S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru

Том 11, стр. 144–144 (2014) УДК 510.652
MSC 11U99

О ГЕНЕРИЧЕСКОЙ СЛОЖНОСТИ ПРОБЛЕМЫ ПРОВЕРКИ ТОЖДЕСТВ В КОНЕЧНЫХ ГРУППАХ И МОНОИДАХ

А.Н. РЫБАЛОВ

Abstract. Kapovich, Myasnikov, Schupp and Shpilrain in 2003 developed generic approach to algorithmic problems, which considers an algorithmic problem on most of the inputs instead of the entire domain and ignores it on the rest of inputs. This approach can be applied to algorithmic problems, which are hard in the classical sense. The problem of checking identities in algebraic structures is the one of the most fundamental problem in algebra. For finite algebraic structures this problem can be decidable in polynomial time, or hard (co-NP-complete). In this paper we present a generic polynomial algorithms for the identity problem in all finite groups, monoids with elements of period greater than 1, and in finite Brandt monoids B_n^1.

Keywords: generic complexity, identity problem, finite groups, finite monoids, Brantd monoids.

1. Введение

Проблема изучения и описания тождеств в различных алгебраических системах является одной из фундаментальных и классических проблем в алгебре. Для полугрупп одним из мировых лидеров в этой области является свердловская школа, много лет возглавляемая Л. Н. Шевриным. Здесь можно отметить работы М. В. Волкова, М. Сапира, А. Н. Трахтмана, Е. И. Клеймана и др. Полученным результатам по тождествам в полугруппах посвящен классический обзор Л. Н. Шеврина и М. В. Волкова [16]. М. Сапир [12] поставил проблему о вычислительной сложности проблемы проверки тождеств в различных конечных алгебраических системах: группах, полугруппах, кольцах, полях и

Rybalov, A.N., On the generic complexity of the identity problem in finite groups and monoids.

т.д. В последующее десятилетие были получены результаты о полиномиальной разрешимости этой проблемы для ассоциативных нильпотентных колец (Х. Хант, Р. Стирнс [8]), нильпотентных групп (С. Баррис, Дж. Лоуренс [3], Г. Хорват, С. Сабо [7]), коммутативных полугрупп (А. Кисилевич [11]), апериодических 0-простых полугрупп (С. Сейф, С. Сабо [18]), моноидов с менее, чем 6 элементами (О. Клима [13]). С другой стороны были получены результаты о co-NP-полноте этой проблемы для ассоциативных не-нильпотентных колец (С. Баррис, Дж. Лоуренс [2]), неразрешимых групп (Г. Хорват, Дж. Лоуренс, Л. Мераи, С. Сабо [6]), некоторых матричных полугрупп (С. Сабо, В. Вертеши [19]), конечных полугрупп с неразрешимыми подгруппами (Ж. Алмейда, М. В. Волков, С. В. Гольдберг [1]), некоторых 0-простых полугрупп (С. В. Плещева, В. Вертеши [15]), 6-элементного монида Брандта B_2^1 (О. Клима [13], С. Сейф [17]). При условии P $\neq$ NP для co-NP-полной проблемы проверки тождеств не существует полиномиального алгоритма.

Генерический подход был предложен в 2003 году И. Каповичем, А. Г. Мясниковым, В. Шпильрайном и П. Шуппом в работе [10]. В рамках этого подхода изучается поведение алгоритмов на множествах входов, асимптотическая плотность которых равна 1 (эти множества называются генерическими), и игнорируется поведение алгоритма на остальных входах, на которых алгоритм может работать медленно или вообще не останавливаться. Исследования вычислительной сложности для «почти всех» входов началось в 1970-80-х годах, после того как был выделен огромный пласт трудноразрешимых алгоритмических проблем – NP-полных проблем, для которых не удалось найти эффективных алгоритмов, работающих за полиномиальное время для всех входов. Оказалось, что если немного ослабить требование эффективности – рассматривать не все входы, а «почти все» или случайные входы, то иногда можно быстро решать задачу для таких типичных входов. Этот подход имеет практический смысл, когда алгоритм должен решать быстро задачу для случайных входных данных: если вероятность «наткнуться» на «плохой» вход пренебрежимо мала, то алгоритм будет быстро работать практически всегда. Ярким примером такого алгоритма является симплекс-метод: этот алгоритм имеет экспоненциальную сложность в худшем случае, но за полиномиальное время решает задачу линейного программирования для почти всех входных данных.

В данной статье предложены полиномиальные генерические алгоритмы для проблемы проверки тождеств во всех конечных группах, в конечных моноидах с элементами периода больше 1 и в конечных моноидах Брандта B_n^1.

2. Генерические алгоритмы

Пусть I – некоторое множество входов. Для подмножества $S \subseteq I$ определим последовательность

$$\rho_n(S) = \frac{|S_n|}{|I_n|},\ n = 1, 2, 3, \ldots,$$

где I_n – множество входов размера n, а $S_n = S \cap I_n$ – множество входов из S размера n. Заметим, что $\rho_n(S)$ это вероятность попасть в S при случайной и равновероятной генерации входов из I_n. *Асимптотической плотностью* S назовем предел (если он существует)

$$\rho(S) = \lim_{n\to\infty} \rho_n(S).$$

Множество S называется *генерическим*, если $\rho(S) = 1$ и *пренебрежимым*, если $\rho(S) = 0$. Очевидно, что S генерическое тогда и только тогда, когда его дополнение $I \setminus S$ пренебрежимо.

