числу t, кратному d и не меньшему $f_n(l_1,\dots,l_m)$. Так как множество $M_{(i,\mu)+(\nu,j)}$ является d-полным, то, в соответствии с (5), можно при любом $p=0,\dots,d-1$ построить путь $[i,j]_p$, длина которого равна любому числу t, сравнимому с p по модулю d и не меньшему $f_n(l_1,\dots,l_m)+d+q_{i,j}-1$. Следовательно, граф Γ является $i\times j$-примитивным и оценка для $\gamma_{i\times j}$ верна. ■

Следствие 2. При выполнении условия *б* теоремы 2 верна оценка

$$\gamma_{i\times j}\leqslant\frac{l_1l_m}{d}-l_1-l_m+2d+n(m+3)-4-\sum_{k=1}^{m}l_k.$$

Доказательство. Следует из оценки числа Фробениуса [7, теорема 3.1.1]

$$g(a_1,\dots,a_m)\leqslant a_1a_m-a_1-a_m,$$

а также оценки $q_{i,j}\leqslant 2n-2$. ■

Замечание 2. Для снижения оценки величины $\gamma_{i\times j}$, полученной в п. *б* теоремы 2, можно использовать эквивалентное подмножество $\{\lambda_1,\dots,\lambda_r\}$ множества $\{l_1,\dots,l_m\}$, такое, что $r\leqslant m$, $(\lambda_1,\dots,\lambda_r)=(l_1,\dots,l_m)=d$ и $g(\lambda_1/d,\dots,\lambda_r/d)=g(l_1/d,\dots,l_m/d)$. Построение эквивалентных подмножеств $\{\lambda_1,\dots,\lambda_r\}$ исследовалось в [6].

Замечание 3. Условие *б* теоремы 2 может быть расширено не только за счёт рассмотрения в орграфе Γ нескольких i,j-связывающих не примитивных подграфов, но и за счёт d-полноты объединения по данным подграфам множеств вида $M_{(i,\mu)+(\nu,j)}$.

Теорема 3. Граф Γ является J^2-примитивным, если и только если в Γ имеется примитивная компонента сильной связности $\tilde{U}$, содержащая множество вершин J; при этом

$$\gamma_{J2}\leqslant\exp\tilde{U}\leqslant\gamma_{J2}+\theta(U\setminus J,J)+\tau(J,U\setminus J).$$

Доказательство. Необходимость. Пусть граф Γ является J^2-примитивным. Тогда существует путь длины t из μ в ν для любых $\mu,\nu\in J$ и любого $t\geqslant\gamma_{J2}$. Значит, $J\subseteq U$, где $\tilde{U}$ — компонента сильной связности графа Γ. Тогда при любом $t\geqslant\gamma_{J2}$ существует путь длины $\lambda=t+\theta(U\setminus J,J)+\tau(J,U\setminus J)$ из i в j для любых $i,j\in U$. Следовательно, граф $\tilde{U}$ примитивный с экспонентом, не превышающим $\gamma_{J2}+\theta(U\setminus J,J)+\tau(J,U\setminus J)$.

Нижняя оценка $\exp\tilde{U}$ следует из п. *а* утверждения 2.

Достаточность. Если $\tilde{U}$ — примитивная компонента сильной связности в графе Γ, где $J\subseteq U$ и A_U — матрица смежности вершин графа $\tilde{U}$, то существует натуральное число γ, такое, что $A_U^t>0$ при любом $t\geqslant\gamma$. Значит, $A_U^t(J^2)>0$ при любом $t\geqslant\gamma$, то есть матрица A_U является J^2-примитивной. Следовательно, граф Γ является J^2-примитивным. ■

Следствие 3. J^2-примитивный граф Γ примитивен, если и только если Γ сильносвязный, при этом

$$\gamma_{J2}\leqslant\exp\Gamma\leqslant\gamma_{J2}+\theta(\bar{J},J)+\tau(J,\bar{J}).$$

Следствие 4. Граф Γ является $*J$-примитивным ($I*$-примитивным), если и только если Γ является J^2-примитивным (I^2-примитивным) и из каждой вершины $i\notin U$ достижимо множество вершин U (из множества вершин U достижима каждая вершина $i\notin U$), при этом

$$\gamma_{J2}\leqslant\gamma_{*J}\leqslant\gamma_{J2}+\theta(\bar{J},J)\qquad(\gamma_{I2}\leqslant\gamma_{I*}\leqslant\gamma_{I2}+\tau(I,\bar{I})).$$

Доказательство. Докажем следствие для $*J$-примитивности графа; для $I*$-примитивности доказательство аналогично.

Необходимость. Если граф Γ является J-примитивным, то по определению он является и J^2-примитивным. Вместе с тем множество вершин J достижимо из любой вершины графа. Так как $\tilde{U}$ — компонента сильной связности и $J \subseteq U$, то из любой вершины графа достижимо множество вершин U.

Достаточность. Если граф Γ является J^2-примитивным, то $A^t(J^2) > 0$ при любом $t \geqslant \gamma_{J2}$. Вместе с тем множество вершин U достижимо из любой вершины $i \notin U$; тогда множество вершин J достижимо из любой вершины $i \notin J$, следовательно, $A^t(*J) > 0$ при всех натуральных $t \geqslant \gamma_{J2} + \theta(\bar{J}, J)$. Значит, граф Γ является J-примитивным и $\gamma_{*J} \leqslant \gamma_{J2} + \theta(\bar{J}, J)$. Нижняя оценка γ_{*J} следует из п. *а* утверждения 2. ■

Замечание 4. Если граф Γ является J^2-примитивным и $\tilde{U}$ — компонента сильной связности графа Γ, такая, что $J \subseteq U$, то Γ является U^2-примитивным.

Если при этом $|U| = m$ и граф Γ примитивен ($*U$-примитивен, $U*$-примитивен), то $\max\{\theta(\bar{U}, U), \tau(U, \bar{U})\} \leqslant n - m$; тогда из следствий 3 и 4 имеем соответственно

$$\exp\Gamma \leqslant \gamma_{U2} + 2(n-m), \quad \gamma_{*U} \leqslant \gamma_{U2} + n - m, \quad \gamma_{U*} \leqslant \gamma_{U2} + n - m.$$

Так как для примитивного подграфа $\tilde{U}$ из п. *в* утверждения 2 следует, что $\gamma_{U2} \leqslant \exp\tilde{U}$, то с помощью известных оценок экспонентов отсюда получаются огрублённые оценки. Например, при использовании оценки Виландта [1, с. 7] для $\exp\tilde{U}$ получаем

$$\exp\Gamma \leqslant m^2 - 4m + 2n + 2, \quad \gamma_{*U} \leqslant m^2 - 3m + n + 2, \quad \gamma_{U*} \leqslant m^2 - 3m + n + 2.$$

4. Локальная квазипримитивность неотрицательных матриц и графов

Матрица A называется $I \times J$-квазипримитивной (J^2-квазипримитивной при $I = J$, $*J$-квазипримитивной при $I = N_n$, $I*$-квазипримитивной при $J = N_n$), если при некотором натуральном числе δ подматрица $A^t(I \times J)$ ($A^t(J^2)$, $A^t(*J)$, $A^t(I*)$) s-положительна для любого $t \geqslant \delta$. Наименьшее такое δ назовём $I \times J$-квазиэкспонентом матрицы A (J^2-квазиэкспонентом при $I = J$, $*J$-квазиэкспонентом при $I = N_n$, $I*$-квазиэкспонентом при $J = N_n$) и обозначим $I \times J$-$\operatorname{qexp} A$ (J^2-$\operatorname{qexp} A$, $*J$-$\operatorname{qexp} A$, $I*$-$\operatorname{qexp} A$). Множество $I \times J$-квазипримитивных (J^2-квазипримитивных, $*J$-квазипримитивных, $I*$-квазипримитивных) матриц обозначим $\Pi(I \times J)$ ($\Pi(J^2)$, $\Pi(*J)$, $\Pi(I*)$).

$I \times J$-квазипримитивность матрицы при $I = J = N_n$ равносильна s-положительности. Определим: граф Γ является $I \times J$-квазипримитивным, если и только если матрица A является $I \times J$-квазипримитивной; соответствующие $I \times J$-квазиэкспоненты матрицы A и графа Γ равны.

Обозначим при любых допустимых I, J:

$\delta_{I\times J}$ — величину $I \times J$-$\operatorname{qexp} A$ для $I \times J$-квазипримитивной матрицы A;

δ_{J2} — величину J^2-$\operatorname{qexp} A$ для J^2-квазипримитивной матрицы A;

δ_{*J} — величину $*J$-$\operatorname{qexp} A$ для $*J$-квазипримитивной матрицы A;

δ_{I*} — величину $I*$-$\operatorname{qexp} A$ для $I*$-квазипримитивной матрицы A.

Замечание 5. При любых допустимых множествах I, J

$$P(I \times J) \subseteq \Sigma(I \times J) \cap \Pi(I \times J),$$

следовательно, для любой матрицы $A \in M_0(n)$

$$\max\{\delta_{I\times J}, I \times J\text{-}\operatorname{sbxp} A\} \leqslant \gamma_{I\times J}.$$

Таким образом, наиболее сложным является описание $I \times J$-квазипримитивных матриц (графов), не являющихся $I \times J$-примитивными.

Матрицу A (граф Γ) будем называть локально примитивной (локально субпримитивной, локально квазипримитивной), если она $I \times J$-примитивная ($I \times J$-субпримитивная, $I \times J$-квазипримитивная) при некоторых допустимых I, J, где $I \cap J \neq N_n$. Соответствующие величины экспонентов назовём локальными экспонентами (локальными субэкспонентами, локальными квазиэкспонентами) матрицы A (графа Γ).

Утверждение 7. Для любой матрицы $A \in M_0(n)$:

а) для любого фиксированного множества I локальный квазиэкспонент $I \times J$-qexp A является антиизотонной функцией $\Omega_n \to \mathbb{N}$;

б) для любого фиксированного множества J локальный квазиэкспонент $I \times J$-qexp A является изотонной функцией $\Omega_n \to \mathbb{N}$;

в) при любых фиксированных подмножествах I, J локальный квазиэкспонент $I \times J$-qexp A является антиизотонной функцией $M_0(n) \to \mathbb{N}$.

Доказательство.

а) Если $J \subseteq J'$ и матрица $A(I \times J)$ s-положительна, то матрица $A(I \times J')$ также s-положительна.

б) Если $I \subseteq I'$ и матрица $A(I' \times J)$ s-положительна, то s-положительна и матрица $A(I \times J)$.

в) Если $B \leqslant A$, то $B^t \leqslant A^t$ при любом $t \geqslant 1$; следовательно, $B^t(I \times J) \leqslant A^t(I \times J)$ при любых $I, J \subseteq N_n$ и $t \geqslant 1$. ■

Утверждение 8.

а) Если орграф Γ является $I \times J$-квазипримитивным, то орграф Γ имеет i, j_i-связывающий подграф для любого $i \in I$ и некоторого $j_i \in J$.

б) $Q_s(J^2) \subset \Pi(J^2)$, и $\delta_{J2} = 1$ для любой матрицы $A \in Q_s(J^2)$.

Доказательство.

а) Если орграф Γ является $I \times J$-квазипримитивным, то матрица A является $I \times J$-квазипримитивной; тогда матрица $A^t(I \times J)$ является s-положительной при любом $t \geqslant \gamma$, где $\gamma \in \mathbb{N}$. Отсюда в Γ имеется путь w длины t из i в j_i для любого $i \in I$ и некоторого $j_i \in J$ при любом $t \geqslant \gamma$. Число вершин в подграфе не более n, значит, при $t > n$ путь w содержит цикл $\tilde{C}$. Следовательно, подграф $\Gamma(C)$ с множеством вершин C является сильносвязным и i, j_i-связывающим подграфом.

б) Если $A \in Q_s(J^2)$, то, согласно п. *б* утверждения 1, $A^t \in Q_s(J^2)$ при любом натуральном t. Следовательно, $A \in \Pi(J^2)$ и $\delta_{J2} = 1$. ■

Следствие 5. Если орграф Γ является $I \times J$-квазипримитивным, то множество J достижимо из всех вершин множества I.

Теорема 4.

а) Пусть $\tilde{U}$ — сильносвязный d-дольный подграф орграфа Γ с блоками U_0, U_1, …, U_{d-1}, где $L(\tilde{U}) = \{l_1, \ldots, l_m\}$; $d = (l_1, \ldots, l_m) > 1$. Граф Γ является $U \times J$-квазипримитивным, если и только если $J \cap U_s \neq \varnothing$, $s = 0, 1, \ldots, d-1$, в этом случае при $|U| = r$ выполнено

$$U \times J\text{-qexp}\,\Gamma \leqslant f_r(l_1, \ldots, l_m) + d - 1,$$

где величина $f_r(l_1, \ldots, l_m)$ определена равенством (3).

б) Пусть в графе Γ подграф $\tilde{J}$ сильносвязный или состоит из компонент сильной связности, где J есть множество вершин Γ, достижимых из любой вершины множе-

ства I. Тогда орграф Γ является $U \times J$-квазипримитивным при $U = I \cup J$ и

$$U \times J\text{-qexp}\,\Gamma = \theta(I, J).$$

Доказательство.

а) Необходимость. Пусть граф Γ является $U \times J$-квазипримитивным, тогда при любом $t \geqslant \delta_{U\times J}$ матрица $A^t(U \times J)$ является s-положительной. Значит, для любого $u \in U$ и любого $t \geqslant \delta_{U\times J}$ в Γ имеется путь длины t из u в одну из вершин J. Вместе с тем, если взять пути длины $t, t+1, \ldots, t+d-1$, то с учётом d-дольности орграфа $\tilde{U}$ получим, что $J \cap U_s \neq \varnothing$, $s = 0, 1, \ldots, d-1$.

Достаточность. Пусть $j_s \in J \cap U_s$, $s = 0, 1, \ldots, d-1$. Не ограничивая общности, рассмотрим блок U_0. В блоке U_0 имеется вершина u_s, из которой вершина j_s достижима за s шагов, $s = 0, \ldots, d-1$. По лемме 1 для любой вершины $u \in U_0$ можно построить путь $[u, u_s]$ длины t, где t — любое число, кратное d и не меньшее $f_r(l_1, \ldots, l_m)$. Тогда имеется путь $[u, j_s]$ длины $t+s$, $s = 1, \ldots, d-1$. Следовательно, для любой вершины $u \in U$ и любого $t \geqslant f_r(l_1, \ldots, l_m) + d - 1$ имеется путь длины t из u в одну из вершин множества J. Это означает, что граф Γ является $U \times J$-квазипримитивным и оценка для $U \times J$-qexp Γ верна.

б) Из любой вершины $i \in I$ существует путь длины $\theta(i, J)$ в некоторую вершину $j_i \in J$. По условию для любой вершины $j \in J$ и любого $r \in \mathbb{N}$ существует путь $[j, v]$ длины r, где $v \in J$. Тогда для любой вершины $i \in U$ в Γ имеется путь $[i, v]$ длины $\theta(i, J) + r$, где $v \in J$ и $r \in \mathbb{N}$ (если $i \in I$, то $[i, v] = [i, j_i] \cdot [j_i, v]$). Следовательно, граф Γ $U \times J$-квазипримитивен и $U \times J$-qexp $\Gamma = \max\limits_{i\in I} \theta(i, J) = \theta(I, J)$. ■

Пример (локально квазипримитивная, но не локально примитивная матрица).

Пусть $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in Q_{sc}(I \times J)$, соответствующий граф $\Gamma(A)$ изображён на рис. 1.

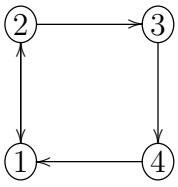


Рис. 1. Граф $\Gamma(A)$

Граф Γ является двудольным с блоками $V_0 = \{1, 3\}$ и $V_1 = \{2, 4\}$. При $J = \{1, 4\}$ оба блока содержат по одной вершине из множества J. Тогда по теореме 4 матрица A является $*J$-квазипримитивной. Вместе с тем по теореме 3 матрица A не является $*J$-примитивной.

Полученные результаты могут быть использованы для изучения перемешивающих свойств композиций криптографических преобразований.

ЛИТЕРАТУРА

1. *Когос К. Г., Фомичев В. М.* Положительные свойства неотрицательных матриц // Прикладная дискретная математика. 2012. № 4(18). С. 5–13.

2. *Кяжин С. Н.* О локальной примитивности графов и неотрицательных матриц // Прикладная дискретная математика. Приложение. 2013. № 6. С. 81–83.
3. *Фомичев В. М.* Методы дискретной математики в криптологии. М.: Диалог-МИФИ, 2010. 424 с.
4. *Берж К.* Теория графов и её применения. М.: ИЛ, 1962. 320 с.
5. *Сачков В. Н., Тараканов В. Е.* Комбинаторика неотрицательных матриц. М.: ТВП, 2000. 448 с.
6. *Фомичев В. М.* Эквивалентные по Фробениусу примитивные множества чисел // Прикладная дискретная математика. 2014. № 1(23). С. 20–26.
7. *Alfonsin J. R.* The Diophantine Frobenius Problem. Oxford University Press, 2005. 243 p.

2014 Прикладная теория графов № 3(25)

DOI 10.17223/20710410/25/7

УДК 519.87:519.6:519.178

К ВОПРОСУ О МАКСИМАЛЬНО ДОСТИЖИМОМ ЧИСЛЕ ВЕРШИН ЦИРКУЛЯНТНЫХ ГРАФОВ ПРИ ЛЮБОМ ДИАМЕТРЕ

Э. А. Монахова, О. Г. Монахов

Институт вычислительной математики и математической геофизики СО РАН, г. Новосибирск, Россия

E-mail: emilia@rav.sscc.ru

Рассматривается задача о максимально достижимом числе вершин при заданных размерности и диаметре неориентированных циркулянтных графов. В 1994 г. Ф. П. Муга доказал теорему о том, что это число является нечётным при любых размерностях и диаметрах циркулянтного графа, что подтверждается для одно-, двух- и трёхмерных циркулянтов. В настоящей работе доказано, что найденное доказательство теоремы некорректно. На основании новых данных скорректирована таблица максимально достижимых порядков циркулянтов размерности четыре.

Ключевые слова: *неориентированные циркулянтные графы, диаметр, максимальный порядок графа.*

Введение

Циркулянтные графы являются графами Кэли абелевых групп и широко изучаются в теории графов и дискретной математике, играя также важную роль в разнообразных приложениях (см. [1–3] и ссылки в них).

Пусть $s_1, s_2, \ldots, s_k, n$ — целые числа, такие, что $1 \leqslant s_1 < s_2 < \ldots < s_k < n$, и $S = \{s_1, s_2, \ldots, s_k\}$. Неориентированный граф $C(n; S)$ с множествами вершин $V = \{0, 1, \ldots, n-1\}$ и рёбер $E = \{(v, (v \pm s_l) \bmod n) : v \in V, l = 1, \ldots, k\}$, называется *циркулянтным*, числа из множества S — *образующими*, k — *размерностью*, n — *порядком* графа. *Диаметром* графа G называется $\mathrm{d}(G) = \max\limits_{i,j \in V} \mathrm{d}(i, j)$, где $\mathrm{d}(i, j)$ — длина кратчайшего пути из вершины i в вершину j графа G.

Пусть $x \in V$ — вершина циркулянтного графа $C(n; S)$, а y — другая его вершина, такая, что $y = \sum\limits_{i=1}^{k} \alpha_i s_i(x)$. Тогда будем говорить, что y достижима из x за $\sum\limits_{i=1}^{k} |\alpha_i|$ шагов. Поскольку циркулянтные графы являются вершинно-транзитивными, достаточно рассматривать в качестве начальной вершины нулевую. В циркулянте размерности k функция $P(d, k)$ определяет максимальное (теоретически) число вершин, которые могут быть достижимы из любой вершины графа самое большее за d шагов. Известно (см. ссылки в [1]), что

$$P(d, k) = \sum_{i=0}^{k} C_k^i C_d^{k-i} 2^{k-i},$$

где значение функции $P(d, k)$ может рассматриваться как граница Мура для циркулянтных графов размерности k.

Для того чтобы в циркулянте диаметра d достигалась эта верхняя граница, необходимо, чтобы различные комбинации кратностей образующих (и их обратных) создавали пути из нулевой вершины длины от 1 до d, которые ведут к различным вершинам.

Отметим, что достижение этой верхней границы эквивалентно достижению плотной упаковки пространства $\mathbb{Z}^k$ k-мерными сферами Ли $S_{k,d}$ радиуса d [4–6]. При этом $S_{k,d}$ определяется для любого заданного диаметра d как множество элементов пространства $\mathbb{Z}^k$, которые могут быть выражены как слова длины не более d через канонические образующие $\mathbf{e}_i$, $1 \leqslant i \leqslant k$, пространства $\mathbb{Z}^k$, взятые положительно или отрицательно. Или если рассматривать метрику L^1 (Manhattan): $S_{k,d}$ есть множество точек в $\mathbb{Z}^k$ на расстоянии не более d от нуля, то есть $S_{k,d} = \{(x_1, \ldots, x_k) \in \mathbb{Z}^k : |x_1| + \ldots + |x_k| \leqslant d\}$.

В работе [7] авторы доказали, что плотная упаковка пространства $\mathbb{Z}^k$ k-мерными сферами Ли возможна для любой размерности $k \geqslant 1$ для радиуса 1 и для размерностей $k = 1, 2$ для любого радиуса. Предположение Голомба — Вельча [7] утверждает, что это невозможно для любой размерности $k > 2$ и радиуса $d > 1$. По подходам к решению этой проблемы см., например, работу [5], в которой получены компактные доказательства невозможности такой упаковки для размерностей $k = 3, 4, 5$, и ссылки в [4–6].

Следуя [8], определим для любых натуральных d и k экстремальную функцию $M(d, k)$ как максимально возможное (достижимое) натуральное n, такое, что существует множество образующих $S = \{1, s_2, \ldots, s_k\}$, при котором $\mathrm{d}(C(n; S)) \leqslant d$. Имеем $M(d, k) = P(d, k)$ для $k \leqslant 2$, а именно $M(d, 1) = 2d + 1$, $M(d, 2) = 2d^2 + 2d + 1$, и $M(d, k) < P(d, k)$ для $k > 2$.

Получение точных значений функции $M(d, k)$ для $k \geqslant 3$ достаточно трудоёмко и сводится к полному перебору параметрических описаний циркулянтного графа. Нижние оценки для $M(d, k)$ в каждом отдельном случае могут зависеть от рассматриваемого диаметра и, как правило, получаются посредством поиска бесконечных семейств графов, достигающих этих оценок.