Алгоритм $\mathcal{A}$ с множеством входов I называется *генерическим*, если множество $\{x \in I : \mathcal{A}(x) \downarrow\}$ генерическое. Генерический алгоритм $\mathcal{A}$ вычисляет функцию $f : I \to J$, если

$$\forall x \in I \ \mathcal{A}(x) \downarrow \Rightarrow \ f(x) = \mathcal{A}(x).$$

Генерический алгоритм $\mathcal{A}$ работает за полиномиальное время, если существует полином $p(n)$ такой, что

$$\forall x \in I \ \mathcal{A}(x) \downarrow \Rightarrow t_{\mathcal{A}}(x) < p(size(x)).$$

Еще такие алгоритмы мы будем называть полиномиальными генерическими.

С практической точки зрения, когда требуется построить алгоритм, решающий конкретную алгоритмическую проблему для почти всех входов, удобнее рассматривать алгоритмы следующего типа. Каждый такой алгоритм останавливается на всех входах, на входах из некоторого генерического множества выдает правильный ответ, а на пренебрежимом множестве остальных входов выдает специальный ответ «?» – «Не знаю». Определение такой эффективной генерической вычислимости можно найти в обзоре [5] и в гораздо более ранней работе [14].

Алгоритм $\mathcal{A}$ с множеством входов I и множеством выходов $J \cup \{?\}$ $(? \notin J)$ называется *эффективно генерическим*, если

(1) $\mathcal{A}$ останавливается на всех входах из I,
(2) множество $\{x \in I : \mathcal{A}(x) =?\}$ пренебрежимо.

Эффективно генерический алгоритм $\mathcal{A}$ вычисляет функцию $f : I \to J$, если

$$\forall x \in I \ \mathcal{A}(x) \neq? \ \Rightarrow \ f(x) = \mathcal{A}(x).$$

Множество $S \subseteq I$ и соответствующая проблема распознавания (S, I) *(эффективно) генерически разрешимы*, если существует (эффективно) генерический алгоритм, вычисляющий характеристическую функцию S.

Легко видеть, что из эффективной генерической разрешимости следует генерическая разрешимость. Действительно, любой эффективный генерический алгоритм можно легко переделать в генерический заменив выдачу ответа «?» на бесконечное зацикливание. В обратную сторону это неверно – см., например, теорему 2.22 и следствие 2.24 в [9]. Однако для полиномиальной (экспоненциальной) сложности верно и обратное: из полиномиальной (экспоненциальной) генерической разрешимости следует полиномиальная (экспоненциальная) эффективная разрешимость. Действительно, если имеется полиномиальная оценка $p(n)$ на время работы генерического алгоритма в случае, когда он останавливается, то можно завести счетчик T числа шагов, и, в случае, если $T > p(n)$, можно обрывать вычисление и выдавать ответ «?», – в этом случае генерический алгоритм уже не остановится. Таким образом получается эффективно генерический полиномиальный алгоритм, решающий ту же проблему.

С учетом вышесказанного, в дальнейшем, при доказательстве существования генерического алгоритма, будут строиться эффективно генерические алгоритмы. Из существования эффективного (полиномиального) генерического алгоритма будет следовать существование (полиномиального) генерического алгоритма.

3. Конечные группы

Пусть G – конечная группа. Рассмотрим счетный алфавиты переменных $X = \{x_1, x_2, \ldots\}$ и обратных к ним $X^{-1} = \{x_1^{-1}, x_2^{-1}, \ldots\}$. Обозначим $X_n = \{x_1, \ldots, x_n\}$ и $X_n^{-1} = \{x_1^{-1}, \ldots, x_n^{-1}\}$. *Термом* над группой G называется конечное слово t над алфавитом $\{X_n \cup X_n^{-1}\}$, где $n = |t|$. Под *размером* терма будем подразумевать его длину. Каждый терм t размера n задает функцию $t : G^n \to G$, определенную следующим образом. Для любого $(a_1, \ldots, a_n) \in G^n$ значение $t(a_1, \ldots, a_n)$ вычисляется подстановкой элемента a_i вместо x_i в слово t для каждого $i = 1, \ldots, n$. Обозначим через $\mathcal{T}$ множество всех термов.

Лемма 1. *Для любого n имеем $\mathcal{T}_n = (2n)^n$.*

Доказательство. На каждое из n мест в терме из $\mathcal{T}_n$ можно поставить одну из переменных $x_1, \ldots, x_n$ или обратную переменную $x_1^{-1}, \ldots, x_n^{-1}$. Итого получаем $\mathcal{T}_n = (2n)^n$. □

Терм t размера n называется *тождеством* в G, если для всех $(a_1, \ldots, a_n) \in G^n$ имеет место $t(a_1, \ldots, a_n) = 1$. *Проблема проверки тождеств в группе G* заключается в следующем: По произвольному заданому терму t определить, является ли t тождеством в группе G.

Определим для произвольного терма $t \in \{X \cup X^{-1}\}^*$ и для всех $i = 1, \ldots, n$, где $n = |t|$, число $d_i(t)$ как сумму степеней буквы x_i в слове t.