В [8] для циркулянтов всех размерностей $k \geqslant 3$ получены нижние оценки функции $M(d, k)$, равные

$$M(d, k) \geqslant n = 2q \sum_{i=0}^{k-1} (4q)^i,$$

где $q = \lfloor (d - k + 3)/k \rfloor$, $d \geqslant k$.

В [8] также приведены найденные компьютерным поиском значения $M(d, 3)$ (и образующих соответствующих графов) для диаметров $2 \leqslant d \leqslant 10$. Отметим, что для $d = 4$ и $d = 10$ указаны не совсем точные значения, но все полученные $M(d, 3)$ являются нечётными числами. На основе этих данных и того факта, что $M(d, 1)$ и $M(d, 2)$ являются нечётными при любых d, авторы выдвинули гипотезу, что значения $M(d, k)$ являются нечётными числами при любых d и k.

Эта гипотеза подтверждена в [9] для $k = 3$ (см. также [10]): найдена экстремальная функция $M(d, 3)$ для любого диаметра d через построение бесконечных семейств трёхмерных циркулянтов с порядком, совпадающим с $M(d, 3)$. Вид этой функции зависит от класса вычетов d по модулю 3, а функция является нечётным числом при любом d. В работе [11] доказано также, что существует граф Кэли абелевой группы с тремя образующими, который имеет диаметр d и размер равный $M(d, 3)$.

В 1994 г. Ф. П. Муга доказал следующую теорему.

Теорема 1 [12]. Значение $M(d, k)$ является нечётным числом при любых $d \geqslant 1$ и $k \geqslant 1$.

В настоящей работе доказано, что доказательство вышеприведенной теоремы из [12] некорректно и соответственно нельзя сделать вывод, что $M(d, k)$ при любых d и k является нечётным числом, что подтверждает опровергающий пример циркулянт-

ного графа размерности четыре. На основании работы [4] скорректирована таблица максимально достижимых порядков циркулянтов размерности четыре, приведённая в [13].

1. Комментарии к теореме 3 из работы [12]

При доказательстве теоремы 3 [12] автор использует следующие рассуждения. Во-первых, перечисляет все пути, ведущие в вершины, которые достижимы из заданной вершины (например, из нуля) самое большее за d шагов. Их число равно $P(d,k)$, которое, как можно заметить, является нечётным при всех d и k. При $k=1$ и $k=2$ эти пути (и соответственно вершины, в которые они ведут) могут быть все различны, как показано ранее. При $k \geqslant 3$ не все пути обязательно ведут в разные вершины. Если некоторые из путей неотличимы, то существуют по крайней мере два пути, которые ведут в одну и ту же вершину. Пусть это будут $\sum\limits_{i=1}^{k} \alpha_i s_i = \sum\limits_{i=1}^{k} \beta_i s_i$. Если $\sum\limits_{i=1}^{k} |\alpha_i| \geqslant \sum\limits_{i=1}^{k} |\beta_i|$, то удаляется больший путь $\sum\limits_{i=1}^{k} \alpha_i s_i$. Тем самым уменьшается на единицу исходный список путей. Но в исходном списке также присутствует другая вершина, в которую ведут пути $\sum\limits_{i=1}^{k} -\alpha_i s_i$ и $\sum\limits_{i=1}^{k} -\beta_i s_i$, получающиеся путем замены на обратные образующие, и они равны. Поэтому необходимо удалить из списка также путь $\sum\limits_{i=1}^{k} -\alpha_i s_i$. После применения этого метода снова до тех пор, пока не останутся все различные пути (вершины), и удаления каждый раз двух путей остаётся нечётное число вершин.

В этом доказательстве рассмотрены не все возможные варианты. Когда автор полагает, что другая вершина, в которую ведут пути $\sum\limits_{i=1}^{k} -\alpha_i s_i$ и $\sum\limits_{i=1}^{k} -\beta_i s_i$, обязательно отлична от рассмотренной, он тем самым неявно предполагает, что в рассматриваемом графе имеется нечётное число вершин, поскольку в случае чётного числа вершин в циркулянтном графе порядка n возможна ситуация, когда рассматриваемая вершина и вершина с путями, образованными заменой на обратные образующие, могут совпадать; например, если это вершина с номером $n/2$ и она находится от нуля на расстоянии, равном диаметру d графа. Таким образом, автор просто не доказал, что такая ситуация не может встретиться в графе с максимально возможным числом вершин, поэтому его доказательство не является корректным. Таким образом, и теорема 3 [12], и её доказательство являются ошибочными.

В [4] найден первый пример, подтверждающий ошибочность рассматриваемой теоремы для циркулянтных графов размерности четыре. Для диаметра $d=3$ найден циркулянтный граф с числом вершин $n=104$ и образующими $S=\{1,16,20,27\}$, который является максимально возможным четырёхмерным циркулянтом диаметра три. Обратим внимание, что при определении экстремальной функции $M(d,k)$ авторы [8] ограничиваются наборами образующих циркулянтного графа, содержащими $s_1=1$ (или образующую, взаимно простую с n). Но нигде не доказано, что другие наборы образующих не могут дать максимальный граф. Поэтому мы дополнительно проверили (на кластере НКС-30Т Сибирского суперкомпьютерного центра) с помощью программы полного перебора параметрических описаний циркулянтных графов все нечётные значения n в диапазоне $105 \leqslant n \leqslant 129$, где $129 = P(3,4)$, и не нашли других графов диаметра три.

2. К вопросу об экстремальной функции $M(d, 4)$

Нижеприведённый результат из [13] улучшает нижнюю оценку экстремальной функции $M(d, 4)$, полученную в работах [8, 11], для любых диаметров $d > 1$.

Теорема 2. Для всех $d \geqslant 2$ существует неориентированный четырёхмерный циркулянтный граф, который имеет диаметр d и порядок равный $n = d^4/2 + d^3 + 5d^2/2 + 2d + 1$. Множество образующих S для данного графа при чётном d есть

$$\{1, d + 1, d^3/2 + d^2/2 + d, d^3/2 + 3d^2/2 + 3d + 2\},$$

при нечётном — $\{1, d, d^3/2 + 3d/2, d^3/2 + d^2 + 3d/2 + 1\}$.

В свою очередь, найденная в [13] оценка функции $M(d, 4)$ улучшена на $\mathrm{O}(d^2/2)$ в [4].

Теорема 3 [4]. Для всех $d \geqslant 2$ существует неориентированный граф Кэли абелевой группы с четырьмя образующими, который имеет диаметр d и размер равный

$$n = \begin{cases} (d^4 + 2d^3 + 6d^2 + 4d)/2, & \text{если } d \equiv 0 \pmod 2, \\ (d^4 + 2d^3 + 6d^2 + 6d + 1)/2, & \text{если } d \equiv 1 \pmod 2. \end{cases}$$

Множество образующих S для данного графа при чётном d есть

$$\{1, (d^3 + 2d^2 + 6d + 2)/2, (d^4 + 4d^2 - 8d)/4, (d^4 + 4d^2 - 4d)/4\},$$

при нечётном — $\{1, (d^3 + d^2 + 5d + 3)/2, (d^4 + 2d^2 - 8d - 11)/4, (d^4 + 2d^2 - 4d - 7)/4\}$.

Таким образом, на сегодняшний день максимально возможными циркулянтными графами размерности четыре и любого диаметра являются графы, полученные в [4].

Основываясь на работе [4], можно скорректировать приведённую в [13] табл. 4 максимально достижимых порядков циркулянтов размерности четыре (в таблице выделены курсивом новые значения порядков графов и их образующие, полученные в [4]).

Описания максимальных циркулянтов диаметра d и размерности 4

d	$n = M(d, 4)$	$S = \{s_1, s_2, s_3, s_4\}$
1	9	{1, 2, 3, 4}
2	35	{1, 6, 7, 10}, {1, 7, 11, 16}
3	*104*	{*1, 16, 20, 27*}
4	*248*	{*1, 61, 72, 76*}
5	*528*	{*1, 89, 156, 162*}
6	*984*	{*1, 163, 348, 354*}

ЛИТЕРАТУРА

1. *Монахова Э. А.* Структурные и коммуникативные свойства циркулянтных сетей // Прикладная дискретная математика. 2011. №3(13). С. 92–115.
2. *Bermond J.-C., Comellas F., and Hsu D. F.* Distributed loop computer networks: a survey // J. Parallel Distributed Comput. 1995. V. 24. P. 2–10.
3. *Hwang F. K.* A survey on multi-loop networks // Theor. Comput. Sci. 2003. No. 299. P. 107–121.
4. *Lewis R. R.* The degree-diameter problem for circulant graphs of degree 8 and 9 // Electron. J. Combinator. `http://web.ArXiv.org/1404.3948.pdf`, 20 April 2014.
5. *Horak P.* Tilings in Lee metric // Eur. J. Combinator. 2009. No. 30. P. 480–489.
6. *Costa S. I. R., Strapasson J. E., Alves M. M. S., and Carlos T. B.* Circulant graphs and tessellations on flat tori // Linear Algebra Applicat. 2010. V. 432. No. 1. P. 369–382.

7. *Golomb S. W. and Welch L. R.* Perfect codes in the Lee metric and the packing of polyominoes // SIAM J. Appl. Math. 1970. V. 18. No. 2. P. 302–317.
8. *Chen S. and Jia X.-D.* Undirected loop networks // Networks. 1993. No. 23. P. 257–260.
9. *Monakhova E. A.* Optimal triple loop networks with given transmission delay: topological design and routing // Inter. Network Optimization Conf. (INOC'2003). Evry/Paris, France. 2003. P. 410–415.
10. *Monakhova E. A.* Triple circulant communication networks of parallel computer systems // Optoelectronics, Instrumentation and Data Processing. N. Y.: Allerton Press Inc., 2006. No. 3. P. 90–101.
11. *Dougherty R. and Faber V.* The degree-diameter problem for several varieties of Cayley graphs, 1: the Abelian case // SIAM J. Discrete Math. 2004. V. 17. No. 3. P. 478–519.
12. *Muga F. P.* Undirected circulant graphs // Inter. Symp. on Parallel Architectures, Algorithms and Networks. IEEE, 1994. P. 113–118.
13. *Монахова Э. А.* О построении циркулянтных сетей размерности четыре с максимальным числом вершин при любом диаметре // Прикладная дискретная математика. 2013. № 3(21). С. 76–85.

2014 Прикладная теория графов № 3(25)

DOI 10.17223/20710410/25/8

УДК 519.171 + 519.175.1

АЛЬТЕРНАТИВНЫЕ ПОДХОДЫ К ОПИСАНИЮ КЛАССОВ ИЗОМОРФНЫХ ГРАФОВ

М. Н. Назаров

Национальный исследовательский университет «МИЭТ», г. Москва, Россия

E-mail: Nazarov-Maximilian@yandex.ru

Предложен алгоритм естественной индексации для классов симметрии вершин и рёбер конечных графов. На основе этой индексации построено альтернативное описание для классов изоморфных графов. Продемонстрировано, что на классы изоморфных графов можно перенести такие классические понятия, как раскраски, подграфы, а также элементарные операции на графах.

Ключевые слова: *изоморфизм графов, классы симметрии вершин, классы симметрии рёбер, инварианты графов.*

Введение

Графом G принято называть пару $G = (V, E)$, где V — множество вершин, а $E \subseteq V \times V$ — множество рёбер, которые задают связи между вершинами. Поскольку основным предметом теории графов является абстрактная структура связей, то названия вершин V и рёбер E, а также их природа в общем случае не играют принципиальной роли. Данный факт находит своё отражение в определении понятия изоморфизма графов (подробнее см. [1 – 3]).

Определение 1. Два графа G и H называются изоморфными, если существует биекция $\varphi : V(G) \to V(H)$, такая, что

$$\forall u, v \in V(G) \quad (u, v) \in E(G) \Leftrightarrow (\varphi(u), \varphi(v)) \in E(H).$$

Два изоморфных графа с точки зрения теории графов считаются неразличимыми, так как имеют одинаковую структуру связей между вершинами. Поскольку отношение изоморфизма разбивает множество[1] всех графов на классы, то можно отождествить класс изоморфных графов $[G]$ с уникальной структурой связей (рис. 1).

$$G = (V, E) \;\; V = \{a, b, c, d, e, f, g\}$$
$$E = \{(a, b), (a, c), (a, f), (a, g), (b, c), (b, d), (b, g), (c, d), (c, f), (d, f), (d, g), (f, g)\}$$

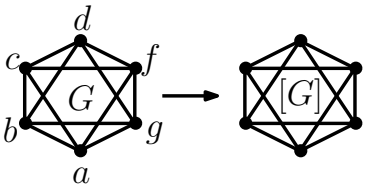


Рис. 1. Переход от обычного графа G к классу изоморфных $[G]$

С понятием изоморфизма графов связано понятие *инварианта* — некоторого набора множеств, которые являются характеристиками структуры графа. У двух изоморфных графов значения инвариантов должны обязательно совпадать, но обратное

[1] Данная терминология корректна только в рамках альтернативных теорий множеств, таких, как NFU (подробнее см. [4]). Для классической теории множеств ZFC принято использовать термин «класс всех графов».

в общем случае не верно. Поэтому дополнительно вводится понятие *полного инварианта*, который совпадает у графов тогда и только тогда, когда они изоморфны (подробнее об инвариантах см. [1, 3]).

Если изоморфизм ψ отображает граф G сам на себя, то такое отображение называется *автоморфизмом*. Две вершины $u, v \in V(G)$ графа G, которые могут быть переведены одна в другую некоторым автоморфизмом $u = \psi(v)$, будем называть *автоморфно эквивалентными* (или симметричными) и обозначать $u \sim v$. При этом граф G можно изобразить так, что две вершины $u \sim v$ переводятся друг в друга геометрической симметрией (на рис. 1 все вершины могут быть переведены друг в друга симметриями, а значит, они все попарно автоморфно эквивалентны). Аналогично определяется отношение автоморфной эквивалентности и для рёбер графа: $(u_1, u_2) \sim (v_1, v_2)$, если найдётся рёберный автоморфизм $\psi : E(G) \to E(G)$, такой, что $(u_1, u_2) = \psi(v_1, v_2)$. Рёберный автоморфизм сохраняет вершины: рёбра e_1 и e_2 имеют общую вершину тогда и только тогда, когда общая вершина есть у рёбер $\psi(e_1)$ и $\psi(e_2)$.

Отметим, что вопрос об автоморфизмах и симметриях интересен не только с теоретической точки зрения, но и с точки зрения приложений теории графов. О роли автоморфизмов графов в теории алгоритмов см. [5, 6]. Дополнительно к этому можно ознакомиться с химическими приложениями графов, а также их автоморфизмов в работах [7 – 9].

Отношения автоморфной эквивалентности разбивают множества вершин $V(G)$ и рёбер $E(G)$ любого графа G на классы симметрии $\overline{v}$ и $(\overline{u, v})$ соответственно. Несмотря на широкий круг приложений для симметрий и автоморфизмов графов, непосредственно изучению классов симметрии вершин $\overline{v}$ посвящено относительно небольшое количество работ. В первую очередь стоит отметить работу M. Спэрроу [10] 1993 г., в которой рассматриваются быстрые алгоритмы поиска классов симметрии, а также диссертационную работу P. Саутвелла [11] 2006 г.

В данной работе мы демонстрируем, что каждому классу автоморфной эквивалентности вершин $\overline{v}$ произвольного графа G можно однозначным образом присвоить порядковый номер $I(\overline{v})$ на основе полного числового инварианта графа — макси-кода. Аналогичное верно и для классов автоморфной эквивалентности рёбер $(\overline{u, v})$: каждому классу можно поставить в соответствие уникальный порядковый номер $I(\overline{u, v})$.

Идентификаторы $I(\overline{v})$ и $I(\overline{u, v})$, в свою очередь, задействованы для построения полного инварианта $I[G]$ — *линейной нотации абстрактного графа* $[G]$. В работе демонстрируется, что полный инвариант $I[G]$ обладает всеми свойствами обычных графов G и, в частности, может быть использован для определения таких классических понятий, как раскраски, подграфы, а также элементарных операций на графах.

При построении инварианта $I[G]$ используется подход, аналогичный алгоритмам построения канонических форм графа (подробнее об этом см. [12]). Можно провести также некоторую аналогию между итоговым инвариантом $I[G]$ и линейной нотацией для молекулярных графов стандарта SMILES [13]. Отметим, что ключевым преимуществом инварианта $I[G]$ по отношению к существующим методам хранения графов является полное описание всех симметрий графа. В частности, это позволяет эффективно реализовать процедуру визуализации графа G с учётом всех возможных симметрий этого графа на плоскости или в пространстве.

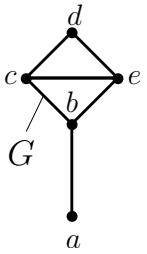
1. Построение индексации для классов симметрии

Определение 2. Пусть дан произвольный конечный граф $G = (V, E)$, у которого число вершин $|V| = n$. Тогда, если выбрать некоторый порядок для множества вершин

$\alpha = (v_1, \ldots, v_n)$, то можно поставить в соответствие графу G матрицу смежности A по стандартному правилу:

$$A(i,j) = 1 \Leftrightarrow (v_i, v_j) \in E \quad \wedge \quad A(i,j) = 0 \Leftrightarrow (v_i, v_j) \notin E.$$

На рис. 2 представлен пример двух матриц смежности, которые построены для разных порядков на множестве вершин графа.



A_1	a	b	c	d	e
a	0	**1**	0	0	0
b	**1**	0	**1**	0	**1**
c	0	**1**	0	**1**	**1**
d	0	0	**1**	0	**1**
e	0	**1**	**1**	**1**	0

A_2	c	e	b	d	a
c	0	**1**	**1**	**1**	0
e	**1**	0	**1**	**1**	0
b	**1**	**1**	0	0	**1**
d	**1**	**1**	0	0	0
a	0	0	**1**	0	0

Рис. 2. Пример двух разных матриц смежности A_1 и A_2 для одного графа

Отметим, что матрицы смежности будут совпадать для тех перестановок вершин, которые задают автоморфизмы графа. В пределе максимально возможное число разных матриц смежности равно $n!$ — числу всех перестановок из n элементов и достигается только у асимметричных графов.

Определение 3. Для произвольного графа G определим *рёберный граф* $L\{G\}$ как граф, чьими вершинами являются рёбра $E(G)$, а связи устанавливаются между теми из них, которые имеют ровно одну общую вершину из $V(G)$:

$$\begin{aligned} L\{G\}: \quad & V(L\{G\}) = E(G), \\ & E(L\{G\}) = \{((u_1,u_2),(v_1,v_2)) : \exists i,j \in \{1,2\} : u_i = v_j \wedge u_j \neq v_i\}. \end{aligned}$$

Пример перехода от графа G к рёберному графу $L\{G\}$ представлен на рис. 3.

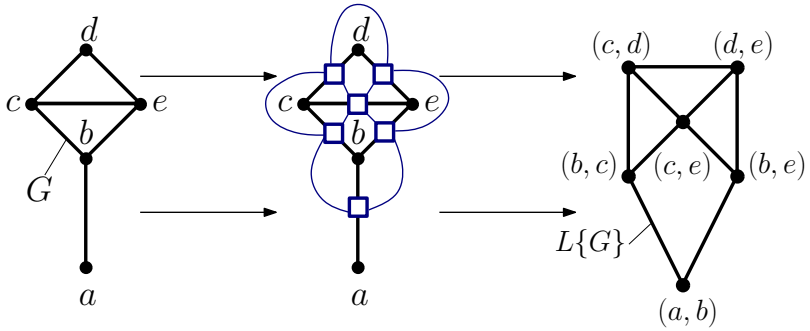


Рис. 3. Переход от G к рёберному графу $L\{G\}$

Замечание 1. Отметим, что любой автоморфизм $\psi : V(L\{G\}) \to V(L\{G\})$ рёберного графа $L\{G\}$ одновременно является рёберными автоморфизмом $\psi : E(G) \to E(G)$ для исходного графа G.

Определение 4. *Кодом матрицы смежности* A конечного графа G будем называть такое число $\mu(A)$, которое получается в результате перевода матрицы смежности в бинарное число $A(1,1)A(1,2)\ldots A(1,n)A(2,1)\ldots A(n,n)$.

На практике для представления числа $\mu(A)$ обычно используются десятичная или шестнадцатеричная системы счисления.

Определение 5. Наибольший из всех возможных кодов матриц смежности графа G назовём *макси-кодом* $\mu_{\max}(G)$. Если для некоторого порядка α вершин код матрицы смежности $\mu(A) = \mu_{\max}(G)$, то будем говорить, что порядок α *соответствует* макси-коду.

Макси-код является полным инвариантом графа, так как по нему можно однозначно восстановить матрицу смежности. Важно отметить, что порядков следования вершин α, которые соответствуют макси-коду, может быть несколько при условии, что у графа есть хотя бы один нетривиальный автоморфизм.

Определение 6. Наибольший из всех возможных кодов матриц смежности рёберного графа $L\{G\}$ назовём *рёберным макси-кодом* $\mu^L_{\max}(G)$.

Утверждение 1. Если два порядка вершин $\alpha_1 = (v^1_1, \ldots, v^1_n)$ и $\alpha_2 = (v^2_1, \ldots, v^2_n)$ соответствуют макси-коду $\mu_{\max}(G)$ графа G, то вершины в этих порядках попарно автоморфны: $v^1_k \sim v^2_k$ для всех $k = 1, \ldots, n$.

Доказательство. Совпадение для двух порядков вершин значения кода $\mu(A)$ возможно лишь в том случае, если совпадают сами матрицы смежности для этих порядков. Это, в свою очередь, означает, что перестановка вершин $\psi = \begin{pmatrix} v^1_1 & \ldots & v^1_n \\ v^2_1 & \ldots & v^2_n \end{pmatrix}$ является автоморфизмом графа G. ■

Утверждение 2. Если два порядка рёбер $\beta_1 = (e^1_1, \ldots, e^1_k)$ и $\beta_2 = (e^2_1, \ldots, e^2_k)$ соответствуют рёберному макси-коду $\mu^L_{\max}(G)$ графа G, то рёбра в этих порядках попарно автоморфны: $e^1_i \sim e^2_i$ для всех $i = 1, \ldots, k$.

Доказательство. Данное утверждение является очевидным следствием утверждения 1, если учесть, что любой автоморфизм рёберного графа $L\{G\}$ является рёберным автоморфизмом графа G. ■

Будем использовать для последовательностей вершин $\alpha = (v_1, \ldots, v_n)$ стандартное индексное обозначение элементов $v_i = \alpha(i)$. Для последовательности рёбер $\beta = (e_1, \ldots, e_n)$ будем использовать аналогичное обозначение $e_i = \beta(i)$.