Лемма 2. *Пусть N – натуральное число больше 1. Множество*

$$\mathcal{T}(N) = \{t \in \mathcal{T} \ : \ d_i(t) \text{ делится на } N, \ i = 1, \ldots, |t|\}$$

пренебрежимо в $\mathcal{T}$.

Доказательство. Для каждого натурального m определим следующее множество

$$\mathcal{T}(N, m) = \{t \in \mathcal{T} \ : \ \Big(\sum_{i \equiv j(\mod m)} d_i(t) \Big) \text{ делится на } N,$$
$$i = 1, \ldots, |t|, \ j = 0, \ldots, m-1\}.$$

Легко видеть, что $\mathcal{T}(N) \subseteq \mathcal{T}(N, m)$ для любого m. Из этого следует, что

$$\rho\Big(\mathcal{T}(N)\Big) \leq \rho\Big(\mathcal{T}(N, m)\Big). \tag{1}$$

Теперь рассмотрим свободную абелеву группу $F(Z_m)$ с m порождающими $Z_m = \{z_1, \ldots, z_m\}$ и определим следующее множество

$$F(Z_m, N) = \{w \in (Z_m \cup Z_m^{-1})^* \ : \ w = \prod_{i=1}^{m} z_i^{a_i}, \text{ в } F(Z_m),$$
$$a_i \text{ делится на } N, \ i = 1, \ldots, m\}.$$

Очевидно, $F(Z_m, N)$ – подгруппа абелевой группы $F(Z_m)$.

Зафиксируем число m. Пусть $n = mk + r$, где $0 \leq r < m$. Для любого $n > m$ определим отображение $\varphi_m : X_n \to Z_m$ по следующему правилу:

$$\varphi_m(x_1) = \varphi_m(x_{m+1}) = \ldots = \varphi_m(x_{mk+1}) = z_1,$$
$$\varphi_m(x_2) = \varphi_m(x_{m+2}) = \ldots = \varphi_m(x_{mk+2}) = z_2,$$
$$\ldots$$
$$\varphi_m(x_r) = \varphi_m(x_{m+r}) = \ldots = \varphi_m(x_{mk+r}) = z_r,$$

$$\varphi_m(x_{r+1}) = \varphi_m(x_{m+r+1}) = \ldots = \varphi_m(x_{m(k-1)+r+1}) = z_{r+1},$$
$$\ldots$$
$$\varphi_m(x_m) = \varphi_m(x_{2m}) = \ldots = \varphi_m(x_{mk}) = z_m,$$

Таким образом, под действием φ_m в переменные $z_1, \ldots, z_r$ переходят $k+1$ переменных из X_n, а в остальные переменные – по k переменных из X_n. В частности, если n делится на m, то в каждую переменную из Z_m переходят поровну, по n/m переменных из X_n. Соответствующие индуцированные отображения для множеств X_n^{-1} и $\mathcal{T}$ тоже будем обозначать через φ_m.

Заметим, что

$$\varphi_m\big(\mathcal{T}(N,m)\big) \subseteq F(Z_m, N).$$

Откуда

$$\mathcal{T}(N,m) \subseteq \varphi_m^{-1}\big(F(Z_m,N)\big).$$

Рассмотрим два случая. Обозначим через r остаток от деления n на m, а через k неполное частное $[n/m]$. То есть $n = mk + r$, $0 \le r < m$.

Случай 1. $r \ge \frac{m}{3}$.

Так как в любой элемент $z_i \in Z_m$ переходят максимум $[n/m]+1$ элементов из X_n, то

$$|\mathcal{T}(N,m)_n| \le |\varphi_m^{-1}\big(F(Z_m,N)_n\big)| \le \big([n/m]+1\big)^n |F(Z_m,N)_n|. \tag{2}$$

Теперь оценим

$$\rho_n\Big(\mathcal{T}(N,m)\Big) = \frac{|\mathcal{T}(N,m)_n|}{|\mathcal{T}_n|} = \frac{|\mathcal{T}(N,m)_n|}{(2n)^n} =$$

$$= \frac{|\mathcal{T}(N,m)_n|}{(2m)^n} \cdot \frac{(2m)^n}{(2n)^n} \le \Big(\frac{m}{n}\Big)^n \cdot \frac{\big(k+1\big)^n |F(Z_m,N)_n|}{|(Z_m \cup Z_m^{-1})_n^*|} =$$

$$= \Big(\frac{m(k+1)}{mk+r}\Big)^n \rho_n\Big(F(Z_m,N)\Big) = \Big(1 + \frac{m-r}{n}\Big)^n \rho_n\Big(F(Z_m,N)\Big) <$$

$$< e^{m-r}\rho_n\Big(F(Z_m,N)\Big) \le e^{2m/3}\rho_n\Big(F(Z_m,N)\Big)$$

для достаточно больших n. Здесь была использована оценка (2), лемма 1 и то, что $|(Z_m \cup Z_m^{-1})_n^*| = (2m)^n$.

Случай 2. $r < \frac{m}{3}$.

Будем называть переменные $z_1, \ldots, z_r$ *плохими*, а $z_{r+1}, \ldots, z_m$ *хорошими*. Теперь переменная x_i из X_n *хорошая*, если $\varphi_m(x_i)$ хорошая, и *плохая*, если $\varphi_m(x_i)$ плохая. Легко видеть, что прообраз $\varphi_m^{-1}(z_i)$ хорошей переменной z_i состоит из k элементов, а для плохой переменной z_i из $k+1$ элементов. Обозначим множество хороших переменных X_n через GX, а плохих через BX. Аналогично GZ и BZ – хорошие и соответственно плохие переменные множества Z_m. Легко видеть, что $|GX| = k(m-r)$, и $|BX| = (k+1)r$. Заметим, что при условии $r < \frac{m}{3}$ хороших переменных в X_n оказывается больше, чем плохих.