Определение 7. Пусть макси-коду графа G соответствует некоторый порядок следования вершин α. Назовём *индексом класса симметрии вершин* $\overline{v}$ натуральное число $I(\overline{v})$, равное первому вхождению в порядок α вершины из класса $\overline{v}$. Формально это можно записать так: $I(\overline{v}) = \min\limits_{i:\, v \sim \alpha(i)} i$.

Корректность определения индексов для классов симметрии вершин непосредственно следует из утверждения 1. Действительно, если макси-коду графа соответствуют два варианта последовательностей вершин $\alpha_1 = (v^1_1, \ldots, v^1_n)$ и $\alpha_2 = (v^2_1, \ldots, v^2_n)$, то $v^1_k \sim v^2_k$ для всех $k = 1, \ldots, n$. Как следствие, индекс первого вхождения вершины из класса $\overline{v}$ одинаков в α_1 и α_2.

Определение 8. Пусть рёберному макси-коду $\mu^L_{\max}(G)$ графа G соответствует некоторый порядок следования рёбер β. Назовём *индексом класса симметрии рёбер* $(\overline{u,v})$ натуральное число $I(\overline{u,v})$, равное первому вхождению в β ребра из класса $(\overline{u,v})$, т. е. $I(\overline{u,v}) = \min\limits_{i:\,(u,v) \sim \beta(i)} i$.

По аналогии с вершинами, корректность определения индексов для классов симметрии рёбер есть прямое следствие утверждения 2.

Теорема 1. Если $G \cong H$, то для двух вершин $u \in V(G)$, $v \in V(H)$ условие $I(\overline{u}) = I(\overline{v})$ выполняется тогда и только тогда, когда существует изоморфизм $\varphi : V(G) \to V(H)$, такой, что $\varphi(u) = v$.

Доказательство. Докажем сначала, что из $\varphi(u) = v$, где φ — изоморфизм, следует совпадение индексов классов симметрии $I(\overline{u}) = I(\overline{v})$. Если два графа изоморфны,

то у них один и тот же макси-код $\mu_{\max}(G) = \mu_{\max}(H)$. В этом случае изоморфизм можно представить в виде $\varphi = \begin{pmatrix} u_1 \dots u_n \\ v_1 \dots v_n \end{pmatrix}$, где $\alpha_1 = (u_1, \dots, u_n)$ — некоторый порядок, который соответствует макси-коду в G, а $\alpha_2 = (v_1, \dots, v_n)$ — порядок, соответствующий макси-коду в H. Поскольку индекс $I(\overline{u})$ — натуральное число, $\alpha_1(I(\overline{u}))$ задаёт некоторую вершину $u^\star = \alpha_1(I(\overline{u}))$. Эта вершина $u^\star$ обладает двумя свойствами: $u \sim u^\star$ и среди всех симметричных по отношению к u вершинам $u^\star$ имеет наименьший индекс в α_1. Если допустим, что $I(\overline{u}) \neq I(\overline{v})$, то получим $\varphi(u^\star) \nsim \varphi(u)$. В результате получаем противоречие: $\varphi(u^\star) \nsim \varphi(u)$ и $u \sim u^\star$.

Пусть теперь известно, что $G \cong H$ и совпали индексы $I(\overline{u}) = I(\overline{v})$. Поскольку $G \cong H$, существует некоторый изоморфизм $\varphi_0 : V(G) \to V(H)$. Для образа $v^* = \varphi_0(u)$ по первой части теоремы получим $I(\overline{v^*}) = I(\overline{u}) = I(\overline{v})$. Из этого можно заключить, что вершины v, v^* симметричны: $v \sim v^*$. Обозначим автоморфизм, переводящий эти вершины одна в другую, через ψ: $\psi(v^*) = v$. Тогда искомый изоморфизм φ можно определить в виде композиции $\varphi = \varphi_0 \circ \psi$, а для вершин u, v получим $v = \psi(\varphi_0(u))$. Отображение φ, как композиция двух изоморфизмов, само является изоморфизмом. Теорема доказана. ■

Теорема 2. Если $G \cong H$, то для двух рёбер $(u_1, u_2) \in E(G)$, $(v_1, v_2) \in E(H)$ условие $I(\overline{u_1, u_2}) = I(\overline{v_1, v_2})$ выполняется тогда и только тогда, когда существует изоморфизм $\varphi : V(G) \to V(H)$, такой, что $\varphi(u_1) = v_1$ и $\varphi(u_2) = v_2$, или $\varphi(u_1) = v_2$ и $\varphi(u_2) = v_1$.

Доказательство. Учтём тот факт, что если $G \cong H$, то вершины у рёберных графов $L\{G\}$ и $L\{H\}$ изоморфны тогда и только тогда, когда изоморфны соответствующие им рёбра у графов G и H. В результате, применив для рёберных графов $L\{G\}$ и $L\{H\}$ теорему 1, получим искомое. ■

Следствие 1. Индексы $I(\overline{v})$ естественным образом порождают *линейный порядок* на классах автоморфных вершин: $\overline{v_1} \leqslant \overline{v_2} \Leftrightarrow I(\overline{v_1}) \leqslant I(\overline{v_2})$.

Доказательство. По определению все индексы $I(v)$ являются натуральными числами, а согласно утверждению 1, индексы могут совпадать только для симметричных вершин $v \sim u$. Отсюда получаем, что множество всех индексов графа $I(\overline{v_1}), \dots, I(\overline{v_k})$ линейно упорядочено и, следовательно, порядок на классах симметрии $\overline{v_1} \leqslant \overline{v_2} \Leftrightarrow I(\overline{v_1}) \leqslant I(\overline{v_2})$ также линейный. ■

Следствие 2. Индексы $I(\overline{u, v})$ порождают *линейный порядок* на классах автоморфных рёбер: $(\overline{u_1, v_1}) \leqslant (\overline{u_2, v_2}) \Leftrightarrow I(\overline{u_1, v_1}) \leqslant I(\overline{u_2, v_2})$.

Доказательство. Полностью аналогично предыдущему доказательству. ■

Для демонстрации изложенных результатов на рис. 4 приводятся примеры графов, чьи вершины маркированы индексами $I(\overline{v})$, а рёбра — индексами $I(\overline{u, v})$.

2. Построение инвариантов на основе индексов классов симметрии

Ранее продемонстрировано, что индексы $I(\overline{v})$ являются уникальными идентификаторами для классов симметрии вершин, а индексы $I(\overline{u, v})$ — для классов симметрии рёбер. Таким образом, индексы $I(\overline{v})$ можно использовать вместо классов вершин $\overline{v}$, а $I(\overline{u, v})$ — вместо классов рёбер $(\overline{u, v})$ при составлении инвариантов графов. Рассмотрим три примера таких инвариантов: мультимножества абстрактных вершин $[V]$, абстрактных рёбер $[E]$ и полный инвариант $I[G]$ — линейную нотацию абстрактного графа.

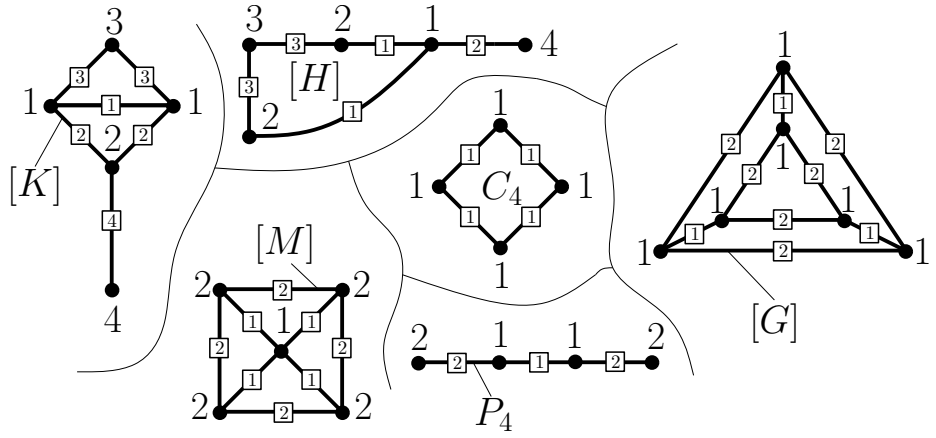


Рис. 4. Графы с вершинами и рёбрами, помеченными $I(\overline{v})$ и $I(\overline{u,v})$

Определение 9. *Мультимножеством абстрактных вершин* графа G назовём такое $[V]$, которое получается из множества $V(G)$ путём замены всех вершин на индексы $I(\overline{v})$. Формально это определение можно выразить так:

$$[V](G) = \{(i) \times n_i : \exists v \in V(G)\,(I(\overline{v}) = i, |\overline{v}| = n_i)\}\,.$$

Замечание 2. Тот факт, что мультимножество $[V](G)$ есть инвариант графа G, непосредственно следует из теоремы 1.

Для примеров классов изоморфных графов, которые изображены на рис. 4, получим следующие мультимножества абстрактных вершин:
$[V](K) = \{(1) \times 2, (2) \times 1, (3) \times 1, (4) \times 1\}$, $[V](H) = \{(1) \times 1, (2) \times 2, (3) \times 1, (4) \times 1\}$, $[V](M) = \{(1) \times 1, (2) \times 4\}$, $[V](C_4) = \{(1) \times 4\}$, $[V](P_4) = \{(1) \times 2, (2) \times 2\}$, $[V](G) = \{(1) \times 6\}$.

По аналогии с $[V]$ можно от обычных рёбер перейти к абстрактным $[E]$. Для этого используем совместно индексы вершин $I(\overline{v})$ с индексами рёбер $I(\overline{u,v})$.

Определение 10. *Мультимножеством абстрактных рёбер* графа G назовём такое $[E]$, которое получается из множества $E(G)$ путём замены всех рёбер $(u,v) \in E(G)$ на тройки $[I(\overline{u,v}) \triangleleft (I(\overline{u}), I(\overline{v}))]$. Формально это определение можно выразить следующим образом:

$$[E](G) = \{[k \triangleleft (i,j)] \times n_{ij} : \exists (u,v) \in E(G)\,(I(\overline{u,v}) = k,\ I(\overline{u}) = i,\ I(\overline{v}) = j,\ |\overline{(u,v)}| = n_{ij}\}.$$

Замечание 3. Тот факт, что мультимножество $[V](G)$ есть инвариант графа G, непосредственно следует из теоремы 2.

Для классов изоморфных графов, которые изображены на рис. 4, получим следующие мультимножества абстрактных рёбер:
$[E](K) = \{[1 \triangleleft (1,1)] \times 1,\ [2 \triangleleft (1,2)] \times 2,\ [3 \triangleleft (1,3)] \times 2,\ [4 \triangleleft (2,4)] \times 1\}$,
$[E](H) = \{[1 \triangleleft (1,2)] \times 2,\ [2 \triangleleft (1,4)] \times 1,\ [3 \triangleleft (2,3)] \times 2\}$,
$[E](M) = \{[1 \triangleleft (1,2)] \times 4,\ [2 \triangleleft (2,2)] \times 4\}$, $[E](C_4) = \{[1 \triangleleft (1,1)] \times 4\}$,
$[E](P_4) = \{[1 \triangleleft (1,1)] \times 1,\ [2 \triangleleft (1,2)] \times 2\}$, $[E](G) = \{[1 \triangleleft (1,1)] \times 3,\ [2 \triangleleft (1,1)] \times 6\}$.

Отметим, что абстрактные вершины $[V]$ вместе с абстрактными рёбрами $[E]$ не задают граф G с точностью до изоморфизма. Рассмотрим ещё один инвариант графа $I[G]$, который построен на основе линейной нотации для графа G. При определении инварианта $I[G]$ воспользуемся техникой, которая в значительной степени аналогична алгоритмам построения канонических форм графа [12] и линейных нотаций для молекулярных графов SMILES [13].

Определение 11. *Симметрическая линейная нотация* $\mathfrak{L}(G)$ для графа G — это строка символов, которая определяется на основе четырёх правил.

П р а в и л о 1. Строка $\mathfrak{L}(G)$ начинается с вершины v первого индекса: $I(v) = 1$.

П р а в и л о 2. Для всех вершин $u \in V(G)$ в нотацию $\mathfrak{L}(G)$ вместе с вершиной u обязательно вводится её окружение в виде записи $u[\xrightarrow{i_1} \ldots; \xrightarrow{i_2} \ldots; \xrightarrow{i_k} \ldots]$, где $i_1, i_2, \ldots, i_k$ — индексы классов симметрии рёбер. При этом в пределах скобок $u[\xrightarrow{i_1} \ldots; \xrightarrow{i_2} \ldots; \xrightarrow{i_k} \ldots]$ **ровно один раз** должны быть учтены все вершины из окружения u и все рёбра $\xrightarrow{i}$, которые их связывают, но не обязательно на одном и том же уровне вложенности скобок (то есть $k \leqslant |\{u^* : (u, u^*) \in E\}|$).

П р а в и л о 3. При повторном появлении на последующих уровнях вложенных скобок $[\ldots[\ldots]\ldots]$ уже встречавшихся вершин u вместо них используются специальные символы $\#1, \#2, \#3, \ldots, \#m$. Обозначение $\#1$ задаёт вершину v перед первой скобкой [в нотации $\mathfrak{L}(G)$; запись $\#2$ означает вершину перед второй скобкой, и так далее до последнего уровня вложенности $\#m$.

П р а в и л о 4. Для каждой вершины u при выборе, в какой последовательности записывать в скобках $u[\ldots;\ldots;\ldots]$ вершины $\xrightarrow{i_1} v_1 \ldots$, $\xrightarrow{i_2} v_2 \ldots$ и специальные символы $\xrightarrow{i} \#m$ из окружения u, действует следующий набор приоритетов:

1) Наиболее приоритетным является код $\#1$, затем код $\#2$, и так далее до последнего уровня вложенности $\#m$.
2) После кодов $\#m$ в нотации идут вершины с наименьшими индексами классов симметрии $I(\overline{v^*})$.
3) Если у двух вершин совпадают индексы $I(\overline{v_1}) = I(\overline{v_2})$, то более приоритетной будет избрана такая вершина v_1, у которой меньше индекс симметрии ребра, соединяющего её с u, то есть $I(\overline{u, v_1}) < I(\overline{u, v_2})$.
4) При совпадении индексов вершин и рёбер предпочтение будет отдано вершине, которая ближе на графе G к вершине кода $\#1$. Если эти расстояния совпадают, то ближе к вершине кода $\#2$, и так далее. Близость здесь понимается в смысле существования пути с меньшим количеством вершин, соединяющего эти две вершины.

Определение 12. *Линейной нотацией абстрактного графа* $[G]$ назовём такую строку $I[G]$, которая получается из любой произвольной симметрической линейной нотации $\mathfrak{L}(G)$ путём замены всех вершин v на индексы классов симметрии $I(\overline{v})$.

Для примеров классов изоморфных графов, которые изображены на рис. 4, можем составить следующие линейные нотации:

$$I[K] = 1\left[\xrightarrow{1} 1\left[\xrightarrow{2} 2\left[\xrightarrow{2} \#1;\ \xrightarrow{4} 4[\,]\right];\ \xrightarrow{3} 3\left[\xrightarrow{3} \#1\right]\right]\right],$$

$$I[H] = 1\left[\xrightarrow{1} 2\left[\xrightarrow{3} 3\left[\xrightarrow{3} 2\left[\xrightarrow{1} \#1\right]\right]\right];\ \xrightarrow{2} 4[\,]\right],$$

$$I[M] = 1\left[\xrightarrow{1} 2\left[\xrightarrow{2} 2\left[\xrightarrow{1} \#1;\ \xrightarrow{2} 2\left[\xrightarrow{1} \#1;\ \xrightarrow{2} 2\left[\xrightarrow{1} \#1;\ \xrightarrow{2} \#2\right]\right]\right]\right]\right],$$

$$I[C_4] = 1\left[\xrightarrow{1} 1\left[\xrightarrow{1} 1\left[\xrightarrow{1} 1\left[\xrightarrow{1} \#1\right]\right]\right]\right], \quad I[P_4] = 1\left[\xrightarrow{1} 1\left[\xrightarrow{2} 2\,[\,]\right];\ \xrightarrow{2} 2[\,]\right],$$

$$I[G] = 1\left[\xrightarrow{1} 1\left[\xrightarrow{2} 1\left[\xrightarrow{1} 1\left[\xrightarrow{2} \#1;\ \xrightarrow{2} 1\left[\xrightarrow{2} \#1;\ \xrightarrow{1} 1\left[\xrightarrow{2} \#2;\ \xrightarrow{2} \#3\right]\right]\right]\right]\right]\right].$$

Определение 13. *Раскраской для линейной нотации* $I[G]$ отдельной абстрактной *вершины* j *в цвет* α будем называть такую строку $I_{(j,\alpha)}[G]$, которая получается из линейной нотации $I[G]$ путём замены первого вхождения j на пару $\langle j, \alpha\rangle$. Раскраску для большего количества вершин и большего количества цветов можно определить индуктивно как последовательную раскраску графа по одной отдельной вершине.

При этом достаточно потребовать, чтобы на каждом шаге раскраски выбирались только те вершины, которые не были раскрашены на предыдущих.

По аналогии можно определить и раскраску для абстрактных рёбер.

Теорема 3. Линейная нотация абстрактного графа $I[G]$ является полным инвариантом для любого связного графа G.

Доказательство. В первую очередь покажем, что $I[G]$ есть инвариант. При построении $I[G]$ единственная ситуация, когда возможна многозначность в определении — это выбор для $u[\ldots;\ldots]$ порядка следования вершин v_1 и v_2 по правилу 4 определения 12 в случае, если их не удалось различить по четырём критериям приоритетности. Фактически это означает, что у двух вершин v_1 и v_2 совпали индексы симметрии вершин $I(\overline{v_1}) = I(\overline{v_2})$ и рёбер $I(\overline{u,v_1}) = I(\overline{u,v_2})$ и количество вершин, отделяющих v_1 и v_2 от вершин, кодированных $\#1,\ldots,\#m$, одинаково. В этом случае коды $\#1,\ldots,\#m$ стоят в скобках $v_1[\ldots]$ и $v_2[\ldots]$ на одних и тех же уровнях вложенности, а сами вершины $v_1 \sim v_2$ и рёбра $(u,v_1) \sim (u,v_2)$ автоморфны по теоремам 1 и 2. Таким образом, после замены вершин на индексы симметрии итоговые скобки $I(v_1)[\ldots]$ и $I(v_2)[\ldots]$ будут абсолютно одинаковыми. В результате определение $I[G]$ не зависит от порядка, в котором выбираются вершины v_1 и v_2, а, следовательно, $I[G]$ является инвариантом графа.

Докажем, что любой связный граф G однозначно восстанавливается на основе линейной нотации $I[G]$ с точностью до изоморфизма.

Если у графа n вершин, то для восстановления G достаточно покрасить абстрактные вершины нотации $I[G]$ в n разных цветов $\alpha_1,\ldots,\alpha_n$. Затем, заменив покрашенные вершины на их краски $\alpha_1,\ldots,\alpha_n$, получим линейную нотацию для графа $\mathfrak{L}(G^*)$, по которой легко восстановить $G^* = (V^*, E^*)$, где $V^* = \{\alpha_1,\ldots,\alpha_n\}$.

Изоморфизм между графом G и G^* можно построить в общем виде, отобразив все вершины, которые встречаются в нотации $\mathfrak{L}(G^*)$, подряд одну за другой в вершины из нотации $\mathfrak{L}(G)$. Если допустить, что у такого отображения не сохраняются рёбра, то это будет противоречить правилу 4 построения симметрической линейной нотации. В результате получаем, что $I[G]$ — это полный инвариант графа. ■

Замечание 4. Линейную нотацию для абстрактного графа $I[G]$ можно использовать как альтернативу для обычных графов G. Например, проверка двух графов на изоморфизм сведётся для данного случая к тривиальной проверке совпадения двух множеств: $G_1 \cong G_2 \Leftrightarrow I[G_1] = I[G_2]$. Отметим также, что мультимножества $[V]$ и $[E]$ можно легко определить на основе $I[G]$ без участия обычных графов G.

3. Определение операций на классах изоморфных графов через индексы классов симметрии вершин и рёбер

В теории графов особую роль играют такие операции на графах, которые можно однозначным образом перенести на классы изоморфных графов. В общем случае произвольная операция на графах переносима на классы изоморфных графов в том и только в том случае, если отношение изоморфизма графов является конгруэнцией для этой операции. С классическими примерами подобных операций можно ознакомиться в [1, 14].

Определение 14. Операция *удаления абстрактного ребра* $[i \lhd (m,k)]$ для абстрактного графа $[G]$ может быть определена на основе стандартной операции удаления ребра у обычных графов. Для этого достаточно в графе G выбрать любые две вершины $(v_1,v_2) \in E(G)$, такие, что $I(\overline{v_1,v_2}) = i$, $I(\overline{v_1}) = m$ и $I(\overline{v_2}) = k$, и удалить

ребро (v_1, v_2) из множества рёбер $E(G)$. Более формально эту операцию $[G] \setminus [i \triangleleft (m, k)]$ можно определить следующим образом:

$$[\tilde{G}]{=}[G] \setminus [i \triangleleft (m, k)] \Leftrightarrow \tilde{G} = G \setminus (v_1, v_2) : I(\overline{v_1}) = m, I(\overline{v_2}) = k, I(\overline{v_1, v_2}) = i, (v_1, v_2){\in}E(G).$$

Замечание 5. Для идентификации абстрактного ребра достаточно одного индекса $I(\overline{v_1, v_2})$. Добавление индексов вершин $I(\overline{v})$ и $I(\overline{u})$ служит лишь для построения определения по аналогии с обычной операцией удаления ребра.

На рис. 5 приводятся примеры использования операции удаления абстрактного ребра на графах, чьи вершины и рёбра помечены индексами классов симметрии $I(\overline{v})$ и $I(\overline{u, v})$ соответственно.