Будем называть терм $t \in \mathcal{T}_n$ *плохим*, если число плохих переменных в t больше $\frac{|t|}{2}$. Иначе терм будем называть *хорошим*. Обозначим через $\mathcal{BT}_n$ множество всех плохих термов в $\mathcal{T}_n$, а через $\mathcal{GT}_n$ множество хороших термов. Также через $\mathcal{BT}(N,m)_n$ и $\mathcal{GT}(N,m)_n$ обозначаются множества плохих и соответственно хороших термов в $\mathcal{T}(N,m)_n$. Очевидно, что

$$\mathcal{T}_n = \mathcal{BT}_n \cup \mathcal{GT}_n,\ \mathcal{T}(N,m)_n = \mathcal{BT}(N,m)_n \cup \mathcal{GT}(N,m)_n.$$

Аналогично определяются хорошие и плохие слова в множестве $(Z_m \cup Z_m^{-1})_n^*$. Множества хороших и плохих слов обозначим через $G(Z_m \cup Z_m^{-1})_n^*$ и $B(Z_m \cup Z_m^{-1})_n^*$ соответственно. Также через $BF(Z_m, N)$ и $GF(Z_m, N)$ обозначаются множества плохих и хороших слов в $F(Z_m, N)$.

Докажем, что для любого n

$$|\mathcal{BT}(N,m)_n| \le \frac{1}{2}|\mathcal{T}(N,m)_n|. \tag{3}$$

Действительно, рассмотрим биекцию ψ на X_n такую, что $\psi(BX) \subset GX$. Такая биекция существует, так как хороших переменных больше, чем плохих. Эта биекция индуцирует биекцию на множестве $\mathcal{T}_n$, которую также будем обозначать ψ. Тогда нетрудно видеть, что

(1) $\psi\big(\mathcal{BT}(N,m)_n\big) \subseteq \mathcal{GT}(N,m)_n$,
(2) $|\psi\big(\mathcal{BT}(N,m)_n\big)| = |\mathcal{BT}(N,m)_n|$,
(3) $\psi\big(\mathcal{BT}(N,m)_n\big) \cap \mathcal{BT}(N,m)_n = \varnothing$.

Откуда следует нужное неравенство (3). Из неравенства (3) следует, что

$$|\mathcal{T}(N,m)_n| \le 2|\mathcal{GT}(N,m)_n|. \tag{4}$$

Так как хорошие термы из $\mathcal{T}(N,m)_n$ под действием φ_m переходят в хорошие слова из $F(Z_m,N)_n$, то

$$\mathcal{GT}(N,m)_n \subseteq \varphi_m^{-1}\big(GF(Z_m,N)_n\big).$$

В любую хорошую букву z_i переходят k букв из X_n, в любую плохую $k+1$. Но плохих букв в любом слове из $GF(Z_m,N)_n$ меньше $\frac{n}{2}$, поэтому можно оценить

$$|\mathcal{GT}(N,m)_n| < k^{n/2}(k+1)^{n/2}|GF(Z_m,N)_n| <$$

$$< k^{n/2}(k+1)^{n/2}|F(Z_m,N)_n|. \tag{5}$$

Теперь оценим

$$\rho_n\Big(\mathcal{T}(N,m)\Big) \le \frac{2|\mathcal{GT}(N,m)_n|}{|\mathcal{T}_n|} = \frac{2|\mathcal{GT}(N,m)_n|}{(2n)^n} =$$

$$= \frac{2|\mathcal{GT}(N,m)_n|}{(2m)^n} \cdot \frac{(2m)^n}{(2n)^n} \le \Big(\frac{m}{n}\Big)^n \cdot \frac{2k^{n/2}(k+1)^{n/2}|F(Z_m,N)_n|}{|(Z_m \cup Z_m^{-1})_n^*|} =$$

$$= 2\Big(\frac{mk}{mk+r}\Big)^{n/2}\Big(\frac{m(k+1)}{mk+r}\Big)^{n/2}\rho_n\Big(F(Z_m,N)\Big) \le$$

$$\le 2\Big(1+\frac{m-r}{n}\Big)^{n/2}\rho_n\Big(F(Z_m,N)\Big) <$$

$$< 2e^{(m-r)/2}\rho_n\Big(F(Z_m,N)\Big) \le 2e^{m/2}\rho_n\Big(F(Z_m,N)\Big)$$

для достаточно больших n. Здесь были использованы оценки (4) и (5).

Итого, с учетом двух рассмотренных случаев, имеем оценку

$$\rho_n\Big(\mathcal{T}(N,m)\Big) < e^{2m/3}\rho_n\Big(F(Z_m,N)\Big).$$

Далее

$$\rho\Big(\mathcal{T}(N,m)\Big) = \lim_{n\to\infty}\rho_n\Big(\mathcal{T}(N,m)\Big) \le$$

$$\leq e^{2m/3} \lim_{n\to\infty} \rho_n\Big(F(Z_m, N)\Big) = e^{2m/3}\rho\Big(F(X_m, N)\Big).$$

Так как $F(X_m, N)$ – нормальная подгруппа $F(X_m)$ индекса N^m, то, по теореме 6.3 из [10]

$$\rho\Big(F(X_m, N)\Big) \leq \frac{2}{N^m}.$$

Поэтому

$$\rho\Big(\mathcal{T}(N, m)\Big) \leq 2\frac{e^{2m/3}}{N^m} = 2\Big(\frac{e^{2/3}}{N}\Big)^m$$

для любого натурального m. Так как $N \geq 2$, а $e^{2/3} < 2$, то увеличением m эту верхнюю оценку можно сделать сколь угодно малой. Это означает, с учетом неравенства (1), что $\rho\Big(\mathcal{T}(N)\Big) = 0$. □

Теорема 1. *Пусть G – конечная группа. Проблема проверки тождеств в G генерически разрешима за полиномиальное время.*

Доказательство. Если $|G| = 1$, то группа тривиальна и проблема проверки тождеств очевидно полиномиально разрешима: все термы являются тождествами. Предположим, что $|G| > 1$. Тогда в группе G найдется элемент a порядка $N > 1$. Рассмотрим следующий эффективно генерический полиномиальный алгоритм $\mathcal{A}$. Алгоритм $\mathcal{A}$ работает на входе $t \in \{X \cup X^{-1}\}^*$ следующим образом.

(1) Вычисляет $d_i(t)$, $i = 1, \ldots, |t|$.
(2) Проверяет, делится ли $d_i(t)$ на N для всех $i = 1, \ldots, |t|$.
(3) Если да, то выдает ответ «?».
(4) Если нет, то выдает ответ «НЕТ».

Чтобы доказать, что этот алгоритм корректно решает проблему проверки тождеств в группе G, предположим, что $d_k(t)$ не делится на N для некоторого k. Вычислим значение терма t на наборе $(a_1, \ldots, a_n)$ таком, что $a_k = a$, где a есть элемент группы G порядка N, и $a_i = 1$ для всех $i \neq k$. Легко видеть, что

$$t(a_1, \ldots, a_n) = a^{d_k(t)} \neq 1.$$

Таким образом, терм t не является тождеством в группе G и алгоритм $\mathcal{A}$ выдает правильный ответ «НЕТ». Генеричность алгоритма следует из леммы 2. □

4. Конечные моноиды с элементами периода > 1

Пусть S – конечная полугруппа. Обозначим через $X = \{x_1, x_2, \ldots\}$ счетный алфавит переменных, а через $X_n = \{x_1, \ldots, x_n\}$. *Термом* над алфавитом X_n в полугруппе S называется конечное слово p над X_n. Каждый терм p над X_n определяет функцию $p : S^n \to S$ следующим образом: для любого набора $(a_1, \ldots, a_n) \in S^n$ значение $p(a_1, \ldots, a_n)$ получается подстановкой элемента a_i вместо of x_i в слово p для каждого $i = 1, \ldots, n$. Пара термов p, q размера $n = |p|+|q|$ над алфавитом X_n задает *тождество* над полугруппой S, если для всех $(a_1, \ldots, a_n) \in S^n$ имеет место $p(a_1, \ldots, a_n) = q(a_1, \ldots, a_n)$. *Проблема проверки тождеств* над полугруппой S состоит в следующем. По произвольной заданой паре термов p, q размера $n = |p| + |q|$ над алфавитом X_n определить, являеется ли $p = q$ тождеством над полугруппой S. Обозначим через $\mathcal{PT}$ множество всех пар термов.

Определим для каждой пары термов $p, q \in X^*$ и для каждого $i = 1, \ldots, n$, где $n = |p| + |q|$, $d_i(p, q)$ – число букв x_i в слове p минус число букв x_i в слове q.

Лемма 3. *Пусть N – натуральное число большее 1. Множество*

$$\mathcal{PT}(N) = \{(p, q) \in \mathcal{PT} \ : \ d_i(p, q) \text{ делится на } N, \ i = 1, \ldots, |p| + |q|\}$$

перенебрежимо в $\mathcal{PT}$.

Доказательство. Для каждого натурального m определим следующее множество

$$\mathcal{PT}(N, m) = \{(p, q) \in \mathcal{PT} \ : \ \Big(\sum_{i \equiv j(\mod m)} d_i(t)\Big) \text{ делится на } N,$$

$$i = 1, \ldots, |t|, \ j = 0, \ldots, m - 1\}.$$

Легко видеть, что $\mathcal{PT}(N) \subseteq \mathcal{PT}(N, m)$ для любого m. Из этого следует, что

$$\rho\Big(\mathcal{PT}(N)\Big) \leq \rho\Big(\mathcal{PT}(N, m)\Big). \tag{6}$$

Пусть $Z_m = \{z_1, \ldots, z_m\}$. Рассмотрим коммутативный моноид с сокращениями

$$M(Z_m, N) = \langle z_1, \ldots, z_m \mid z_1^N = 1, \ldots, z_m^N = 1 \rangle.$$

Легко видеть, что $M(Z_m, N)$ – конечная группа и $|M(Z_m, N)| = N^m$.