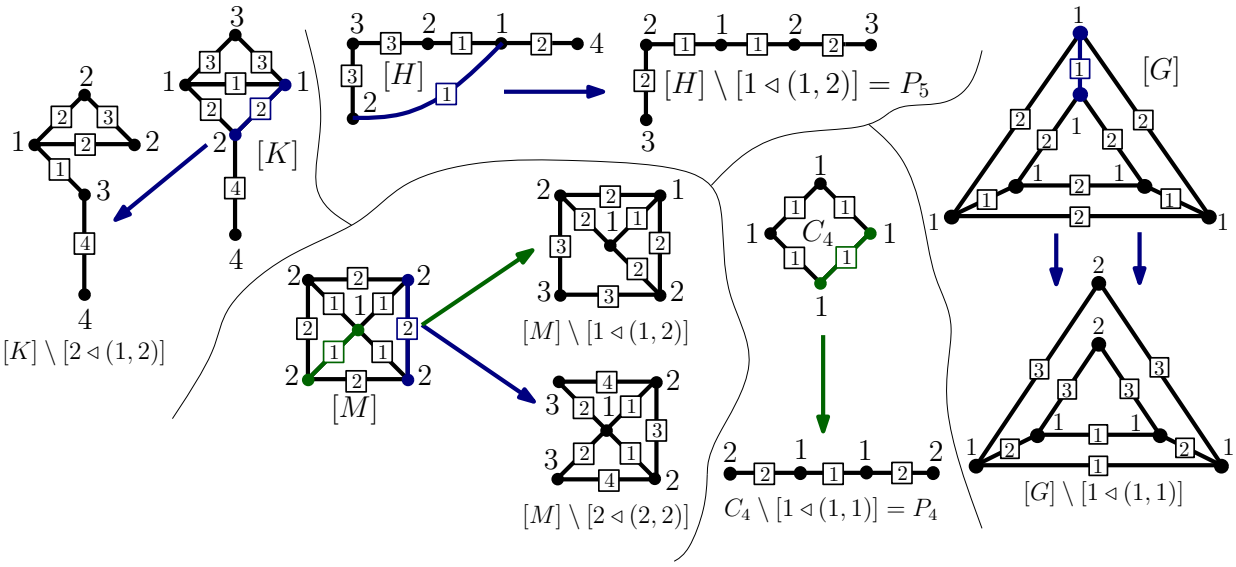


Рис. 5. Примеры использования операции удаления абстрактного ребра

Замечание 6. Актуальной задачей является определение операции удаления ребра $[G] \setminus [i \triangleleft (m, k)]$ без использования для промежуточного представления обычных графов G, а с помощью одной только линейной нотации $I[G]$.

Проверим, что определение операции $[G] \setminus [i \triangleleft (m, k)]$ корректно. Пусть $G_1 \cong G_2$ и выбраны $\tilde{G}_1 = G_1 \setminus (u_1, u_2)$, $\tilde{G}_2 = G_2 \setminus (v_1, v_2)$, где у вершин попарно совпадают индексы $I(\overline{u_1, u_2}) = I(\overline{v_1, v_2})$, $I(\overline{u_1}) = I(\overline{v_1})$, $I(\overline{u_2}) = I(\overline{v_2})$, а сами вершины соединены рёбрами: $(u_1, u_2) \in E(G_1)$ и $(v_1, v_2) \in E(G_2)$.

Из теоремы 2 следует, что найдётся изоморфизм $\varphi : V(G_1) \to V(G_2)$, для которого либо $\varphi(u_1) = v_1$ и $\varphi(u_2) = v_2$, либо $\varphi(u_2) = v_1$ и $\varphi(u_1) = v_2$. Из существования изоморфизма φ по [11, теорема 2.3.1] следует, что $\tilde{G}_1 \cong \tilde{G}_2$.

Замечание 7. Полностью аналогично можно определить и операцию удаления абстрактной вершины $[G] \setminus \{i\}$ через операцию удаления вершины для обычных графов $G \setminus \{v\}$, полагая индекс класса симметрии $I(\overline{v}) = i$.

Определение 15. Два подграфа H_1 и H_2 графа G будем называть *симметричными* ($H_1 \sim H_2$), если они изоморфны и найдётся автоморфизм ψ на G, такой, что $\psi(V(H_1)) = V(H_2)$.

Отношение симметричности для подграфов обобщает понятие симметричности для вершин и рёбер. С геометрической точки зрения это отношение означает, что граф G

можно изобразить так, что два симметричных подграфа $H_1 \sim H_2$ переводятся один в другой с помощью некоторой симметрии графа G.

В работе [11] исследованы свойства отношения симметричности подграфов, в частности, доказана теорема 2.3.1, согласно которой $H_1 \sim H_2$ тогда и только тогда, когда $G \setminus H_1 \cong G \setminus H_2$ и $H_1 \cong H_2$.

Если взять индексы $I(\overline{v})$ и $I(\overline{u,v})$ от графа G и применить их для раскраски вершин и рёбер подграфа $H \subseteq G$, то можно добиться однозначного позиционирования подграфа в графе G с точностью до автоморфизма, что, в свою очередь, эквивалентно полному описанию для симметричных подграфов.

Определение 16. Графом с *раскрашенными вершинами и рёбрами* назовём тройку (H, f, g), где $H = (V, E)$ — обычный граф, а $f : V(H) \to \mathbb{N}$ и $g : E(H) \to \mathbb{N}$ — функции раскраски.

Раскрасив вершины подграфа $H \subseteq G$ индексами $I_G(v)$, а рёбра — индексами $I_G(u, v)$ от большего графа G, получим описание для класса симметричных подграфов. Поскольку данное описание является альтернативой определению класса симметрии, будем использовать для него другое название.

Определение 17. Назовём *классом топологической эквивалентности* графа G по (H, f, g) такое множество $G/_{(H,f,g)}$ подграфов:

$$G/_{(H,f,g)} = \big\{\tilde{H} \subseteq G : \exists \varphi \in ISO(\tilde{H}, H) \forall u, v \in V(\tilde{H}) \\ (f(\varphi(v)) = I_G(\overline{v}),\ g(\varphi(u), \varphi(v)) = I_G(\overline{u,v}))\big\}.$$

В данном определении запись $\varphi \in ISO(\tilde{H}, H)$ обозначает изоморфизм φ графа $\tilde{H}$ на граф H.

Пример позиционирования подграфов $H \subseteq G$ с помощью раскраски его вершин и рёбер индексами $I_G(v)$ и $I_G(u, v)$ основного графа G показан на рис. 6.

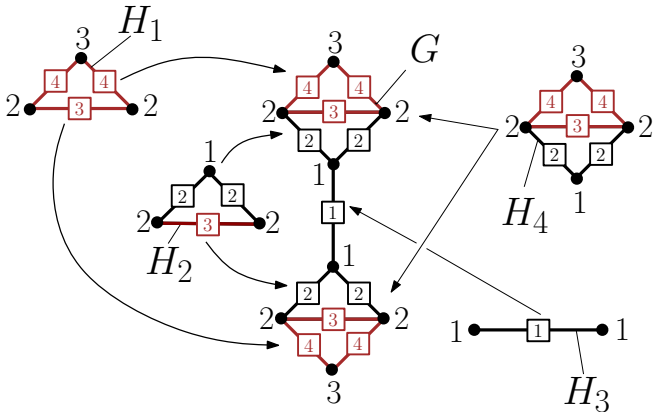


Рис. 6. Пример позиционирования подграфов на основе $G/_{(H,f,g)}$

Нетрудно показать, что подграфы из одного класса $G/_{(H,f,g)}$ являются симметричными; более того, набор $G/_{(H,f,g)}$ в точности совпадает с классами симметрии подграфов. Данный факт позволяет определить операции на абстрактных графах через подграфы, маркированные индексами классов $I(v)$ и $I(u, v)$. В частности, можно определить такие операции, как удаление подграфа или стягивание подграфа в одну вершину.

Замечание 8. Поскольку раскраска определена и для линейных нотаций абстрактных графов $I[G]$, то исследование симметричных подграфов $[H]$ можно проводить, не привлекая обычные графы G, а пользуясь только раскрашенными инвариантами $I[G]$.

Заключение

При построении инвариантов графов на основе индексов $I(v)$ и $I(u,v)$ особенно интересными представляются полные инварианты $I[G]$, которые можно рассматривать как альтернативное описание для классов изоморфных графов $[G]$. Поскольку для полного инварианта $I[G]$ получилось определить процедуру раскраски, это позволяет работать с подграфами для $[G]$, а также реализовать алгоритм восстановления обычного графа на основе $I[G]$.

Одной из важных проблем для практического приложения графов является возможность построения полиномиального алгоритма проверки изоморфизма графов [15]. Поскольку для $I[G]$ удалось определить полиномиальный алгоритм раскраски, $I[G]$ можно использовать на практике в качестве альтернативного способа кодирования классических графов G. В частности, проверка двух графов на изоморфизм сведётся для данного случая к тривиальной проверке совпадения двух множеств: $G_1 \cong G_2 \Leftrightarrow$ $\Leftrightarrow I[G_1] = I[G_2]$. Таким образом, задача проверки на изоморфизм имеет сложность $\mathrm{O}(n)$ для любых связных графов из n вершин. Однако подобный выигрыш в производительности достигается лишь в том случае, если графы уже хранятся в памяти в альтернативной кодировке $I[G]$, поскольку задачи поиска макси-кода $\mu_{\max}(G)$ и формирования инварианта $I[G]$ на основе графа G относятся к классу NP.

Полученные результаты можно коротко представить тремя пунктами:

1) переход от G к $I[G]$ — задача NP-класса;
2) проверка на изоморфизм для $I[G]$ — задача сложности $\mathrm{O}(n)$;
3) переход от $I[G]$ к G на основе раскраски — задача сложности $\mathrm{O}(n^2)$.

Описание на основе $I[G]$ будет особенно ценным с практической точки зрения, если для $I[G]$ удастся определить операции полиномиальной сложности на классах изоморфных графов без использования обычных графов G. Если это окажется возможным, то это потенциально позволит конструировать произвольные абстрактные графы на основе базовых классов графов $[G]$ с помощью некоторого набора операций для $I[G]$. В качестве базовых классов абстрактных графов можно рассматривать, например, пустые графы O_n и полные графы K_n, а также цепи C_n и пути P_n.

Требование полиномиальной сложности для реализаций этих операций на классах изоморфных графов является принципиальным для практики, потому что иначе они не дадут существенного выигрыша по сравнению с определениями, требующими вычисления инварианта $I[G]$ для итогового графа (NP-задачей). Другая важная проблема, для которой $I[G]$ мог бы оказаться удобной альтернативой, — это проблема восстановления графов по подграфам [16].

Рассмотренные в работе конструкции можно обобщить на упорядоченные графы, а также на гиперграфы. Поскольку бинарные группоиды можно определить с помощью гиперграфов, порядок для классов симметрии графов можно перенести на классы автоморфизма элементов группоидов. В частности, это означает, что можно потенциально найти альтернативное описание для класса изоморфных группоидов.

Нужно отметить, что способ индексации $I(v)$ на основе макси-кода является отнюдь не единственно возможным. Очевидно, что похожую индексацию обратного порядка $U(v)$ можно легко получить для мини-кода графа (минимального кода $\mu(A)$ для

матриц смежности A графа). Вполне возможно, что существуют и другие варианты индексаций, которые могут оказаться потенциально удобнее для определения раскрасок и операций на классах изоморфных графов.

Автор выражает благодарность И. Б. Кожухову, Н. В. Суворовой и А. В. Решетникову за критические замечания и помощь в работе над статьёй.

ЛИТЕРАТУРА

1. *Diestel R.* Graph Theory. 3rd edition. Heidelberg: Springer Verlag, 2005. 451 p.
2. *Белов В. В., Воробьев Е. М., Шаталов В. Е.* Теория графов. М.: Высшая школа, 1976. 391 с.
3. *Зыков А. А.* Основы теории графов. М.: Наука, 1987. 383 с.
4. *Holmes M. R.* Elementary Set Theory with a Universal Set. Louvain-la-Neuve: Academia, 1998. 241 p.
5. *Aloul F. A., Ramani A., Markov I. L., and Sakallah K. A.* Solving difficult instances of Boolean satisfiability in the presence of symmetry // IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems. 2003. V. 22. No. 9, P. 1117–1137.
6. *Darga P., Sakallah K., and Markov I. L.* Faster symmetry discovery using sparsity of symmetries // Proc. 45st Design Automation Conference. N. Y.: ACM, 2008. P. 149–154.
7. *Фларри Р.* Группы симметрии. Теория и химические приложения. М.: Мир, 1983. 400 с.
8. *Bonchev D. and Rouvray D. H.* Chemical Graph Theory: Introduction and Fundamentals. N. Y.: Gordon and Breach science publishers, 1991. 300 p.
9. *Кинг Р.* Химические приложения топологии и теории графов. М.: Мир, 1987. 560 с.
10. *Sparrow M. K.* A linear algorithm for computing automorphic equivalence classes: the numerical signatures approach // Social Networks. 1993. V. 15. P. 151–170.
11. *Southwell R.* Finding Symmetries in Graphs. Ph. D. thesis. University of York, UK, 2006. 109 p.
12. *Katebi H., Sakallah K., and Markov I. L.* Graph symmetry detection and canonical labelling: differences and synergies // Proc. Turing-100, EPIC 2012. V. 10. P. 181–195.
13. *Weininger D., Weininger A., and Weininger J.* SMILES. 2. Algorithm for generation of unique SMILES notation // J. Chem. Inf. Comput. Sci. 1989. V. 29. No. 2, P. 97–101.
14. *Харари Ф.* Теория графов. М.: КомКнига, 2006. 296 с.
15. *Kobler J., Schoning U., and Toran J.* The Graph Isomorphism Problem: its Structural Complexity. Berlin: Birkhauser, 1993. 160 p.
16. *Harary F.* A survey of the reconstruction conjecture // Graphs and Combinatorics. Lecture Notes in Mathematics. 1974. V. 406. P. 18–28.

DOI 10.17223/20710410/25/9

УДК 519.17

ОБ ОДНОМ КОНТРПРИМЕРЕ ДЛЯ Т-НЕПРИВОДИМЫХ РАСШИРЕНИЙ СВЕРХСТРОЙНЫХ ДЕРЕВЬЕВ

Д. Ю. Осипов

Саратовский государственный университет им. Н. Г. Чернышевского, г. Саратов, Россия

E-mail: st_hill@mail.ru

Т-неприводимым расширением графа G называется его расширение, получаемое из тривиального расширения данного графа удалением максимально возможного набора добавленных при построении тривиального расширения рёбер. Рассматривается один из способов построения оптимального расширения графа — Т-неприводимое расширение (ТНР). Приводится контрпример для схемы из работы Ф. Харари и М. Хурума «One node fault tolerance for caterpillars and starlike trees», которая описывает построение одного ТНР для произвольного сверхстройного дерева. Рассматривается способ построения всех неизоморфных ТНР для подкласса сверхстройных деревьев — равнолучевых звезд.

Ключевые слова: *граф, Т-неприводимое расширение, сверхстройные деревья, равнолучевые звезды.*

Все понятия и определения соответствуют понятиям и определениям в [1].

Определение 1. Расширением n-вершинного графа G называется граф H с $n+1$ вершинами, такой, что граф G вкладывается в каждый максимальный подграф графа H.

Простейшим примером расширения графа G является его тривиальное расширение — соединение графа G с одноэлементным графом (т. е. к графу G добавляется вершина, которая соединяется ребром с каждой вершиной графа G).

Понятие расширения графа тесно связано с вопросами отказоустойчивости дискретных систем. Если граф G рассматривать как функциональную модель некоторого устройства Σ, то расширение H графа G можно воспринимать как схему отказоустойчивой реализации этого устройства: при отказе любого элемента (что истолковывается как удаление из H соответствующей вершины и всех связанных с нею рёбер) в неповреждённой части обнаруживается работоспособная модель для Σ.

При таком подходе естественно возникает вопрос об оптимальности отказоустойчивой реализации для данной системы, т. е. о получении такого расширения H графа G, которое не содержало бы «лишних» ребер. Один из способов — конструкция минимального расширения графа [2], другой — конструкция его Т-неприводимого расширения [3].

Определение 2. Минимальным расширением графа G называется его расширение с минимальным количеством рёбер.

В общем случае при построении минимального расширения возникает необходимость добавлять рёбра в исходный граф, т. е. менять всю систему, моделируемую этим графом. Но иногда технически важно найти решение следующей задачи: построить оптимальное расширение данного графа, сохраняя его первоначальную конструкцию (т. е. не меняя связей внутри него). Существует следующая процедура:

— построить тривиальное расширение исходного графа;

— удалять из полученного графа рёбра до тех пор, пока выполняется свойство расширения.

Полученные графы назовем THP графа G. Для произвольного графа количество неизоморфных THP неизвестно.

Определение 3. Деревом называется связный граф, в котором нет циклов.

Определение 4 [2]. Дерево называется сверхстройным, если в точности одна его вершина имеет степень больше 2. Эту вершину будем называть корнем сверхстройного дерева.

Сверхстройное дерево можно рассматривать как объединение k цепей с общей концевой вершиной. При этом дерево можно закодировать вектором, состоящим из длин цепей в порядке невозрастания: $(m_1, \ldots, m_k)$, где $m_1 \geqslant \ldots \geqslant m_k$. Очевидно, что такое кодирование сверстройных деревьев при $k > 2$ является взаимно однозначным.

Определение 5. Вершина v_{ij} сверхстройного дерева T (i — номер цепи сверхстройного дерева (цепи пронумерованы в некотором порядке); j — номер вершины i-й цепи, нумерация начинается с 1 от корня) называется сложной, если среди длин цепей дерева T нет цепи длины $j-1$ или $m_i - j$, где $i = 1, \ldots, k$; $m_i > 1$; $j = 2, \ldots, m_i$.

До сих пор остаётся нерешённой следующая задача: построить все неизоморфные THP для произвольного сверхстройного дерева. Однако попытка построить одно из THP для произвольного сверхстройного дерева описывается в [4].

В соответствии со схемой из [4] для построения одного из THP произвольного сверхстройного дерева необходимо:

— добавить новую вершину к исходному графу;
— соединить ребром добавленную вершину с корнем и со всеми листьями исходного сверхстройного дерева;
— если в исходном сверхстройном дереве нет сложных вершин, то полученный граф и является искомым THP. Если есть некоторая сложная вершина v_{ij}, то соединить ребром добавленную вершину и вершину v_{ij-1}. Так поступаем для всякой сложной вершины.

Отметим, что ранее в работе [5] уже приводились контрпримеры для утверждения из [4], что минимальное вершинное 1-расширение сверхстройного дерева с k цепями и p сложными вершинами содержит в точности $k + p + 1$ дополнительных ребер.

Рассмотрим сверхстройное дерево (3, 2, 2) (граф G на рис. 1). У данного сверхстройного дерева только одна сложная вершина — вершина v_{12}. Построим THP для данного сверхстройного дерева по описанной схеме (граф G' на рис. 2).

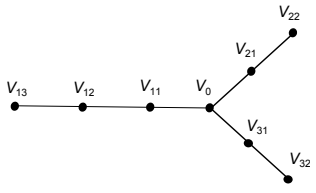


Рис. 1. Граф G

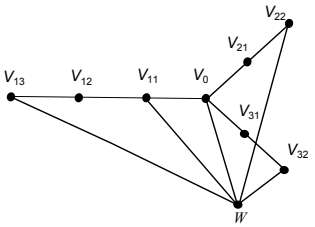


Рис. 2. Граф G'

Нетрудно заметить, что граф G' не является THP для сверхстройного дерева G. На самом деле THP для графа G является граф H (рис. 3).

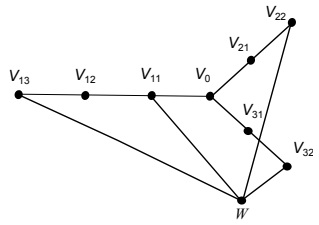


Рис. 3. Граф H

Графы G' и H отличаются на одно ребро. Докажем, что граф H действительно является THP для графа G, представленного на рис. 1.

Легко проверяется, что граф H является расширением графа G. Докажем свойство неприводимости графа H, т. е. что при удалении любого ребра из H полученный граф не будет расширением графа G. Очевидно, при удалении ребра wv_6 (или wv_8, или wv_1) полученный граф не будет расширением для графа G, так как достаточно удалить из полученного графа вершину v_5 (или v_7, или v_2 соответственно), чтобы получить вершину v_6 (или v_8, или v_1 соответственно) степени 0, чего не может быть в графе G. При удалении ребра wv_3 достаточно удалить вершину v_4. В этом случае центром графа будет вершина w, а вершина v_8 будет иметь степень 1, но так как вершина v_8 смежна только с w, то мы не сможем получить двухвершинную цепь, а, следовательно, в полученный граф не вкладывается граф G. Таким образом, никакой граф, полученный из H удалением ребра, не является расширением для G, и свойство неприводимости графа H доказано. Следовательно, граф H является THP для графа G.

Граф G — сверхстройное дерево с наименьшим числом вершин, которое является контрпримером для описанной в [4] схемы. Среди сверхстройных деревьев с числом вершин 9 такого контрпримера нет. Среди сверхстройных деревьев с числом вершин 10 существует одно такое дерево: (5, 2, 2). Среди сверхстройных деревьев с числом вершин 11 существуют два таких дерева: (4, 3, 3) и (6, 2, 2). Можно предположить, что с ростом числа вершин количество таких контрпримеров будет возрастать, и такие графы можно выделить в некий подкласс сверхстройных деревьев.

Поскольку задача построения одного из THP для произвольного сверхстройного дерева не решена, можно решить эту задачу для некоторого подкласса сверхстройных деревьев. Таким подклассом, например, являются равнолучевые звезды [6].

Определение 6. Граф называется равнолучевой звездой с m лучами, каждый из которых состоит из n вершин, если $V = \{v_0, v_1^1, \ldots, v_n^1, \ldots, v_1^m, \ldots, v_n^m\}$, а $\alpha = \{v_i^j v_{i+1}^j : i = 1, \ldots, n-1; j = 1, \ldots, m\} \cup \{v_0 v_1^j : j = 1, \ldots, m\}$, где v_0 — центр равнолучевой звезды.

Теорема 1. Пусть граф S_n^m — равнолучевая звезда с m лучами, каждый из которых состоит из n вершин ($n \geqslant 2$). Тогда единственным THP для графа S_n^m является граф, полученный из тривиального расширения графа S_n^m удалением рёбер wv_{n-1}^j, $j = 1, \ldots, m$, где w — вершина, добавленная при построении тривиального расширения графа S_n^m.

Доказательство. Пусть граф $G = (V, \alpha)$ — равнолучевая звезда с m лучами, каждый из которых состоит из n вершин. Положим $H' = (G)$, $H' = (V', \alpha')$, где $V' = V \cup \{w\}$. Стоит отметить, что граф G можно рассматривать как объединение m цепей одинаковой длины n, имеющих одну общую начальную вершину v_0, т. е. эти цепи можно записать следующим образом: $v_0 v_1^j \ldots v_n^j$, $j = 1, \ldots, m$.

1. Будем удалять ребра из графа H'. Рассмотрим следующие случаи:

а) $H' - v_0w$.
Удалим вершину v_1^1 из полученного графа. Очевидно, что центром полученного графа будет w, так как v_0 имеет степень только $m-1$. Тогда из вершин $v_0, w, v_2^1, \ldots, v_n^1$ мы должны получить цепь из $n+1$ вершин с началом в вершине w, но так как вершина v_0 не смежна ни с w, ни с вершинами $v_2^1, \ldots, v_n^1$, то цепь из $n+1$ вершин получить невозможно. Граф G не вкладывается в полученный граф, а, следовательно, граф $H' - v_0w$ не является расширением графа G.

б) $H' - v_n^1w$.
Удалим вершину v_{n-1}^1 из полученного графа. Тогда вершина v_n^1 имеет степень 0. Следовательно, граф G не вкладывается в полученный граф и $H' - v_n^1w$ не является расширением для G. Аналогично доказывается, что графы $H' - v_n^jw$, $j = 2, \ldots, m$, также не являются расширениями для G.

в) $H' - v_i^1w$, $i = 1, \ldots, n-2$.
Удалим вершину v_{i+1}^1 из полученного графа. Так как v_i^1 будет иметь степень 1, то v_i^1 должна быть концом цепи. Началом цепи может быть вершина v_0 или w.
— v_0 — начальная вершина цепи. Тогда цепь максимальной длины имеет вид $v_0wv_1^1 \ldots v_i^1$. В эту цепь может входить максимум n вершин. Очевидно, что граф G не вкладывается в полученный граф.
— w — начальная вершина цепи. Тогда цепь максимальной длины имеет вид $wv_0v_1^1 \ldots v_i^1$. Очевидно, что граф G не вкладывается в полученный граф.
Таким образом, графы $H' - v_i^1w$, $i = 1, \ldots, n-2$, не являются расширениями для G. Аналогично доказывается, что графы $H' - v_i^jw$, $i = 1, \ldots, n-2$, $j = 2, \ldots, m$, также не являются расширениями для G.

2. Удалим из графа H' рёбра wv_{n-1}^j, $j = 1, \ldots, m$. Полученный граф назовём H. Покажем, что H является расширением для G.

а) $H - w$. Очевидно, что граф G вкладывается в граф $H - w$.

б) $H - v_0$. В этом случае центром графа станет вершина w и цепи можно записать следующим образом: $wv_n^j \ldots v_1^j$, $j = 1, \ldots, m$.

в) $H - v_i^j$, $i = 1, \ldots, n-1$, $j = 1, \ldots, m$. В этом случае «повреждённая» цепь имеет вид $v_0v_1^j \ldots v_{i-1}^jwv_{i+1}^j \ldots v_n^j$, $j = 1, \ldots, m$ (т. е. удалённую вершину v_i^j заменяет вершина w).

г) $H - v_n^j$, $j = 1, \ldots, m$. В этом случае «повреждённая» цепь имеет вид $v_0wv_1^j \ldots v_{n-1}^j$, $j = 1, \ldots, m$.

Таким образом, граф H является расширением графа G.

3. Докажем, что граф H является THP для G. Для этого докажем свойство неприводимости. Поскольку в п. 1 доказательства показано, что из тривиального расширения графа G невозможно удалить никакое ребро так, чтобы свойство расширения сохранялось (кроме рёбер wv_{n-1}^j, $j = 1, \ldots, m$), а граф H получен из тривиального расширения удалением рёбер wv_{n-1}^j, $j = 1, \ldots, m$, то и из графа H невозможно удалить никакое ребро так, чтобы свойство расширения сохранялось. Свойство неприводимости доказано.

4. Так как из тривиального расширения невозможно удалить никакие рёбра, кроме wv_{n-1}^j, $j = 1, \ldots, m$, так, чтобы свойство расширения сохранялось, то граф H — единственное THP для G. ■

ЛИТЕРАТУРА

1. *Богомолов А. М., Салий В. Н.* Алгебраические основы теории дискретных систем. М.: Наука, 2009.
2. *Абросимов М. Б.* Минимальные расширения объединения некоторых графов // Теоретические проблемы информатики и ее приложений. 2001. № 4. С. 3–11.
3. *Салий В. Н.* Доказательства с нулевым разглашением в задачах о расширениях графов // Вестник Томского государственного университета. Приложение. 2003. № 6. С. 63–65.
4. *Harary F. and Khurum M.* One node fault tolerance for caterpillars and starlike trees // Internet J. Comput. Math. 1995. V. 6. P. 135–143.
5. *Абросимов М. Б., Комаров Д. Д.* Об одном контрпримере для минимальных вершинных 1-расширений сверхстройных деревьев // Прикладная дискретная математика. Приложение. 2012. № 5. С. 83–84.
6. *Осипов Д. Ю.* О T-неприводимых расширениях сверхстройных деревьев // Прикладная дискретная математика. Приложение. 2013. № 6. С. 85–86.

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ДИСКРЕТНОЙ МАТЕМАТИКЕ

DOI 10.17223/20710410/25/10

УДК 510.52

ВЫЧИСЛИТЕЛЬНАЯ СЛОЖНОСТЬ ПОСТРОЕНИЯ КОМПОЗИЦИОННЫХ МОДЕЛЕЙ ЛИПШИЦ-ОГРАНИЧЕННЫХ ОТОБРАЖЕНИЙ

И. С. Калинников

Национальный исследовательский университет «МИЭТ», г. Москва, Россия

E-mail: gaminot@gmail.com

Работа посвящена вопросам численного построения композиционных моделей липшиц-ограниченных сюръективных функций одного аргумента. Композиционные модели являются частным случаем функциональной аппроксимации, получаемым путём композиции функций из заданного множества. Доказывается NP-трудность задачи построения оптимальной композиционной модели при заданном множестве функций, используемых для построения модели, и определённой приближаемой функции. Рассматриваются различные алгоритмы нахождения приближённых композиционных моделей, часть из которых имеет полиномиальную сложность; оцениваются возможности применения данных подходов.

Ключевые слова: *композиция функций, композиционные модели, NP-полнота, липшиц-ограниченность, вычислительная сложность.*

Введение

Композиционные модели используются в различных областях науки и техники, в том числе при представлении и анализе экспериментальных данных, для оптимизации процессов вычислений, анализа программного обеспечения методом «чёрного ящика», получения эквивалентных функциональных преобразований. Примеры применения композиционных моделей приводятся в работах [1 – 3]. Существенным недостатком композиционных моделей является отсутствие вычислительно эффективных алгоритмов их получения в общем случае. Известны эффективные алгоритмы для некоторых подклассов отображений, например для полиномов [4].

Точной композиционной моделью длины n для отображения f, построенной по системе функций $F = \{g_1, g_2, \ldots, g_m\}$, назовём композицию функций $g_{i_1}(\ldots g_{i_n}(x))$, такую, что $g_{i_1}(\ldots g_{i_n}(x)) = f$. Оптимальной композиционной моделью длины n будем называть композицию $g_{i_1}(\ldots g_{i_n}(x))$, такую, что по заданной метрике μ достигается $\min\left(\mu(g_{i_1}(\ldots g_{i_n}(x)), f)\right)$. Приближённой композиционной моделью будем называть результат процесса минимизации функции $H(i_1, \ldots, i_n) = \mu\left(g_{i_1}(\ldots g_{i_n}(x)), f\right)$ по набору индексов $(i_1, \ldots, i_n)$, $i_j \in \{1, \ldots, m\}$, $j = 1, \ldots, n$, с использованием методов, не гарантирующих глобальной оптимальности найденного решения в общем случае.

Основной целью работы является доказательство отсутствия эффективных алгоритмов (NP-трудности) построения оптимальных композиционных моделей в случае, если функции из $F \cup \{f\}$ являются липшиц-ограниченными сюръективными отображениями $[0, 1] \to [0, 1]$. Под липшиц-ограниченными отображениями понимаются такие

функции, что существует константа L, для которой верно

$$\forall x_1, x_2 \in [0,1]\ (|f(x_1) - f(x_2)| \leqslant L|x_1 - x_2|).$$

В работе NP-полнота понимается в смысле сводимости к ней за полиномиальное время остальных задач класса NP [5, 6].

Рассматриваются также вопросы построения алгоритмов поиска приближённой композиционной модели, приводятся имеющиеся методы решения данной задачи, кратко проводится их сравнение. В заключении подводятся итоги и перечисляются некоторые открытые вопросы.

1. NP-трудность построения оптимальной и точной композиционной модели

В настоящее время доказана NP-трудность задачи MGS — поиска минимальной генерирующей последовательности перестановок для заданной перестановки. NP-трудность MGS доказывается редукцией к ней проблемы 3XC, входящей в список 21 проблемы Карпа [5, 6]. В приводимом далее доказательстве NP-трудности задачи построения композиционной модели используются некоторые идеи работы [7], где проводится доказательство NP-трудности задачи MGS.

Дадим формулировки задач, которые фигурируют в доказательстве:

— 3XC(S, U) (точное покрытие 3-множествами). Дано множество $S = \{u_1, \ldots, u_{3n}\}$ и подмножество $U \subset S^3$, $|U| = m$. Определить, существует ли подмножество $S' \subset U$, $|S'| = n$, такое, что для любого $u \in S$ есть ровно одно 3-множество $s' \in S'$, содержащее u.

П р и м е ч а н и е: элементы из U — неупорядоченные множества из трёх различных элементов S, называемые 3-множествами.

— COMP-L(F, f, μ, ε) (распознавание композиции липшиц-ограниченных функций, приближающей целевую функцию с заданной погрешностью). Дан набор функций $F = \{g_1, \ldots, g_m\}$ и целевая функция f, причём функции f, g_i являются сюръективными липшиц-ограниченными отображениями $[0,1] \to [0,1]$. Определить, существует ли набор $(i_1, \ldots, i_n)$, такой, что $\mu\left(g_{i_1}\left(\ldots g_{i_n}\right), f\right) \leqslant \varepsilon$.

В качестве метрики μ может рассматриваться любая метрика, для которой вычисление (или приближение с устанавливаемой погрешностью) значения между определяемыми в доказательстве функциями может быть произведено за полиномиальное время.

Теорема 1. 3XC(S, U) редуцируется к COMP-L(F, f, μ, ε).

Доказательство. Каждый элемент $s \in U$ можно задать как бинарный вектор b_s длины $3n$, где $b_s(i) = 1$, если $u_i \in s$, а все прочие позиции в b_s заняты нулями. Очевидно, получение таких векторов может быть выполнено за время, полиномиальное от n и m.

Каждому бинарному вектору b_s поставим в соответствие кусочно-определённую липшиц-ограниченную функцию из $C^1[0,1]$. В качестве базовой функции выберем $t(x) = \dfrac{9}{2}\operatorname{sign}(x)\left((1-|x|)^3 - (1-|x|)^2\right)$, $x \in [-1,1]$, введём также функцию окна

$$w(i, N, x) = \begin{cases} 1, x \in \left[\dfrac{i-1}{N}, \dfrac{i}{N}\right] \\ 0, x \notin \left[\dfrac{i-1}{N}, \dfrac{i}{N}\right] \end{cases},$$

и функции

$$g_s(x) = x + \frac{1}{3n}\sum_{i=1}^{3n} b_s(i)w(i,3n,x)t(6nx-2i+1).$$

Систему функций F для задачи COMP-L определим так: $F = \{g_s : s \in U\}$, $|F| = m$; а целевую функцию f как

$$f(x) = x + \frac{1}{3n}\sum_{i=1}^{3n} w(i,3n,x)t(6nx-2i+1).$$

Построение и вычисление функций из F и функции f выполняется за время, линейное от n и m. Константа Липшица всех функций из F и функции f равна 4.

Построенные функции g_{s1}, g_{s2} при выполнении композиции ведут себя следующим образом:

1) Если s_1 и s_2 не содержат одинаковых элементов u_i, то $g_{s1}(g_{s2}) = g_{s2}(g_{s1})$ соответствуют функции g, построенной для бинарного вектора $b = b_{s1} \vee b_{s2}$, так как

$$\begin{aligned} g_{s1}\left(g_{s2}\left(x\right)\right) &= x + \frac{1}{3n}\sum_{i=1}^{3n} b_{s1}\left(i\right) w\left(i,3n,x\right) t\left(6nx-2i+1\right) + \\ &+ \frac{1}{3n}\sum_{i=1}^{3n} b_{s2}\left(i\right) w\left(i,3n,x\right) t\left(6nx-2i+1\right) = \\ &= x + \frac{1}{3n}\sum_{i=1}^{3n} \left(b_{s1}\left(i\right) \vee b_{s2}\left(i\right)\right) w\left(i,3n,x\right) t\left(6nx-2i+1\right). \end{aligned}$$

2) Если s_1 и s_2 содержат одинаковые элементы, то функция в соответствующих частях изменяется следующим образом: если $b_{s1}(i) = b_{s2}(i) = 1$ для некоторого i, то для $x \in \left[\frac{i-1}{3n}, \frac{i}{3n}\right]$ выполняется равенство

$$g_{s1}(g_{s2}(x)) = x + \frac{1}{3n}t\left(6nx-2i+1\right) + \frac{1}{3n}t\left(6n\left(x+\frac{1}{3n}t\left(6nx-2i+1\right)\right)-2i+1\right),$$

что отличается от значений g_{s1} и g_{s2}. При дальнейших композициях $g_{s1}(g_{s2}(x))$ с функциями $g_s(x)$, такими, что $b_s(i) = 1$, на отрезке $\left[\frac{i-1}{3n}, \frac{i}{3n}\right]$ будут образовываться многократные подстановки $t(x)$ в саму себя, не совпадающие с исходной функцией.

Подобные свойства позволяют определить отсутствие, однократное или многократное покрытие элемента u_i в результате объединения 3-множеств $s \in S$, представляемых функциями из F. Приведём возможные результаты композиции функций $g_{s1}(x)$ и $g_{s2}(x)$ в таблице:

$b_{s1}(i)$ — бит покрытия элемента 3-множеством s_1	$b_{s2}(i)$ — бит покрытия элемента 3-множеством s_2	Функция результата композиции на отрезке $\left[\frac{i-1}{3n}, \frac{i}{3n}\right]$
0	0	x
1	0	$x + \frac{t(6nx-2i+1)}{3n}$
0	1	$x + \frac{t(6nx-2i+1)}{3n}$
1	1	$x + \frac{1}{3n}t\left(6nx-2i+1\right) + \frac{1}{3n}t\left(6n\left(x+\frac{1}{3n}t(6nx-2i+1)\right)-2i+1\right)$

Свойства построенной системы функций показывает рис. 1, на котором сверху вниз изображены функция g_{s1} для $s_1 = \{u_1, u_2, u_4\}$; функция g_{s2} для $s_2 = \{u_2, u_3, u_5\}$; композиция функций $g_{s1}(g_{s2})$ при $m = 2$.

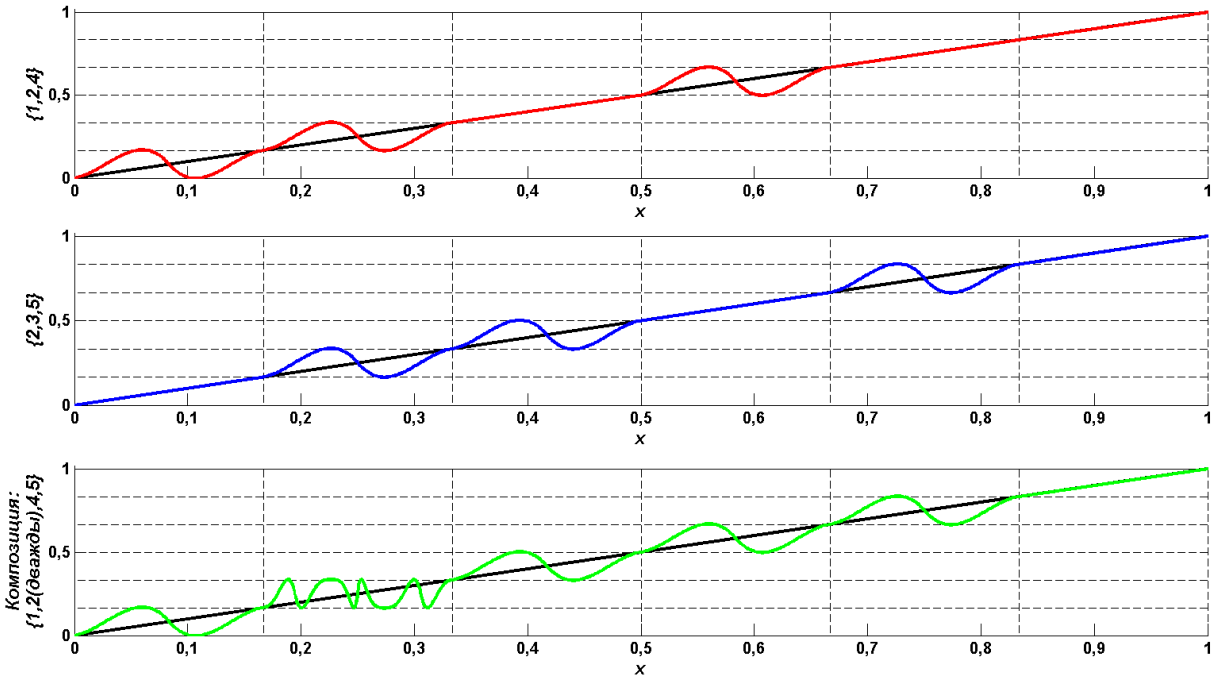


Рис. 1. Функции, представляющие 3-множества, и их композиция

Пусть определено существование решения задачи COMP-L(F, f, μ, ε) при $\varepsilon = 0$, тогда если оно существует, то покрытие для задачи 3XC также существует, и найденный набор индексов $(i_1, \ldots, i_m)$ соответствует S' — решению задачи 3XC. Отсутствие решения задачи COMP-L ведёт к отсутствию решения задачи 3XC. Для получения покрытия из набора индексов используется обратное преобразование функций с данными индексами в бинарные векторы, а их — в 3-множества. Сложность данного преобразования, очевидно, является полиномиальной от n и m. ■

Замечание 1. Может показаться, что задача может быть NP-полной в силу необходимости вычисления значений метрики между липшиц-ограниченными функциями, но в данном случае метрика Чебышева для приведенных наборов функций может быть вычислена за полиномиальное от m и n время на основе таблицы расстояний между подстановками функции $t(x)$. Аналогично может рассматриваться метрика Чебышева по любой дискретной системе узлов, позволяющей различать композиции функции $t(x)$.

Замечание 2. Задача построения точной композиционной модели является частным случаем задачи COMP-L(F, f, μ, ε) при $\varepsilon = 0$. Значит, в доказательстве можно совместно рассматривать задачу распознавания композиционной модели, приближающей целевую функцию с заданной погрешностью, и задачу построения точной композиционной модели.

Замечание 3. Задача построения оптимальной композиционной модели NP-трудна, так как задача распознавания композиционной модели COMP-L(F, f, μ, ε) NP-полна. Предположим, решена задача построения оптимальной композиционной модели и известен $\min\left(\mu\left(g_{i_1}\left(\ldots g_{i_n}\right), f\right)\right)$; тогда задача распознавания решается путём

сравнения ε с найденным минимумом. Следовательно, задача построения оптимальной композиционной модели не может решаться проще задачи распознавания композиционной модели.

Таким образом, задача построения точной композиционной модели липшиц-ограниченной функции является NP-полной, а задача построения оптимальной композиционной модели — NP-трудной. Названные задачи скорее всего не имеют эффективных алгоритмов решения (проблема $\mathrm{P} \neq \mathrm{NP}$). Однако данный факт не влияет на возможность выполнять поиск приближённой композиционной модели.

2. Методы поиска приближённой композиционной модели

Рассмотрим различные методы поиска приближённых композиционных моделей и их свойства. Методы поиска приближённых композиционных моделей по основному принципу работы можно разделить на использующие:

— аппроксимации специальными видами функций;
— параметрическую оптимизацию;
— теорию поиска в метрических пространствах;
— случайный или генетический поиск.

При аппроксимации специальными функциями необходимо подобрать класс отображений C так, что:

1) возможно аппроксимировать f и $g_i \in F$ функциями данного класса;
2) в выбранном классе существуют эффективные алгоритмы построения композиционных моделей.

Если такой класс C отображений находится, то функция f заменяется на $\hat{f} \in C$ (входящую в выбранный класс), а функции $g_i \in F$ заменяются на $\hat{g}_i \in \hat{F} \subset C$. Далее решается задача построения точной/оптимальной композиционной модели в классе C. Пусть решение в классе C найдено, погрешность равна $\varepsilon = \mu\left(\hat{f}, \hat{g}_{i_1}\left(\ldots \hat{g}_{i_n}\right)\right)$. Рассмотрим погрешность решения $(i_1, \ldots, i_n)$ для исходной задачи: $\varepsilon \leqslant \mu\left(f, \hat{f}\right) + \mu\left(\hat{f}, \hat{g}_{i_1}\left(\ldots \hat{g}_{i_n}\right)\right) + \mu\left(\hat{g}_{i_1}\left(\ldots \hat{g}_{i_n}\right), g_{i_1}\left(\ldots\ g_{i_n}\right)\right) = \mu\left(f, \hat{f}\right) + \hat{\varepsilon} + \mu\left(\hat{g}_{i_1}\left(\ldots \hat{g}_{i_n}\right), g_{i_1}\left(\ldots\ g_{i_n}\right)\right)$. Таким образом, погрешность ограничена сверху функцией от расстояний между исходными функциями и их функциями-представителями из класса C, а также погрешностью решения задачи в классе C. Более точный анализ погрешностей можно провести для конкретной метрики, например метрики Чебышева:

$$\varepsilon = \mu\left(f, g_{i_1}\left(\ldots g_{i_n}\right)\right) \leqslant \mu\left(f, \hat{f}\right) + \hat{\varepsilon} + \sum_{j=1}^{n}\left(\prod_{k=1}^{j-1}\left[\min\left(L_{g_{i_k}}, L_{\hat{g}_{i_k}}\right)\right]\mu\left(g_{i_j}, \hat{g}_{i_j}\right)\right),$$

при этом, если обозначить $\varepsilon_c = \max\left(\max_j\left(\mu\left(\hat{g}_j, g_j\right)\right), \mu(\hat{f}, f)\right)$, а $L = \max_i\left(L_{g_i}\right)$, то получаем оценку $\hat{\varepsilon} \leqslant \varepsilon_c\left(\dfrac{L^n - 1}{L - 1} + 1\right) + \varepsilon$.