Пусть $\varphi_m : X_n \to Z_m$ – отображение из доказательства леммы 2. Нетрудно видеть, что

$$(p, q) \in \mathcal{PT}(N, m) \Leftrightarrow \varphi_m(p) = \varphi_m(q) \text{ в моноиде } M(Z_m, N).$$

Определим множество

$$EM(Z_m, N) = \{(w_1, w_2) \in Z_m^* \ : \ w_1 = w_2 \text{ в } M(Z_m, N)\}.$$

Аналогично тому, как это делалось в доказательстве леммы 2, можно получить оценку

$$\rho_n\Big(\mathcal{PT}(N, m)\Big) < e^{2m/3} \rho_n(EM(Z_m, N)). \tag{7}$$

Для дальнейшего нам потребуется теория конечных цепей Маркова. Необходимые сведения из этой теории можно найти в монографии [4].

Построим по моноиду $M(Z_m, N)$ цепь Маркова $MC(M(Z_m, N))$. Состояния этой цепи – элементы моноида $s_1, \ldots, s_{N^m}$. Матрица вероятностей переходов P цепи состоит из вероятностей p_{ij}, где

$$p_{ij} = \begin{cases} \frac{1}{m}, \text{ если } s_i z_k = s_j, \ z_k \in Z_m, \\ 0, \text{ иначе.} \end{cases}$$

Это можно представлять как случайное блуждание по графу Кэли моноида $M(Z_m, N)$ с началом в 1, когда в каждый момент времени, находясь в вершине s_i равновероятно из всех выходящих ребер (соответствующих всем порождающим из Z_m) выбирается ребро, по которому происходит переход в следующую вершину.

Обозначим через $p_n(s_i)$ – вероятность попасть в состояние s_i из начального состояния s_1 за n шагов. *Периодом* состояния s_i цепи Маркова называется число

$$t = \text{НОД}(t_1, t_2, t_3, \ldots, t_k, \ldots),$$

где для любого k имеет место

(1) $p_{t_k}(s_i) > 0$,
(2) $p_l(s_i) = 0$ для любого l такого, что $t_k < l < t_{k+1}$.

То есть через t_1 шагов мы можем впервые с ненулевой вероятностью попасть в состояние s_i, через t_2 – во второй раз, через t_k – в k-й раз и так далее. Для состояния s_i цепи $MC(M(Z_m, N))$ это означает, что в графе Кэли моноида $M(Z_m, N)$ существуют ориентированные пути из вершины s_1 в вершину s_i длин $t_1, t_2, \ldots, t_k, \ldots$ и для любого k не существуют путей длины l, где $t_k < l < t_{k+1}$.

Нетрудно видеть, что цепь $MC(M(Z_m, N))$ имеет период N и разбивается на подцепи $C_0, \ldots, C_{N-1}$, причем

$$|C_0| = \ldots = |C_{N-1}| = N^{m-1}.$$

Эти подцепи $C_0, \ldots, C_{N-1}$ получаются следующим образом. Полагаем $s_1 \in C_0$. Далее в C_0 добавляются все состояния, достижимые из s_1 ровно за N шагов. Потом добавляются все состояния, достижимые ровно за N шагов из добавленных на предыдущем шаге и так до тех пор, пока есть еще не добавленные. Получается цепь C_0. Теперь цепь C_1 состоит из всех состояний, достижимых из всех состояний цепи C_0 ровно за 1 шаг, цепь C_2 – ровно за 2 шага, и так далее. Последняя цепь C_{N-1} содержит состояния, достижимые из C_0 ровно за $N-1$ шагов. Матрицы подцепей $C_0, \ldots, C_{N-1}$ равны P^N.

Матрица называется *стохастической*, если ее элементы неотрицательны и сумма элементов в каждой строке равна 1. Легко видеть, что матрица P цепи Маркова $MC(M(Z_m, N))$ (как и любой другой конечной цепи) является стохастической. Матрица называется *дважды стохастической*, если ее элементы неотрицательны и сумма элементов в каждой строке и в каждом столбце равна 1. Легко проверяется, что произведение двух дважды стохастических матриц является опять дважды стохастической матрицей.

Заметим, что матрицы вероятностей переходов цепи $MC(M(Z_m, N))$ и подцепей $C_0, \ldots, C_{N-1}$ являются дважды стохастическими, так как в любую вершину s графа Кэли моноида $M(Z_m, N)$ входят ровно m ребер, помеченных порождающими $z_1, \ldots, z_m$. Действительно, для каждого порождающего z_i одно ребро, помеченное z_i, входит в вершину s из вершины, соответствующей элементу sz_i^{N-1}. Если два ребра, помеченные z_i, входят в s из двух разных вершин s_l и s_t, то $s = s_l a_i = s_t a_i$, откуда $s_l = s_t$.

Для цепи с дважды стохастической матрицей предельные вероятности одинаковы (см. [4], стр. 386). Поэтому для любого $s \in C_l$ имеет место

(1) $p_t(s) = 0$, если $t \neq l(\mod N)$,
(2) $\lim_{t\to\infty} p_{tN+l}(s) = p(s) = \frac{1}{|C_l|} = \frac{1}{N^{m-1}}$.

Здесь $p_t(s)$ – вероятность того, что после t шагов цепь будет находиться в состоянии s. С другой стороны, это вероятность того, что случайное слово в алфавите Z_m длины t равно элементу s в моноиде $M(Z_m, N)$, то есть

$$p_t(s) = \frac{|eq(s)_t|}{m^t},$$

где $eq(s) = \{w \in Z_m^* \ : \ w = s \text{ в } M(Z_m, N)\}$.