Так как $L > 1$ (по условию на отображения f, $g_i \in F$), то верхняя оценка погрешности возрастает очень быстро с ростом длины композиционной модели n, при этом на практике L редко бывает меньше 3. Таким образом, аппроксимационный подход может применяться, например, для классов C полиномиальных [4] и рациональных [8] функций, для которых существуют полиномиальные алгоритмы построения композиционных моделей. При этом аппроксимационный подход будет эффективен, когда

функции f и F с малой погрешностью ε_c приближаются функциями из C, а длина n композиционной модели мала.

Методы поиска приближённой композиционной модели, основанные на параметрическом подходе, традиционно включают две стадии:

— параметрическую оптимизацию с использованием традиционных (классических) методов оптимизации;
— прямой перебор вариантов с целью определения подходящего.

Данные стадии могут быть сгруппированы различным образом. Например, в работе [3] автор предлагает, рассматривая бесконечное множество F(с семействами функций, зависящими от параметров), изначально, на основе приближённых значений параметров, выбрать перебором подходящую композиционную модель. При выполнении перебора поиск выполняется с точностью до определения семейства каждой входящей в композицию функции. После этого автор [3] предлагает применить параметрическую оптимизацию для определения конкретных функций в каждом семействе. Полагается, что $F = F_1 \cup F_2 \cup \ldots \cup F_m$; на первом этапе перебором ищется

$$\min_{\{(i_1,\ldots,i_n):i_1,\ldots,i_n\in\{1,\ldots,m\}\}} \left(\mu\left(f, g_1 = \text{select}\left(f, F_{i_1}\right) \in F_{i_1}\left(\ldots \right.\right.\right.$$
$$\left.\left.\left. \ldots g_n = \text{select}\left(g_{n-1}^{-1}\left(\ldots g_1^{-1}\left(f\right), F_{i_n}\right)\right) \in F_{i_n}\right)\right)\right),$$

где $\text{select}\left(f, F_i\right)$ выбирает некоторым вычислительно простым алгоритмом параметры функции из F_i для того, чтобы она приближала f (например, на основе метода МНК). На втором этапе ищется минимум функции

$$H\left(\bar{p}_1, \ldots, \bar{p}_n\right) = \mu\left(f, g_1 = F_{i_1}\left(\bar{p}_1\right)\left(\ldots g_n = F_{i_n}\left(\bar{p}_n\right)\right)\right),$$

где $(i_1, \ldots, i_n)$ — индексы, выбранные в результате поиска на первом этапе. Альтернативой является изначальное сведение задачи к задаче параметрической оптимизации, а затем поиск лучшего целочисленного округления полученного решения, подобный подход рассмотрен в работе [9]. Недостатками этих подходов является сложность предсказания качества получаемого решения и быстрое возрастание числа параметров, по которым проводится оптимизация, с ростом длины композиционной модели. Последний недостаток не позволяет использовать многие методы оптимизации, поэтому в большинстве случаев производится поиск только локального экстремума.

Поскольку на множестве $\{f\} \cup \{g_{i_1}(\ldots g_{i_n}) : i_1, \ldots, i_n \in \{1, \ldots, m\}\} = \{f\} \cup F^n$ в задаче построения композиционной модели определена метрика, то могут использоваться методы теории поиска в метрических пространствах, рассматриваемые в работах [10, 11]. Оптимизация поиска в метрических пространствах достигается за счёт применения неравенства треугольника, на основе предвычислений расстояний во множестве F^n, дальнейшая оптимизация с ограничением точности получаемого результата может быть достигнута с использованием методов либо раннего прекращения поиска, либо методом ослабления условий сравнений в ходе поиска. Проблема данного подхода заключается в необходимости хранения предвычислений размера $\mathrm{O}(|F|^n)$ для оптимизации процесса поиска. Это объясняется тем, что оценка расстояния $\mu\left(f, g\right)$ за счёт неравенства треугольника может быть произведена, только если существует t, такое, что $\mu\left(f, t\right)$ и $\mu\left(t, g\right)$ могут быть оценены. Таким образом, отношение «оцениваемости» расстояния $U_{\text{est}} \subseteq \left[\{f\} \cup F^n\right]^2$ является транзитивным замыканием отношения

«определённости» $U_{\text{def}} \subseteq [\{f\} \cup F^n]^2$ расстояния на множестве $\{f\} \cup F^n$ при условии, что оценки строятся за счёт неравенства треугольника с использованием «определённых» значений расстояния. Следовательно, $U_{\text{est}} \subseteq [\{f\} \cup F^n]^2$, только если отношение «определённости» расстояния U_{def} включает дерево на вершинах $\{f\} \cup F^n$, что требует минимум $\mathrm{O}(|F|^n)$ вычисленных расстояний. С учётом необходимости хранения предвычисленных значений расстояний невозможно организовать поиск для больших значений n и множеств F со значительным количеством функций. Также следует иметь в виду, что если f заранее не известна, а предвычисления проводятся исключительно над множеством функций F^n, что чаще всего и необходимо в практических задачах, то, пока не будет найдено $(i_1, \ldots, i_n) : \mu\left(f, g_{i_1}\left(\ldots g_{i_n}\right)\right) \ll 1$, предвычисленные значения не могут быть использованы для эффективной оценки расстояний $\mu\left(f, g_{j_1}\left(\ldots g_{j_n}\right)\right)$, используемых в работе алгоритма. Таким образом, на первых этапах работы алгоритма эффективность предвычислений незначительна.

Методы, основанные на случайном или генетическом поиске, применяются к задачам, аналогичным задаче поиска приближённой композиционной модели, в работах [1, 2]. Поскольку методология самих этих процедур хорошо известна [1, 12], обратим внимание лишь на проблемы их применения к задаче поиска приближённой композиционной модели. Случайный поисковый алгоритм для повышения эффективности работы требует определения перечня параметров распределения вектора модификации текущего решения, которые должны адаптироваться, исходя из истории поиска. Пока объём параметров адаптации незначителен, их выбор, как и определение правил адаптации, составляет значительную проблему, а возрастание эффективности незначительно. Если параметров адаптации много, то адаптация происходит крайне медленно и возрастание эффективности также незначительно. В случае применения генетического алгоритма проблема его использования заключается в определении функции смешивания элементов, такой, что коэффициент корреляции средней метрики элемента, получающегося на выходе функции смешивания, со средней метрикой элементов, попадающих на вход функции смешивания, не равен нулю. Это условие следует из теоремы Прайса [13] об эффективности генетического алгоритма, где под метрикой элемента понимается значение метрики между данным вариантом композиционной модели (названным здесь элементом) и целевой функцией f. На текущий момент для задачи построения приближённой композиционной модели такая процедура смешивания не известна.

Заключение

Задача построения точной композиционной модели липшиц-ограниченной функции является NP-полной, а задача построения оптимальной композиционной модели — NP-трудной. Автору не известны неперебо́рные алгоритмы решения данных задач, например алгоритмы с константами $c, d < 1$ и сложностью $\mathrm{O}\left((c|F|)^{dn}\right)$, хотя их существование вполне возможно.

Алгоритмы поиска приближённой композиционной модели могут иметь полиномиальную сложность (например, алгоритмы, основанные на аппроксимации специальными видами функций и параметрической оптимизации), но применимость этих алгоритмов ограничена. Более универсальные алгоритмы, основанные на теории поиска в метрических пространствах, генетическом и случайном поисках, не гарантируют объёма вычислений (или предвычислений) менее $\mathrm{O}(|F|^n)$, поэтому актуальным является построение и математическое обоснование алгоритмов, позволяющих снизить объём вычислений. Отдельную область исследований представляют алгоритмы, использующие

функции распределения расстояний между элементами $\{f\} \cup F^n$ и применяющие эту информацию для оптимизации поиска. Для таких алгоритмов могут быть вычислены оценки на достигаемое повышение эффективности и вероятностные характеристики точности решения, исходя из способов оценки распределения расстояний.

ЛИТЕРАТУРА

1. *Koza J. R.* Genetic Programming: on the Programming of Computers by Means of Natural Selection. London: A Bradford Book, 1998. 815 p.
2. *Luke S.* Essentials of metaheuristics. `http://cs.gmu.edu/~sean/book/metaheuristics.` 2009.
3. *Лабутин С. А., Пугин М. В.* Анализ сигналов и зависимостей: учеб. пособие. Н. Новгород: Нижегород. гос. тех. ун-т, 2001. 158 с.
4. *Seong J. K., Elber G., and Kim M. S.* Polynomial decomposition and its applications. `http://cana.kaist.ac.kr/seong/decomposition.pdf.` 2003.
5. *Karp R. M.* Reducibility among combinatorial problems // GJ-474 report. 1971. P. 87–103. `http://www.seas.upenn.edu/~bhusnur4/cit596_spring2014/karp-1971.pdf`
6. *Пападимитриу Х. Х., Стайглиц К.* Комбинаторная оптимизация: Алгоритмы и сложность. М.: Мир, 1987. 520 с.
7. *Even S. and Goldreich O.* The minimum-length generator sequence problem is NP-hard // J. Algorithms. 1981. No. 2. P. 311–313.
8. *Alonso C., Gutierrez J., and Recio T.* A rational function decomposition algorithm by near-separated polynomials // J. Symbolic Comput. 1995. V. 19. P. 527–544.
9. *Калинников И. С.* Алгоритм построения декомпозиции непрерывной функции одного аргумента по заданному множеству функций // Инновации в науке, образовании и бизнесе: X Междунар. научн. конф. Калининград: КГТУ, 2012. Ч. 2. С. 160–163.
10. *Chavez E., Navarro G., Baeza-Yates R., and Marroquin J. L.* Search in metric spaces // ACM Computing Surveys. 2001. V. 33. No. 3. P. 273–321.
11. *Zezula P., Amato G., Dohnal V., and Batko M.* Similarity Search: The Metric Space Approach. N. Y.: Springer Verlag, 2006. 220 p.
12. *Растригин Л. А., Рипа К. К., Тарасенко Г. С.* Адаптация случайного поиска. Рига: Зинатне, 1978. 243 с.
13. *Alenberg L.* The schema theorem and Price's theorem // Foundations of Genetic Algorithms 3. San Francisco: Morgan Kaufmann, 1995. P. 23–49.

2014 Вычислительные методы в дискретной математике № 3(25)

DOI 10.17223/20710410/25/11

УДК 510.25+510.52+519.688

ВЫЧИСЛЕНИЕ ВЕЩЕСТВЕННОЙ W-ФУНКЦИИ ЛАМБЕРТА W_0 В ПРЕДЕЛАХ FP//LINSPACE

М. А. Старицын, С. В. Яхонтов

Санкт-Петербургский государственный университет, г. Санкт-Петербург, Россия

E-mail: m.staritzyn2012@yandex.ru, SergeyV.Yakhontov@gmail.com

Строится FP//LINSPACE алгоритмический аналог вещественной W-функции Ламберта $W_0(x)$ на отрезке $[-(re)^{-1}, (re)^{-1}]$ FP//LINSPACE алгоритмических вещественных чисел, где r — рациональное, $r > 4/3$ (в качестве r можно брать любое рациональное с таким условием). Для построения алгоритмического аналога вещественной W-функции Ламберта $W_0(x)$ предлагается алгоритм WLE расчёта двоично-рациональных приближений данной функции на отрезке $[-(re)^{-1}, (re)^{-1}]$ с полиномиальной временной и линейной емкостной сложностью на машине Тьюринга. Алгоритм WLE строится на основе разложения в ряд Тейлора данной функции, при этом показывается и используется в алгоритме линейная сходимость ряда Тейлора W-функции Ламберта $W_0(x)$ на отрезке $[-(re)^{-1}, (re)^{-1}]$.

Ключевые слова: *вещественная W-функция Ламберта W_0, алгоритмические вещественные функции, машина Тьюринга, полиномиальная временная сложность, линейная емкостная сложность.*

Введение

В работе предлагается алгоритм расчёта вещественной W-функции Ламберта W_0 [1] на отрезке $[-(re)^{-1}, (re)^{-1}]$, где r — рациональное, $r > 4/3$ (говоря более точно, основной ветви W_0 вещественной W-функции Ламберта W) с полиномиальной временной и линейной емкостной сложностью на машине Тьюринга. Алгоритм строится на основе разложения в ряд Тейлора данной функции с использованием алгоритма вычисления линейно сходящихся степенных рядов в пределах FP//LINSPACE из [2] в качестве базового алгоритма.

При построении алгоритмического аналога вещественной W-функции Ламберта W_0 берётся модель алгоритмических чисел и функций, изложенная в [3]. Посредством FP//LINSPACE будем обозначать класс алгоритмов, полиномиальных по времени и линейных по памяти при вычислении на машине Тьюринга [4].

W-функция Ламберта W является трансцендентной функцией и относится к классу специальных функций (т. е. как специальная функция W-функция Ламберта W не выражается через элементарные функции). W-функция Ламберта W интересна как с практической точки зрения, так как используется, например, в математических задачах физики, так и с теоретической точки зрения, например при рассмотрении вычислительной сложности констант и функций математического анализа в теоретической информатике.

Имеется достаточно большое количество алгоритмов расчёта констант и функций математического анализа: алгоритмы на основе AGM [5], метод Карацубы быстрого вычисления экспоненциальной функции [6], метод `binary splitting` [7] как вариант метода Карацубы, алгоритм `bit-burst` расчёта голономных функций (D-finite на англ.) [8], метод Ньютона вычисления обратных функций и др. Но на данный мо-

мент не известно никаких результатов относительно использования этих методов вычисления констант и функций математического анализа для вычисления вещественной W-функции Ламберта W_0 на отрезке $[-(re)^{-1}, (re)^{-1}]$ с линейной памятью на машине Тьюринга. Кроме того, некоторые из перечисленных методов не могут быть применены для расчёта W-функции Ламберта W_0; например, алгоритм `bit-burst` расчёта голономных функций [8] неприменим к расчету W-функции Ламберта W_0, так как данная функция не принадлежит к классу голономных функций. Поэтому результат, изложенный в данной работе, является новым в области вычислительной сложности алгоритмических чисел и функций.

Частично результат работы представлялся на конференции СПИСОК-14 [9].

1. Алгоритмические вещественные числа и функции

В данной работе основу представления конструктивных объектов (чисел и функций) составляет понятие алгоритмической последовательности φ, сходящейся по Коши [3], при этом в качестве вычислительной модели берётся машина Тьюринга. Такая последовательность определяется на множестве всех натуральных чисел $\mathbb{N}$, включая 0, а областью аппроксимирующих значений является всюду плотное в $\mathbb{R}$ естественное подмножество множества рациональных чисел. Для последовательности, сходящейся по Коши и задающей вещественное число x, требуют, чтобы выполнялось

$$|\varphi(n) - x| \leqslant 2^{-n}$$

для любого натурального n.

В качестве множества аппроксимирующих значений берётся множество двоично-рациональных чисел $\mathbb{D}$ [3]. Рациональное число d называется двоично-рациональным, если $d = m/2^n$ для некоторого целого m и натурального n. Двоично-рациональные числа имеют конечное двоичное представление: строка s, равная

$$\pm u_p u_{p-1} \ldots u_0 . v_1 v_2 \ldots v_r,$$

обозначает число

$$d = \pm \left(\sum_{i=0}^{p} u_i 2^i + \sum_{j=1}^{r} v_j 2^{-j} \right).$$

Длина представления двоично-рационального числа определяется как количество символов в строке s, равное, с учётом знака и двоичной точки, $p+r+3$, и обозначается $l(s)$. Под точностью представления $\mathrm{prec}(s)$ понимается число битов справа от двоичной точки, то есть r. С точки зрения изучения вычислительной сложности двоично-рациональные числа удобны тем, что для любого n двоично-рациональные числа с точностью n равномерно распределены на вещественной прямой [3].

Последовательность $\varphi : \mathbb{N} \to \mathbb{D}$ двоично-рационально сходится к вещественному числу x, если для любого $n \in \mathbb{N}$ выполняется $\mathrm{prec}(\varphi(n)) = n + 1$ и

$$|\varphi(n) - x| \leqslant 2^{-n}.$$

Множество всех функций, двоично-рационально сходящихся к вещественному числу x, обозначается CF_x.

Вещественное число x называется CF-алгоритмическим [3], если CF_x содержит вычислимую функцию φ.

Вещественная функция f, заданная на отрезке $[a, b]$, называется алгоритмической функцией [3] на этом отрезке, если существует машина Тьюринга M с оракульной функцией, такая, что для любого $x \in [a, b]$ и любой вычислимой функции $\varphi \in CF_x$ функция ψ, вычисляемая M с оракульной функцией φ, принадлежит $CF_{f(x)}$.

Фактически это означает, что для любой вычислимой функции $\varphi \in CF_x$ и любого $n \in \mathbb{N}$ машина M последовательно вычисляет $m \in \mathbb{N}$ и $d \in \mathbb{D}$, такие, что

$$|\varphi(m) - x| \leqslant 2^{-m}, \quad |d - f(x)| \leqslant 2^{-n}.$$

Сложность расчёта двоично-рациональных приближений алгоритмических чисел и функций определяется в [3] на основе длины двоичного представления точности вычисления. Память ленты запроса и ленты ответа оракульной функции при оценке емкостной вычислительной сложности алгоритма не учитывается. Обращение к оракульной функции $\varphi \in CF_x$ аргумента x алгоритмической функции осуществляется следующим образом:

— на ленту запроса оракульной функции записывается точность вычисления аргумента 2^{-m} в виде 0^m (унарная запись);
— рассчитывается значение $\varphi(m)$ оракульной функции, и результат записывается на ленту ответа;
— значение $\varphi(m)$ считывается с ленты ответа в промежуточную память.

Определение 1 [2]. Число $x \in \mathbb{R}$ назовём FP//LINSPACE алгоритмическим вещественным числом, если существует функция $\varphi \in CF_x$, вычислимая в пределах FP//LINSPACE.

Определение 2 [2]. Вещественную функцию f, заданную на отрезке $[a, b]$, назовём FP//LINSPACE алгоритмической вещественной функцией на отрезке $[a, b]$, если для любого $x \in [a, b]$ функция ψ (указанная в определении алгоритмической функции) из $CF_{f(x)}$ является FP//LINSPACE вычислимой.

Множества FP//LINSPACE алгоритмических вещественных чисел и функций будем обозначать FP//LINSPACE_{CF} и $\text{FP//LINSPACE}_{C[a,b]}$ соответственно. Здесь использование индекса $C[a, b]$ обусловлено тем, что алгоритмические функции являются непрерывными на всей области определения [3]. Построение алгоритмического аналога вещественной функции f на отрезке $[a, b]$ означает описание алгоритма, вычисляющего двоично-рациональные приближения с произвольной точностью значений $f(x)$ для $x \in [a, b]$.

2. Вычисление W-функции Ламберта W_0

Вещественная W-функция Ламберта W определяется как решение функционального уравнения

$$x = W(x)e^{W(x)}.$$

Данное решение является функцией, обратной к функции $f(x) = xe^x$. Вещественная W-функция Ламберта W определена на полуинтервале $[-e^{-1}, \infty)$ и имеет две ветви, верхнюю W_0 и нижнюю W_{-1}.

Будем строить алгоритм расчёта верхней ветви W_0, который обозначим через WLE, на основе разложения в ряд Тейлора данной функции с использованием алгоритма вычисления линейно сходящихся степенных рядов в пределах FP//LINSPACE из [2] в качестве базового алгоритма.

Напомним, что степенной ряд $S = \sum\limits_{i=0}^{\infty} a_i$ называется линейно сходящимся степенным рядом, если его частичная сумма $S_{\mu(k)} = \sum\limits_{i=0}^{\mu(k)} a_i$, такая, что $\mu(k)$ — линейная функция от k, отличается от точного значения не более чем на 2^{-k}: $|S - S_{\mu(k)}| \leqslant 2^{-k}$.

Рассмотрим ряд Тейлора функции W_0 [1] в окрестности точки $x = 0$:

$$W_0(x) = \sum_{k=1}^{\infty} a_k x^k = \sum_{k=1}^{\infty} \frac{(-k)^{(k-1)}}{k!} x^k, \tag{1}$$

радиусом сходимости данного ряда является величина e^{-1}. Перепишем ряд (1) в виде

$$W_0^{(1)}(x) = \sum_{k=1}^{\infty} a_k^{(1)} x^k = \sum_{k=1}^{\infty} \frac{(-k)^{(k-1)}}{e^k \cdot k!} (ex)^k, \tag{2}$$

рассмотрим данный ряд для $x \in [-(re)^{-1}, (re)^{-1}]$ и оценим сверху модуль n-го остатка

$$R_n^{(1)}(x) = \sum_{k=n+1}^{\infty} a_k^{(1)} x^k \tag{3}$$

данного ряда. В силу неравенства $k! \geqslant 2^2 \cdot k^{k+1/2} e^{-k}$, следующего из формулы Стирлинга [10], получаем

$$|a_k^{(1)} x^k| < \frac{k^k}{e^k \cdot 2^2 \cdot k^{k+1/2} e^{-k}} \cdot \frac{e^k}{(re)^k} < 2^{-C_1 \cdot k}$$

для $x \in [-(re)^{-1}, (re)^{-1}]$. Здесь C_1 — константа, зависящая от рационального r. Следовательно,

$$R_n^{(1)}(x) < \sum_{k=n+1}^{\infty} 2^{-C_1 \cdot k} = C_2 \cdot 2^{-C_1 (n+1)},$$

то есть ряд (2) линейно сходится.

Далее, покажем FP//LINSPACE вычислимость коэффициентов $a_k^{(1)}$ ряда (2) (отметим, что входными данными для алгоритма вычисления коэффициентов $a_k^{(1)}$ является двоичная запись точности вычисления 2^{-m}, что является точностью вычисления аргумента x [2]).

Для этого запишем коэффициенты $a_k^{(1)}$ в виде произведения $a_k^{(1)} = \dfrac{1}{ek} \prod\limits_{j=1}^{k-1} (-1)^{j-1} b_j$, где $b_j = \dfrac{k}{ej}$; обозначим $a_{(p,k)} = \prod\limits_{j=1}^{p} (-1)^{j-1} b_j$. Будем вычислять величину $a_{(k-1,k)}$ (равную $a_k^{(1)}$) последовательно в цикле для $p \in \{1, \ldots, k-2\}$, на каждом шаге выполняя произведение $a^*_{(p,k)} b^*_{p+1}$ (при этом $a_{(1,k)} = b_1$), где $a^*_{(p,k)}$ — приближённое значение величины $a_{(p,k)}$ с некоторой точностью ε_p; b^*_{p+1} — приближённое значение величины b_{p+1} с той же точностью $\varepsilon_p = 2^{-q} < 2^{-1}$; q — некоторое натуральное. Произведение $\zeta = a^*_{(p,k)} b^*_{p+1}$ будем округлять с точностью ε_p, то есть отбрасывать биты числа ζ после двоичной точки, начиная с q-го бита.