Пусть n достаточно большое, чтобы для $t \geq \sqrt{n}$ выполнялось неравенство

$$\frac{|eq(s)_t|}{m^t} < \frac{2}{N^{m-1}}.$$

Оценим

$$\rho_n(EM(Z_m,N)) = \frac{|EM(Z_m,N)_n|}{(Z_m^* \times Z_m^*)_n} = \frac{|EM(Z_m,N)_n|}{nm^n} =$$

$$= \frac{1}{n}\sum_{k=0}^{n}\sum_{i=1}^{N^m}\Big(\frac{|eq(s_i)_k|}{m^k}\cdot\frac{|eq(s_i)_{n-k}|}{m^{n-k}}\Big) <$$

$$< \frac{2\sqrt{n}}{n} + \frac{1}{n}\sum_{k=[\sqrt{n}]}^{n-[\sqrt{n}]}\sum_{i=1}^{N^m}\Big(\frac{|eq(s_i)_k|}{m^k}\cdot\frac{|eq(s_i)_{n-k}|}{m^{n-k}}\Big) <$$

$$< \frac{2\sqrt{n}}{n} + \frac{1}{n}\sum_{k=[\sqrt{n}]}^{n-[\sqrt{n}]}\sum_{i=1}^{N^m}\Big(\frac{2}{N^{m-1}}\cdot\frac{2}{N^{m-1}}\Big) < \frac{2\sqrt{n}}{n} + \frac{n-2\sqrt{n}}{n}\cdot\frac{4}{N^{m-2}}.$$

То есть имеем оценку

$$\rho_n(EM(Z_m,N)) < \frac{2}{\sqrt{n}} + \frac{4}{N^{m-2}}.$$

А с учетом оценки (7) получаем

$$\rho_n\Big(\mathcal{PT}(N,m)\Big) < \frac{2e^{2m/3}}{\sqrt{n}} + \frac{4e^{2m/3}}{N^{m-2}}.$$

Теперь

$$\rho\Big(\mathcal{PT}(N,m)\Big) = \lim_{n\to\infty}\rho_n\Big(\mathcal{PT}(N,m)\Big) \le$$

$$\le \lim_{n\to\infty}\Big(\frac{2e^{2m/3}}{\sqrt{n}} + \frac{4e^{2m/3}}{N^{m-2}}\Big) = \frac{4e^{2m/3}}{N^{m-2}}.$$

А с учетом (6) имеем

$$\rho\Big(\mathcal{PT}(N)\Big) < \frac{4e^{2m/3}}{N^{m-2}}$$

для любого натурального m. Так как $N \ge 2$, а $e^{2/3} < 2$, то увеличением m эту верхнюю оценку можно сделать сколь угодно малой. Это означает, что $\rho\Big(\mathcal{PT}(N)\Big) = 0$. □

Напомним, что периодом элемента a полугруппы называется минимальное натуральное число m такое, что $a^k = a^{k+m}$ для некоторого k.

Теорема 2. *Пусть S – конечный моноид такой, что существует элемент $a \in S$ периода больше 1. Тогда проблема проверки тождеств над S генерически разрешима за полиномиальное время.*

Доказательство. Пусть $a \in S$ – элемент периода $N > 1$. Полиномиальный эффективно генерический алгоритм $\mathcal{A}$, решающий проблему проверки тождеств, работает на входе $(p,q) \in X^*$ следующим образом.

(1) Вычисляет $d_i(p,q)$, $i = 1,\dots,|p|+|q|$.
(2) Проверяет, делится ли $d_i(p,q)$ на N для всех $i = 1,\dots,|p|+|q|$.
(3) Если ДА, то выдает ответ «?».
(4) Если НЕТ, то выдает ответ «НЕТ».

Для того, чтобы убедиться, что алгоритм $\mathcal{A}$ корректно решает проблему проверки тождеств в моноиде S, допустим, что $d_k(p,q)$ не делится на N для некоторого k. Вычислим значения термов p и q на наборе элементов $(a_1,\dots,a_n)$, где $a_k = a$ и a – элемент моноида S периода N, и $a_i = 1$ для всех остальных $i \neq k$. Легко видеть, что

$$p(a_1,\dots,a_n) = a^m \neq q(a_1,\dots,a_n) = a^k$$

потому что $m-k$ не делится на N. Таким образом, $p = q$ не является тождеством над S и алгоритм $\mathcal{A}$ выдает правильный ответ «НЕТ». Пренебрежимость множества входов, на которых алгоритм $\mathcal{A}$ выдает ответ «?», следует из леммы 3. □

5. Конечные моноиды Брандта

Полугруппа Брандта B_n есть множество пар натуральных чисел

$$B_n = \{(i,j) \ : \ 1 \leq i,j \leq n\}$$

с нулевым элементом 0. Операция умножения элементов множества B_n определяется следующим образом:

$$(i,j)(k,l) = \begin{cases} (i,l), \text{ если } j = k, \\ 0, \text{ иначе,} \end{cases}$$

для всех $i,j,k,l \leq n$. Также положим $0x = 0$ и $x0 = 0$ для всех $x \in B_n$. *Моноид Брандта* B_n^1 – это полугруппа Брандта B_n с присоединенной единицей 1.