Используя метод математической индукции по $p \in \{1, \ldots, k-2\}$, покажем, что если взять $\varepsilon_1 \leqslant 2^{-C_3 k + (\log_2(k) + 2)}$, где C_3 — некоторая константа, то

$$\varepsilon_p \leqslant 2^{-C_3 k + \sum\limits_{j=1}^{p} (\log_2(k) - \log_2(j) + 2)}$$

для любого $p \in \{1, \ldots, k-2\}$. База индукции: $p = 1$; в этом случае величина $a^*_{(1,k)}$, равная b^*_1, вычисляется с точностью ε_1. Индукционный переход: пусть для $p \in \{1, \ldots, k-3\}$ выполняется $|a^*_{(p,k)} - a_{(p,k)}| \leqslant \varepsilon_p$. Тогда

$$|a^*_{(p+1,k)} - a_{(p+1,k)}| \leqslant |a^*_{(p,k)} b^*_{p+1} - a_{(p,k)} b_{p+1}| + \varepsilon_p =$$
$$= |a^*_{(p,k)} b^*_{p+1} - a^*_{(p,k)} b_{p+1} + a^*_{(p,k)} b_{p+1} - a_{(p,k)} b_{p+1}| + \varepsilon_p \leqslant$$
$$\leqslant |a^*_{(p,k)}(b^*_{p+1} - b_{p+1})| + |b_{p+1}(a^*_{(p,k)} - a_{(p,k)})| + \varepsilon_p < 2\varepsilon_p + \frac{k}{p+1}\varepsilon_p < 2^{\log_2(k) - \log_2(p+1) + 2}\varepsilon_p$$

(здесь используется оценка $a_{(p,k)} < 2^{-1}$), то есть $\varepsilon_{p+1} < 2^{-C_3 k + \sum\limits_{j=1}^{p+1} (\log_2(k) - \log_2(j) + 2)}$.

Теперь оценим сверху величину

$$\nu = \sum_{j=1}^{p} (\log_2(k) - \log_2(j)).$$

В силу неравенства $p! \geqslant 2p^{p+1/2}e^{-p}$, следующего из формулы Стирлинга, получаем

$$\nu = \sum_{j=1}^{p} \left(\log_2 \left(\frac{k}{j} \right) \right) = \log_2 \left(\prod_{j=1}^{p} \frac{k}{j} \right) = \log_2 \left(\frac{k^p}{p!} \right) \leqslant \log_2 \left(\frac{k^p}{2p^{p+1/2}e^{-p}} \right) =$$
$$= \log_2 \left(\frac{(2k)^p}{2p^{p+1/2}e^{-p}2^p} \right) \leqslant C_4 k,$$

применив тот факт, что функция $f(x) = x(\log_2(2k) - \log_2(x))$ возрастает на отрезке $[1, k]$.

В результате имеем оценку $\varepsilon_{p+1} < 2^{-C_5 k}$. Так как $k \leqslant n$ (где n берётся из формулы (3)), то для вычислений $a_{(p,k)}$ можно взять точности ε_{p+1}, такие, что $\varepsilon_{p+1} < 2^{-C_5 n}$. Это означает FP//LINSPACE вычислимость коэффициентов $a^{(1)}_k$ ряда Тейлора (2) функции $W^{(1)}_0(x)$ (так как m линейно зависит от n для линейно сходящихся степенных рядов, а m такое, что 2^{-m} — точность вычисления аргумента x [2]).

Так как ряд Тейлора (2) линейно сходится на отрезке $[-(re)^{-1}, (re)^{-1}]$, где r — рациональное, $r > 4/3$, то воспользуемся алгоритмом $SeriesSum_1$ из [2] для вычисления данного ряда в пределах FP//LINSPACE. Алгоритм $SeriesSum_1$ позволяет вычислять с полиномиальным временем и линейной памятью на машине Тьюринга линейно сходящиеся степенные ряды вида $S(x) = \sum\limits_{i=0}^{\infty} a_i x^i$ на любом отрезке $\sigma \subseteq [-\varrho, \varrho]$ при условии, что $|a_i| \leqslant 1$, $\varrho \leqslant 3/4 + 2^{-5}$ и величины a_k являются FP//LINSPACE вычислимыми. Так как все условия, при которых алгоритм $SeriesSum_1$ применим, выполняются для ряда Тейлора функции $W^{(1)}_0(x)$ (а значит, и для ряда Тейлора функции W_0), то верна следующая теорема.

Теорема 1. Основная ветвь W_0 вещественной W-функции Ламберта является FP//LINSPACE алгоритмической вещественной функцией на любом отрезке $[-(re)^{-1}, (re)^{-1}]$ FP//LINSPACE алгоритмических вещественных чисел, где r — рациональное, $r > 4/3$.

Заключение

Алгоритм WLE расчета вещественной W-функции Ламберта W_0 можно применять в информатике как основу FP//LINSPACE алгоритмической вещественной W-функции Ламберта W_0, заданной на отрезке $[-(re)^{-1}, (re)^{-1}]$ FP//LINSPACE алгоритмических вещественных чисел, где r — рациональное, $r > 4/3$.

Из дальнейших исследований стоит отметить задачу построения алгоритмов, основанных на разложении в ряды, для FP//LINSPACE алгоритмических аналогов других вещественных функций, не выражающихся через элементарные функции.

ЛИТЕРАТУРА

1. *Дубинов А. Е., Дубинова И. Д., Сайков С. К.* W-функция Ламберта и ее применение в математических задачах физики. Саров: Изд-во ФГУП «РФЯЦ-ВНИИЭФ», 2006. 160 с.
2. *Яхонтов С. В., Косовский Н. К., Косовская Т. М.* Эффективные по времени и памяти алгоритмические приближения чисел и функций. Учеб. пособие. СПб.: Изд-во СПбГУ, 2012. 256 с.
3. *Ko K.* Complexity Theory of Real Functions. Boston: Birkhauser, 1991. 310 p.
4. *Du D. and Ko K.* Theory of Computational Complexity. N. Y.: John Wiley & Sons, 2000. 491 p.
5. *Brent R. P.* Fast multiple-precision evaluation of elementary functions // J. ACM. 1976. V. 23. No. 2. P. 242–251.
6. *Карацуба Е. А.* Быстрые вычисления трансцендентных функций // Проблемы передачи информации. 1991. Т. 27. Вып. 4. С. 76–99.
7. *Haible B. and Papanikolaou T.* Fast multiprecision evaluation of series of rational numbers // Proc. Third Intern. Symposium on Algorithmic Number Theory, Portland, Orgeon, USA, June 21–25, 1998. P. 338–350.
8. *Mezzarobba M.* A note on the space complexity of fast D-finite function evaluation // Computer Algebra in Scientific Comput. 2012. V. 7442. P. 212–223.
9. *Старицын М. А., Яхонтов С. В.* Эффективное по времени и памяти вычисление W-функции Ламберта // Четвертая Всерос. науч. конф. по проблемам информатики СПИСОК-14. СПб., 2014 (в печати).
10. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления. Т. 2. М.: Физматлит, 2003. 680 с.

ДИСКРЕТНЫЕ МОДЕЛИ РЕАЛЬНЫХ ПРОЦЕССОВ

DOI 10.17223/20710410/25/12

УДК 519.8

ПОСТОПТИМАЛЬНЫЙ АНАЛИЗ ИНВЕСТИЦИОННОЙ ЗАДАЧИ С КРИТЕРИЯМИ КРАЙНЕГО ОПТИМИЗМА[1]

В. А. Емеличев, Е. В. Устилко

Белорусский государственный университет, г. Минск, Беларусь

E-mail: emelichev@bsu.by, ustilko@tut.by

Получены нижняя и верхняя оценки радиуса устойчивости многокритериальной инвестиционной булевой задачи с критериями крайнего оптимизма в случае, когда в пространстве состояний финансового рынка и критериальном пространстве экономической эффективности инвестиционных проектов задана произвольная метрика Гельдера, а в пространстве проектов — метрика Чебышева.

Ключевые слова: *многокритериальная инвестиционная задача, критерий крайнего оптимизма, множество Парето, радиус устойчивости задачи, метрика Гельдера, метрика Чебышева.*

Введение

Многие проблемы принятия многоцелевых решений (индивидуальных или групповых) в управлении, планировании и проектировании могут быть сформулированы как многокритериальные (векторные) задачи дискретной оптимизации. Решение таких задач сводится к выбору лучших в том или ином смысле значений переменных из некоторой дискретной совокупности, что определяется экономическим или физическим смыслом изучаемых проблем. Характерной особенностью подобных задач, возникающих на практике, является неточность исходных данных. Эта неточность обусловлена влиянием различных факторов неопределённости и случайности. Порой сколь угодные малые погрешности в исходной информации влекут значительные искажения искомых решений. Такие задачи обычно называются некорректно поставленными, т. е. являются неустойчивыми к малым изменениям исходных данных, их решение может быть лишено смысла [1]. При этом естественно возникает вопрос: в каких пределах можно варьировать (возмущать) исходные данные задачи, чтобы множество оптимальных решений обладало некоторым свойство инвариантности? Этой проблематике и посвящена настоящая работа, где для многокритериальной задачи формирования оптимального портфеля с критериями крайнего оптимизма по доходности получены нижняя и верхняя оценки радиуса устойчивости задачи в случае, когда в пространстве состояний финансового рынка и критериальном пространстве эффективности инвестиционных портфелей задана произвольная метрика Гельдера l_p, $1 \leqslant p \leqslant \infty$.

Отметим, что ранее в [2, 3, 4, 5] аналогичные оценки (снизу и сверху) радиуса устойчивости многокритериальных инвестиционных задач с критериями Вальда и Сэвиджа

[1]Работа частично поддержана грантом Белорусского республиканского фонда фундаментальных исследований № Ф13К-078.

были получены лишь в частных случаях, когда в упомянутом трехмерном пространстве параметров задач задавались линейная l_1 и чебышевская l_∞ метрики в различных комбинациях. Кроме того, в [6] дан обзор результатов, связанных с оценками радиуса устойчивости парето-оптимальных и лексикографических оптимальных портфелей тех же инвестиционных задач с разнообразными метриками, заданными в пространствах их параметров.

1. Постановка задачи и определения

Рассмотрим многокритериальный вариант задачи управления инвестициями. Для этого введём ряд обозначений.

Пусть известны m возможных состояний финансового рынка $(A_1, A_2, \dots, A_m)$, n альтернативных инвестиционных проектов $(B_1, B_2, \dots, B_n)$ и s видов (показателей) экономической эффективности проекта $(C_1, C_2, \dots, C_s)$. Задана ожидаемая оценка экономической эффективности e_{ijk} любого вида C_k всякого инвестиционного проекта B_j в случае, когда рынок находится в состоянии A_i. Через E будем обозначать трехмерную матрицу $[e_{ijk}] \in \mathbb{R}^{m\times n\times s}$, а через $E_k \in \mathbb{R}^{m\times n}$ — её k-е сечение. Пусть $x = (x_1, x_2, \dots, x_n)^{\mathrm{T}} \in \mathbb{E}^n$ — инвестиционный портфель, где $\mathbb{E} = \{0, 1\}$; $x_j = 1$, если инвестор выбирает проект B_j, и $x_j = 0$ в противном случае; $X \subseteq \mathbb{E}^n$ — множество всех возможных инвестиционных портфелей, т. е. тех, реализация которых не превосходит начального капитала инвестора. Отметим, что существует несколько подходов при оценке эффективности инвестиционных проектов (см., например, библиографию в [7]).

На множестве портфелей X зададим векторную целевую функцию

$$f(x, E) = (f_1(x, E_1), f_2(x, E_2), \dots, f_s(x, E_s)),$$

компонентами которой являются широко известные в теории принятия решений критерии крайнего оптимизма (MAXMAX):

$$f_k(x, E_k) = \max_{1\leqslant i\leqslant m} e_{ik}x = \max_{1\leqslant i\leqslant m} \sum_{j=1}^{n} e_{ijk}x_j \to \max, \quad k \in N_s = \{1, 2, \dots, s\},$$

где $e_{ik} = (e_{i1k}, e_{i2k}, \dots, e_{ink})$ — i-я строка сечения E_k. С помощью такого критерия азартный инвестор оптимизирует эффективность $e_{ik}x$ портфеля x в предположении, что рынок находится в самом выгодном для него состоянии, а именно, когда доходность портфеля максимальна. Очевидно, что подобный подход основан на стереотипе поведения безоглядного оптимизма («или пан или пропал», «кто не рискует, тот не выигрывает» и т. п.). Следует отметить, что ситуации, требующие применения такого критерия, в экономической практике не являются редкими, и пользуются им не только крайние оптимисты, но и инвесторы, поставленные в безвыходное положение.

Под многокритериальной инвестиционной булевой задачей $Z^s(E)$, $s \in \mathbb{N}$, будем понимать задачу поиска множества Парето $P^s(E)$, т. е. множества парето-оптимальных инвестиционных портфелей

$$P^s(E) = \{x \in X : X(x, E) = \varnothing\},$$

где $X(x, E) = \{x' \in X : f(x, E) \leqslant f(x', E) \quad \& \quad f(x, E) \neq f(x', E)\}$.

Очевидно, что $P^s(E) \neq \varnothing$ при любой матрице $E \in \mathbb{R}^{m\times n\times s}$.

Для всякого натурального числа d в действительном пространстве $\mathbb{R}^d$ зададим метрику Гельдера l_p, $p \in [1, \infty]$, т. е. под нормой вектора $y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$ будем

понимать число

$$\|y\|_p = \begin{cases} \left(\sum\limits_{i=1}^{d} |y_i|^p\right)^{1/p}, & \text{если} \quad 1 \leqslant p < \infty, \\ \max\limits_{1 \leqslant i \leqslant d} |y_i|, & \text{если} \quad p = \infty. \end{cases}$$

В пространствах $\mathbb{R}^m$ и $\mathbb{R}^s$ зададим произвольную метрику Гельдера l_p, $1 \leqslant p \leqslant \infty$, а в пространстве $\mathbb{R}^n$ — метрику Чебышева l_∞, т. е. полагаем

$$\|E\|_{\infty pp} = \|(\|E_1\|_{\infty p}, \|E_2\|_{\infty p}, \ldots, \|E_s\|_{\infty p})\|_p,$$
$$\|E_k\|_{\infty p} = \|(\|e_{1k}\|_\infty, \|e_{2k}\|_\infty, \ldots, \|e_{mk}\|_\infty)\|_p, \quad k \in N_s.$$

Ясно, что

$$\|e_{ik}\|_\infty \leqslant \|E_k\|_{\infty p} \leqslant \|E\|_{\infty pp}, \quad i \in N_m, \quad k \in N_s.$$

Поэтому для любых x, $x' \in X$ и $E \in \mathbb{R}^{m \times n \times s}$ очевидны неравенства

$$e_{ik}x - e_{i'k}x' \geqslant -(\|e_{ik}\|_\infty \|x\|_1 + \|e_{i'k}\|_\infty \|x'\|_1) \geqslant -\|E\|_{\infty pp} \|x + x'\|_1, \;\; i, i' \in N_m, \; k \in N_s. \quad (1)$$

Следуя [2, 3, 4, 5], радиусом устойчивости задач $Z^s(E)$, $s \in \mathbb{N}$, назовём число

$$\rho = \rho(m, n, s, p) = \begin{cases} \sup \Xi_p, & \text{если} \quad \Xi_p \neq \varnothing, \\ 0, & \text{если} \quad \Xi_p = \varnothing, \end{cases}$$

где $\Xi_p = \{\varepsilon > 0 : \forall E' \in \Omega_p(\varepsilon) \; (P^s(E + E') \subseteq P^s(E))\}$; $\Omega_p(\varepsilon) = \{E' \in \mathbb{R}^{m \times n \times s} : \|E'\|_{\infty pp} < \varepsilon\}$ — множество возмущающих матриц; $P^s(E + E')$ — множество Парето возмущённой задачи $Z^s(E + E')$. Таким образом, радиус устойчивости задач $Z^s(E)$ — это предельный уровень возмущений элементов матрицы E в нормированном пространстве $\mathbb{R}^{m \times n \times s}$, которые не приводят к появлению новых парето-оптимальных портфелей. Очевидно, что при $P^s(E) = X$ радиус устойчивости задачи следует считать бесконечным. Задачу, для которой $P^s(E) \neq X$, будем называть нетривиальной.

2. Оценки радиуса устойчивости задачи

Для нетривиальной задачи $Z^s(E)$ положим

$$\varphi = \varphi(m, n, s) = \min_{x \notin P^s(E)} \max_{x' \in P(x,E)} \frac{\gamma(x', x)}{\|x' + x\|_1},$$
$$\psi = \psi(m, n, s) = \min_{x \notin P^s(E)} \max_{x' \in P(x,E)} \frac{\gamma(x', x)}{\|x' - x\|_1},$$

где $\gamma(x', x) = \min\{f_k(x', E_k) - f_k(x, E_k) : k \in N_s\}$; $P(x, E) = X(x, E) \cap P^s(E)$. Легко видеть, что $\varphi, \; \psi \geqslant 0$.

Теорема 1. При любых $m, n, s \in \mathbb{N}$ и $p \in [1, \infty]$ для радиуса устойчивости $\rho(m, n, s, p)$ многокритериальной нетривиальной инвестиционной задачи $Z^s(E)$ справедливы следующие оценки:

$$\varphi(m, n, s) \leqslant \rho(m, n, s, p) \leqslant (ms)^{1/p} \psi(m, n, s). \quad (2)$$

Здесь и далее считаем, что $1/p = 0$, если $p = \infty$.

Доказательство. Сначала покажем справедливость неравенства $\rho \geqslant \varphi$. При $\varphi = 0$ оно очевидно. Пусть $\varphi > 0$ и возмущающая матрица $E' \in \mathbb{R}^{m\times n\times s}$ с сечениями E'_k, $k \in N_s$, принадлежит множеству $\Omega_p(\varphi)$, т. е. $\|E'\|_{\infty pp} < \varphi$. Согласно определению числа φ, для любого портфеля $x \notin P^s(E)$ существует такой портфель $x^0 \in P(x, E)$, что

$$\gamma(x^0, x) \geqslant \varphi\|x^0 + x\|_1,$$

т. е. выполняются неравенства

$$f_k(x^0, E_k) - f_k(x, E_k) \geqslant \varphi\|x^0 + x\|_1, \quad k \in N_s.$$

Поэтому, учитывая неравенства (1), для всякого индекса $k \in N_s$ получаем

$$\begin{gathered} f_k(x^0, E_k + E'_k) - f_k(x, E_k + E'_k) = \max_{1\leqslant i\leqslant m}(e_{ik} + e'_{ik})x^0 - \max_{1\leqslant i\leqslant m}(e_{ik} + e'_{ik})x = \\ = \min_{1\leqslant i\leqslant m}\max_{1\leqslant i'\leqslant m}(e_{i'k}x^0 - e_{ik}x + e'_{i'k}x^0 - e'_{ik}x) \geqslant \\ \geqslant \min_{1\leqslant i\leqslant m}\max_{1\leqslant i'\leqslant m}(e_{i'k}x^0 - e_{ik}x) - \|E'\|_{\infty pp}\|x^0 + x\|_1 = \\ = f_k(x^0, E_k) - f_k(x, E_k) - \|E'\|_{\infty pp}\|x^0 + x\|_1 \geqslant (\varphi - \|E'\|_{\infty pp})\|x^0 + x\|_1 > 0, \end{gathered}$$

где e'_{ik} — i-я строка сечения E'_k. Таким образом, любой портфель x, не содержащийся в $P^s(E)$, не является парето-оптимальным портфелем возмущённой задачи $Z^s(E+E')$. Поэтому заключаем, что при любой возмущающей матрице $E' \in \Omega_p(\varphi)$ справедливо включение $P^s(E + E') \subseteq P^s(E)$. Следовательно, верно неравенство $\rho \geqslant \varphi$.

Докажем неравенство $\rho \leqslant (ms)^{1/p}\psi$. В соответствии с определением величины ψ найдётся такой портфель $x^0 \notin P^s(E)$, что для любого портфеля $x \in P(x^0, E)$ существует индекс $l \in N_s$, при котором

$$f_l(x, E_l) - f_l(x^0, E_l) \leqslant \psi\|x - x^0\|_1. \tag{3}$$

Полагая $\varepsilon > (ms)^{1/p}\psi$, зададим элементы e^0_{ijk} любого k-го сечения E^0_k, $k \in N_s$, возмущающей матрицы E^0 по правилу

$$e^0_{ijk} = \begin{cases} \delta, & \text{если} \quad i \in N_s,\ x^0_j = 1, \\ -\delta & \text{в остальных случаях,} \end{cases}$$

где $\varepsilon/(ms)^{1/p} > \delta > \psi$. Тогда получаем

$$\|e^0_{ik}\|_\infty = \delta, \quad \|E^0_k\|_{\infty p} = m^{1/p}\delta, \quad i \in N_m, \quad k \in N_s; \quad \|E^0\|_{\infty pp} = (ms)^{1/p}\delta.$$

Это значит, что $E^0 \in \Omega_p(\varepsilon)$. Кроме того, все строки e^0_{ik}, $i \in N_m$, любого сечения E^0_k, $k \in N_s$, одинаковы и состоят из компонент δ и $-\delta$. Поэтому, положив $A = e^0_{ik}$, $i \in N_m$, $k \in N_s$, имеем

$$A(x - x^0) = -\delta\|x - x^0\|_1. \tag{4}$$

Отсюда, учитывая (3), выводим, что для любого портфеля $x \in P(x^0, E)$ существует индекс $l \in N_s$, удовлетворяющий соотношениям

$$\begin{gathered} f_l(x, E_l + E^0_l) - f_l(x^0, E_l + E^0_l) = \max_{1\leqslant i\leqslant m}(e_{il} + e^0_{il})x - \max_{1\leqslant i\leqslant m}(e_{il} + e^0_{il})x^0 = \\ = \min_{1\leqslant i\leqslant m}\max_{1\leqslant i'\leqslant m}(e_{i'l}x - e_{il}x^0 + e^0_{i'l}x - e^0_{il}x^0) = f_l(x, E_l) - f_l(x^0, E_l) + A(x - x^0) \leqslant \\ \leqslant (\psi - \delta)\|x - x^0\|_1 < 0. \end{gathered}$$

Таким образом, справедлива формула

$$\forall x \in P(x^0, E) \quad (x \notin X(x^0, E + E^0)). \tag{5}$$

Если $X(x^0, E + E^0) = \varnothing$, то $x^0 \in P^s(E + E^0)$. Напомним, что $x^0 \notin P^s(E)$.

Допустим теперь, что $X(x^0, E + E^0) \neq \varnothing$. Тогда благодаря внешней устойчивости множества $P^s(E + E^0)$ (см., например, [8]) найдётся портфель $x^* \in P(x^0, E + E^0)$. Покажем, что $x^* \notin P^s(E)$.

Допустим обратное: $x^* \in P^s(E)$. Согласно (5), выполняется включение

$$x^* \in P^s(E) \setminus P(x^0, E).$$

Поэтому возможны лишь следующие два случая.

С л у ч а й 1. $f(x^*, E) = f(x^0, E)$. Тогда для любого $k \in N_s$ равенства (4) влекут

$$f_k(x^*, E_k + E_k^0) - f_k(x^0, E_k + E_k^0) = f_k(x^*, E_k) - f_k(x^0, E_k) + A(x^* - x^0) = -\delta\|x^* - x^0\|_1 < 0.$$

С л у ч а й 2. Существует такой индекс $q \in N_s$, что $f_q(x^*, E_q) < f_q(x^0, E_q)$. Тогда, вновь используя (4), приходим к соотношениям

$$f_q(x^*, E_q + E_q^0) - f_q(x^0, E_q + E_q^0) = f_q(x^*, E_q) - f_q(x^0, E_q) + A(x^* - x^0) < 0.$$

В результате и тот, и другой случай противоречат включению $x^* \in P(x^0, E + E^0)$. Тем самым доказано, что $x^* \notin P^s(E)$. Напомним, что $x^* \in P^s(E + E^0)$.

Итак, при любом числе $\varepsilon > (ms)^{1/p}\psi$ гарантируется существование такой возмущающей матрицы $E^0 \in \Omega_p(\varepsilon)$, что найдётся портфель (x^0 или x^*), который одновременно, не являясь парето-оптимальным портфелем задачи $Z^s(E)$, является таковым в возмущённой задаче $Z^s(E + E^0)$. Таким образом, справедлива формула

$$\forall \varepsilon > (ms)^{1/p}\psi \quad \exists E^0 \in \Omega_p(\varepsilon) \quad (P^s(E + E^0) \nsubseteq P^s(E)).$$

Следовательно, $\rho \leqslant (ms)^{1/p}\psi$. ■

Из теоремы 1 вытекает следующий известный результат.

Следствие 1 [2]. $\varphi(m, n, s) \leqslant \rho(m, n, s, \infty) \leqslant \psi(m, n, s)$.

О достижимости этих оценок свидельствует слудующее очевидное утверждение.

Следствие 2. Если для любой пары портфелей $x \notin P^s(E)$ и $x' \in P(x, E)$ выполняется равенство

$$\{j \in N_n : x_j = x'_j = 1\} = \varnothing,$$

то справедлива формула

$$\rho(m, n, s, \infty) = \varphi(m, n, s) = \psi(m, n, s).$$

В случае $m = 1$ многокритериальная инвестиционная задача $Z^s(E)$ превращается в s-критериальную задачу линейного булева программирования $Z_L^s(E)$:

$$Ex = (e_1 x, e_2 x, \ldots, e_s x)^{\mathrm{T}} \to \max_{x \in X},$$

где $e_k = (e_{k1}, e_{k2}, \ldots, e_{kn})$ — k-я строка матрицы $E = [e_{kj}] \in \mathbb{R}^{s \times n}$, $X \in \mathbb{E}^n$. Очевидно, что такой случай можно интерпретировать как ситуацию, при которой состояние рынка не вызывает сомнений. При этом, как и ранее, в критериальном пространстве $\mathbb{R}^s$ задана произвольная метрика Гельдера l_p, а в пространстве портфелей $\mathbb{R}^n$ — метрика Чебышева l_∞.

Следствие 3. При любых n, $s \in \mathbb{N}$ и $p \in [1, \infty]$ справедливы оценки

$$\rho(1, n, s, p) \leqslant s^{1/p}\psi(1, n, s) = s^{1/p} \min_{x \notin P^s(E)} \max_{x' \in P(x,E)} \min_{k \in N_s} \frac{e_k(x' - x)}{\|x' - x\|_1}.$$

Достижимость этой оценки докажем путём построения соответствующего класса задач.

Теорема 2. Существует такой класс задач линейного булева программирования $Z_L^s(E)$, что справелива формула

$$\rho(1, n, s, p) = s^{1/p}\psi(1, n, s), \quad n, s \in \mathbb{N}, \quad p \in [1, \infty]. \tag{6}$$

Доказательство. Согласно следствию 3, для доказательства равенств (6) достаточно указать класс задач $Z^s(E)$ с условием $\rho(1, n, s, p) \geqslant s^{1/p}\psi(1, n, s)$.

Пусть $X = \{x^1, x^2, \ldots, x^n\} \subset \mathbb{E}^n$, где $n = s + 1$; x^j — j-й столбец единичной $(n \times n)$-матрицы. Пусть матрица $E = [e_{kj}] \in \mathbb{R}^{s \times n}$ со строками e_k, $k \in N_s$, имеет вид

$$E = \begin{pmatrix} 0 & M & \cdots & M & -2\alpha \\ M & 0 & \cdots & M & -2\alpha \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ M & M & \cdots & 0 & -2\alpha \end{pmatrix},$$

где $M \gg \alpha > 0$; M — достаточно большое число. Тогда

$$\begin{gathered} Ex^1 = (0, M, \ldots, M, M)^{\mathrm{T}} \in \mathbb{R}^s, \\ Ex^2 = (M, 0, \ldots, M, M)^{\mathrm{T}} \in \mathbb{R}^s, \\ \cdots \\ Ex^{n-1} = (M, M, \ldots, M, 0)^{\mathrm{T}} \in \mathbb{R}^s, \\ Ex^n = (-2\alpha, -2\alpha, \ldots, -2\alpha)^{\mathrm{T}} \in \mathbb{R}^s. \end{gathered}$$

Поэтому $x^n \notin P^s(E)$, $x^j \in P(x^n, E)$, $j \in N_s$. Кроме того, справедливы равенства

$$\begin{gathered} P^s(E) = X \setminus \{x^n\} = \{x^1, x^2, \ldots, x^s\}, \\ \psi(1, n, s) = \max_{j \in N_s} \min_{k \in N_s} \frac{e_k(x^j - x^n)}{2} = \alpha. \end{gathered}$$

Пусть $E' = [e'_{kj}] \in \Omega_p(s^{1/p}\alpha)$ — произвольная возмущающая матрица со строками $e'_1, e'_2, \ldots, e'_s$, т. е. $E' \in \mathbb{R}^{s \times n}$, $\|E'\|_{\infty p} < s^{1/p}\alpha$. Методом от противного легко доказать, что существует индекс $v \in N_s$, подчинённый неравенству $\|e'_v\|_\infty < \alpha$. Поэтому $|e'_{vj}| < \alpha$ при любом индексе $j \in N_n$. Отсюда имеем

$$(e_v + e'_v)(x^v - x^n) = 2\alpha + e'_{vv} - e'_{vn} > 2\alpha - |e'_{vv}| - |e'_{vn}| > 0,$$

а для каждого индекса $k \in N_s \setminus \{v\}$ выводим

$$(e_k + e'_k)(x^v - x^n) = e_k(x^v - x^n) + e'_k(x^v - x^n) = M + 2\alpha + e'_{kv} - e'_{kn} > 0.$$

Резюмируя, заключаем, что $x^n \notin P^s(E + E')$ при любой возмущающей матрице $E' \in \Omega_p(s^{1/p}\alpha)$. Следовательно, $\rho(1, n, s, p) \geqslant s^{1/p}\psi(1, n, s)$. ■

О достижимости верхней оценки (2) при $m = 1$ и $p = \infty$ свидетельствует следующая известная теорема.

Теорема 3 [9]. $\rho(1, n, s, \infty) = \psi(1, n, s)$, n, $s \in \mathbb{N}$.

Заключение

Поскольку рыночной экономике присущ динамизм и высокая степень неопределённости, то фактор риска — неотъемлемый атрибут финансового рынка. В настоящей работе на основе портфельной теории сформулирована многокритериальная (векторная) инвестиционная булева задача с паретовским принципом оптимальности, в которой эффективность выбираемого инвестором портфеля оценивается векторной целевой функцией, состоящей из критериев крайнего оптимизма, присущего безоглядному игроку. При этом фактор неопределённости и неточности исходной информации предлагается учитывать путём указания пределов надёжности принимаемых инвестором решений, т. е. с помощью оценок радиуса устойчивости множества Парето. В результате проведённого параметрического анализа задачи получены нижняя и верхняя оценки радиуса устойчивости в случае, когда в пространстве состояний финансового рынка $\mathbb{R}^m$ и критериальном пространстве показателей экономической эффективности проектов $\mathbb{R}^s$ задана произвольная метрика Гельдера l_p, $1 \leqslant p \leqslant \infty$, а в пространстве проектов $\mathbb{R}^n$ — метрика Чебышева l_∞. Оказалось, что нижняя оценка не зависит от величины p, а верхняя с возрастанием числа p от 1 до ∞ уменьшается в ms раз.

ЛИТЕРАТУРА

1. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. М.: Наука, 1986. 286 с.
2. *Емеличев В. А., Коротков В. В.* О радиусе устойчивости векторной инвестиционной задачи с критериями минимаксного риска Сэвиджа // Кибернетика и системный анализ. 2012. № 3. С. 68–77.
3. *Емеличев В. А., Коротков В. В.* Устойчивость векторной инвестиционной булевой задачи с критериями Вальда // Дискретная математика. 2012. Т. 24. № 3. С. 3–16.
4. *Емеличев В. А., Коротков В. В.* О мере устойчивости многокритериальной инвестиционной задачи с критериями эффективности Вальда // Известия НАН Азербайджана. Сер. физ.-тех. и матем. наук. 2012. Т. 32. № 6. С. 88–98.
5. *Emelichev V. and Korotkov V.* On stability radius of the multicriteria variant of Markowitz's investment portfolio problem // Bulletin of the Academy of Sciences of Moldova. Mathematics. 2011. No. 1. P. 83–94.
6. *Емеличев В. А., Котов В. М., Кузьмин К. Г. и др.* Устойчивость и эффективные алгоритмы решения задач дискретной оптимизации с многими критериями и неполной информацией // Проблемы управления и информатики. 2014. № 1. С. 53–67.
7. *Емеличев В. А., Коротков В. В.* Исследование устойчивости решений векторной инвестиционной булевой задачи в случае метрики Гельдера в критериальном пространстве // Прикладная дискретная математика. 2012. № 4. С. 61–72.
8. *Подиновский В. В., Ногин В. Д.* Парето-оптимальные решения многокритериальных задач. 2-е изд., испр. и доп. М.: Физматлит, 2007. 256 с.
9. *Емеличев В. А., Подкопаев Д. П.* О количественной мере устойчивости векторной задачи целочисленного программирования // Журнал вычислительной математики и математической физики. 1998. Т. 38. № 11. С. 1801–1805.

СВЕДЕНИЯ ОБ АВТОРАХ

БОНДАРЬ Евгения Алексеевна — аспирантка кафедры математического анализа и алгебры Луганского национального университета им. Тараса Шевченко, г. Луганск, Украина. E-mail: **bondareug@gmail.com**

ЕМЕЛИЧЕВ Владимир Алексеевич — профессор, доктор физико-математических наук, профессор кафедры математической кибернетики Белорусского государственного университета, г. Минск. E-mail: **emelichev@bsu.by**

ЖАРКОВА Анастасия Владимировна — кандидат физико-математических наук, доцент кафедры теоретических основ компьютерной безопасности и криптографии Саратовского государственного университета им. Н. Г. Чернышевского, г. Саратов. E-mail: **VAnastasiyaV@gmail.com**

ЗАЕЦ Мирослав Владимирович — сотрудник ФГУП НИИ «КВАНТ», г. Москва. E-mail: **mirzaets@hotmail.com**

КАЛИННИКОВ Иван Сергеевич — аспирант Национального исследовательского университета «МИЭТ», г. Москва. E-mail: **gaminot@gmail.com**

КОЛОМЕЕЦ Николай Александрович — аспирант Института математики им. С. Л. Соболева СО РАН, г. Новосибирск. E-mail: **nkolomeec@gmail.com**

КЯЖИН Сергей Николаевич — аспирант кафедры криптологии и дискретной математики Национального исследовательского ядерного университета «МИФИ», г. Москва. E-mail: **s.kyazhin@kaf42.ru**

МОНАХОВ Олег Геннадьевич — кандидат технических наук, старший научный сотрудник, ведущий научный сотрудник Института вычислительной математики и математической геофизики СО РАН, г. Новосибирск. E-mail: **monakhov@rav.sscc.ru**

МОНАХОВА Эмилия Анатольевна — кандидат технических наук, доцент, старший научный сотрудник Института вычислительной математики и математической геофизики СО РАН, г. Новосибирск. E-mail: **emilia@rav.sscc.ru**

НАЗАРОВ Максим Николаевич — ассистент кафедры ВМ-1 Национального исследовательского университета «МИЭТ», г. Москва. E-mail: **Nazarov-Maximilian@yandex.ru**

ОСИПОВ Дмитрий Юрьевич — аспирант Саратовского государственного университета им. Н. Г. Чернышевского, г. Саратов. E-mail: **st_hill@mail.ru**

СТАРИЦЫН Максим Анатольевич — студент кафедры информатики математико-механического факультета Санкт-Петербургского государственного университета, г. Санкт-Петербург. E-mail: **m.staritzyn2012@yandex.ru**

УСТИЛКО Екатерина Валерьевна — студентка 5-го курса кафедры математической кибернетики Белорусского государственного университета, г. Минск. E-mail: **ustilko@tut.by**

ФОМИЧЕВ Владимир Михайлович — доктор физико-математических наук, профессор, профессор Финансового университета при Правительстве Российской Федерации, профессор Национального исследовательского ядерного университета «МИФИ», г. Москва. E-mail: **fomichev@nm.ru**

ШОЛОМОВ Лев Абрамович — доктор физико-математических наук, профессор, главный научный сотрудник Института системного анализа РАН, г. Москва.
E-mail: **sholomov@isa.ru**

ЯХОНТОВ Сергей Викторович — кандидат физико-математических наук, доцент кафедры информатики математико-механического факультета Санкт-Петербургского государственного университета, г. Санкт-Петербург.
E-mail: **SergeyV.Yakhontov@gmail.com**

АННОТАЦИИ СТАТЕЙ НА АНГЛИЙСКОМ ЯЗЫКЕ

Bondar E. A. **ON THE REGULARITY OF SOME SUBSEMIGROUPS OF EQUIVALENCE RELATION'S ENDOMORPHISM MONOID.** For the set of halfstrong (locally strong, quasi-strong) endomorphisms of an equivalence relation graph, the conditions to form a semigroup are found. Thus the answer to the question put by M. Böttcher and U. Knauer's is given. The conditions for regularity of such semigroups are found too.

Keywords: *regularity, semigroup, endomorphism, equivalence.*

Zaets M. V. **FUNCTIONS WITH VARIATIVE-COORDINATE POLYNOMIALITY OVER PRIMARY RINGS OF RESIDUES.** A new class of functions over primary ring of residues called the functions with variative-coordinate polynomiality is considered. This class generalizes the class of polynomial functions with the property that every system of equations composed of functions in the class may be solved by the coordinate linearization method.

Keywords: *primary ring of residues, polynomial functions, formal derivative, system of equations, VCP-functions.*

Kolomeec N. A. **AN UPPER BOUND FOR THE NUMBER OF BENT FUNCTIONS AT THE DISTANCE 2^k FROM AN ARBITRARY BENT FUNCTION IN $2k$ VARIABLES.** An upper bound for the number of bent functions at the distance 2^k from an arbitrary bent function in $2k$ variables is obtained. The bound is reached only for quadratic bent functions. A notion of completely affine decomposable Boolean function is introduced. It is proved that only affine and quadratic Boolean functions can be completely affine decomposable.

Keywords: *Boolean functions, bent functions, quadratic bent functions.*

Sholomov L. A. **ON THE CONCEPT OF UNDERDETERMINED ALPHABETS OF EQUAL STRENGTH.** For underdetermined alphabets, the following two concepts are defined: a) one alphabet is stronger than another, and b) two alphabets have equal strength. To define concepts (a) and (b), several approaches are used. The functional approach is based on expressibility of one alphabet via another; three other approaches — combinatorial, probabilistic, and algorithmic — are terminologically connected with the Kolmogorov's approaches to the notion of the amount of information. It is proved that all these approaches to the comparison of alphabets are equivalent. In case (b), a solution of the optimal compression problem for one of the alphabets, in fact, solves the same problem for the other. It is shown that the concepts (a) and (b) allow polynomial time verification.

Keywords: *underdetermined alphabet, alphabets of equal strength, entropy of underdetermined data, Kolmogorov complexity.*

Zharkova A. V. **ATTRACTORS IN FINITE DYNAMIC SYSTEMS OF BINARY VECTORS ASSOCIATED WITH PALMS ORIENTATIONS.** Attractors in finite dynamic systems of binary vectors associated with palms orientations are described and states belonging to attractors are characterized. The states of such a system are all the possible orientations of some palm, and evolutionary function transforms a given palm

orientation by reversing all arcs that enter the sinks.
Keywords: *attractor, binary vector, finite dynamic system, palm, starlike tree.*

Kyazhin S. N., Fomichev V. M. **LOCAL PRIMITIVENESS OF GRAPHS AND NONNEGATIVE MATRICES.** Some important properties of objects simulated by nonnegative matrices (graphs) are revealed when their submatrices are positive (subgraphs are complete). For this reason, the primitiveness and the exponent of a matrix (graph) are generalized to the local primitiveness and to the quasiprimitiveness of nonnegative matrices and graphs. Conditions for matrix local primitiveness and quasiprimitiveness are obtained. A relation between local exponent and exponent is established.
Keywords: *exponent, local exponent, local subexponent, local quasiexponent, primitive matrix, local primitiveness.*

Monakhova E. A., Monakhov O. G. **ON THE PROBLEM OF CIRCULANT NETWORKS WITH THE MAXIMAL NUMBER OF NODES FOR ANY DIAMETER.** For undirected circulant networks, the problem of the maximal reachable number of nodes under given dimension and diameter of a graph is considered. In 1994, F. P. Muga proved the theorem that this number is odd for any dimension and any diameter of a circulant graph. Later, R. R. Lewis has presented a counterexample of four-dimensional circulant. In the present paper, a mistake in the proof of this theorem is pointed. Based on the new results, the early presented table of the maximal reachable orders of four-dimensional circulants is corrected.
Keywords: *undirected circulant graphs, diameter, maximum order of a graph.*

Nazarov M. N. **ALTERNATIVE APPROACHES TO THE DESCRIPTION OF CLASSES OF ISOMORPHIC GRAPHS.** An algorithm for natural indexing of automorphic equivalence classes of vertices and edges in finite graphs is proposed. Using this indexing, the alternative description of graph isomorphism classes is constructed. It is also demonstrated that one can apply such classical concepts as colouring, operations on graphs and subgraphs to the graph isomorphism classes.
Keywords: *graph isomorphism, automorphic equivalence classes of vertices, automorphic equivalence classes of edges, graph invariants.*

Osipov D. U. **ON A COUNTEREXAMPLE FOR A T-IRREDUCIBLE EXTENSIONS OF STARLIKE TREES.** T-irreducible extension of a graph G is an extension of the graph G which is obtained by removing maximal set of edges from the trivial extension of G. Here, counterexample is shown for the method by F. Harary and M. Khurum for constructing one of T-irreducible extensions for star-like trees. Besides, all nonisomorphic T-irreducible extensions are constructed for star-like trees with rays of equal length.
Keywords: *graph, T-irreducible extension, star-like trees, star-like trees with rays of equal length.*

Kalinnikov I. S. **COMPUTATIONAL COMPLEXITY OF THE SYNTHESIS OF COMPOSITE MODELS FOR LIPSCHITZ-BOUNDED FUNCTIONS.** The paper is devoted to the analysis of computational complexity of the synthesis of composite models for Lipschitz-bounded surjective functions of single variable. Composite models are some function approximation methods based on approximating via composition of functions taken from a given set. In this paper, it is proved that the problem of building strictly optimal composite model for a target functions via a given set of functions is NP-complete. Methods that are capable to build a near-optimal composition model are discussed. Some of these methods can be realized as algorithms with the polynomial computational com-

plexity but they have a limited application.
Keywords: *function composition, composition models, NP-completeness, Lipschitz-bounded, computational complexity.*

Staritsyn M. A., Yakhontov S. V. **FP//LINSPACE EVALUATION OF REAL LAMBERT W-FUNCTION W_0.** In the paper, we construct FP//LINSPACE algorithmic analog of real Lambert W-function $W_0(x)$ on segment $[-(re)^{-1}, (re)^{-1}]$ of FP//LINSPACE algorithmic real numbers, where r is a rational number, $r > 4/3$ (any such rational number is suitable). To construct algorithmic analog of real Lambert W-function $W_0(x)$, we consider algorithm WLE for the evaluation of dyadic rational approximations of the function on segment $[-(re)^{-1}, (re)^{-1}]$ on Turing machine using polynomial time and linear space. Algorithm WLE is based on the Taylor series expansion of the function; it is shown that the Taylor series of real Lambert W-function $W_0(x)$ on segment $[-(re)^{-1}, (re)^{-1}]$ converges linearly. This fact is used in the algorithm.
Keywords: *real Lambert W-function W_0, algorithmic real functions, Turing machine, polynomial time complexity, linear space complexity.*

Emelichev V. A., Ustilko E. V. **POSTOPTIMAL ANALYSIS OF MULTICRITERIA INVESTMENT PROBLEM WITH THE EXTREME OPTIMISM CRITERIA.** For the stability radius of the multicriteria investment Boolean problem with the extreme optimism criteria, some lower and upper bounds are obtained in the case of the Hölder metric in the criteria and financial market state space and the Chebyshev metric in the portfolio space.
Keywords: *multicriteria investment problem, extreme optimism criterion, Pareto set, stability radius of problem, the Hölder metric, the Chebyshev metric.*