Заметим, что все элементы моноида B_n^1 являются идемпотентами, а потому их периоды равны 1. Таким образом, к моноиду B_n^1 не применим алгоритм из теоремы 2.

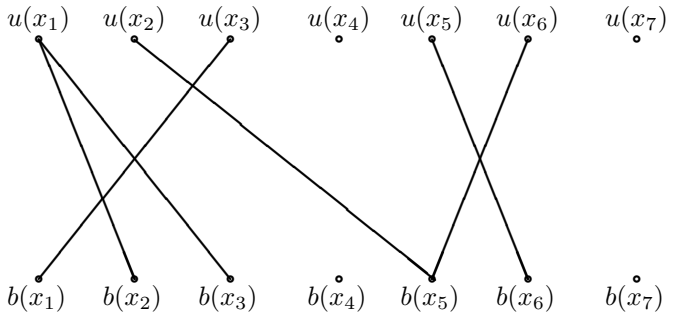
Обозначим через $\mathcal{T}$ множество всех полугрупповых термов, а через $\mathcal{PT}$ множество пар термов.

Важнейшую роль в изучении тождеств в конечных моноидах Брандта играют некоторые специальным образом построенные графы. Следуя [18], для терма $t(x_1,\dots,x_n)$ определим двудольный граф $B(t)$ следующим образом. Граф $B(t)$ имеет n верхних вершин $u(x_1),\dots,u(x_n)$ и n нижних вершин $b(x_1),\dots,b(x_n)$. Каждой переменной в терме t соответствуют две вершины $u(x_i)$ и $b(x_i)$. Вершины $u(x_i)$ и $b(x_j)$ соединены ребром в графе $B(t)$ тогда и только тогда, когда в терме t есть подслово x_ix_j. Обозначим через

$$C(B(t)) = \{C_1, C_2, \dots, C_k\}$$

множество всех связных компонент (заданных множествами вершин) графа $B(t)$. Будем говорить, что связная компонента C_k of $B(t)$ *зависит от переменной* x_i, если в C_k присутствует вершина $u(x_i)$ или вершина $b(x_i)$.

Например, так выглядит граф $B(t)$ для терма $t = x_1x_3x_1x_2x_5x_6x_5$.



В этом графе имеется 8 связных компонент:

(1) $\{u(x_1), b(x_2), b(x_3)\}$,
(2) $\{b(x_1), u(x_3)\}$,
(3) $\{u(x_2), b(x_5), u(x_6)\}$,
(4) $\{u(x_5), b(x_6)\}$,
(5) $\{u(x_4)\}$,
(6) $\{b(x_4)\}$,
(7) $\{u(x_7)\}$,
(8) $\{b(x_7)\}$.

Сейф и Сабо [18] (Предложение 4.12) доказали, что $t_1 = t_2$ является тождеством в полугруппе Брандта B_n тогда и только тогда, когда

(1) $C(B(t_1)) = C(B(t_2))$,
(2) связная компонента $B(t_1)$, содержащая $u(x_i)$, где x_i – это первая буква t_1, совпадает со связной компонентой $B(t_2)$, содержащей $u(x_j)$, где x_j – это первая буква t_2,
(3) связная компонента $B(t_1)$, содержащая $b(x_k)$, где x_k – последняя буква t_1, совпадает со связной компонентой $B(t_2)$, содержащей $b(x_l)$, где x_l – последняя буква t_2.

Заметим, что наборы множеств $C(B(t_1))$ и $C(B(t_2))$ могут быть построены за полиномиальное время, поэтому проблема проверки тождеств над полугруппой Брандта B_n разрешима за полиномиальное время. В тоже время, как было доказано Климой [13] и Сейфом [17], проблема проверки тождеств над моноидом Брандта B_n^1, $n \geq 2$, является co-NP-полной, и не может быть решена за полиномиальное время при условии P $\neq$ NP.

Определим множество пар термов

$$\mathcal{EPT} = \{(t_1, t_2) \in \mathcal{PT} \ : \ C(B(t_1)) = C(B(t_2))\}.$$

Лемма 4. *Множество $\mathcal{EPT}$ пренебрежимо.*

Доказательство. Зафиксируем размер пары термов n. Рассмотрим разбиение

$$\mathcal{EPT}_n = \mathcal{EPT}(1, n-1) \cup \mathcal{EPT}(2, n-2) \cup \ldots \cup \mathcal{EPT}(n-1, 1),$$

где

$$\mathcal{EPT}(l, n-l) = \{(t_1, t_2) \in \mathcal{EPT}_n \ : \ |t_1| = l, |t_2| = n-l\}.$$

Теперь

$$\rho_n\Big(\mathcal{EPT}\Big) = \frac{|\mathcal{EPT}_n|}{|\mathcal{PT}_n|} = \frac{1}{n-2}\sum_{l=1}^{n-1}\frac{|\mathcal{EPT}(l,n-l)|}{n^n}. \tag{8}$$

Обозначим через $\mathcal{T}(n,l)$ множество всех термов размера l от переменных $x_1,\ldots,x_n$. Заметим, что $|\mathcal{T}(n,l)| = n^l$.

Зафиксируем терм $t_1(x_1,\ldots,x_n)$ размера $l<n$. Пусть

$$C(B(t_1)) = \{C_1, C_2, \ldots, C_k\}.$$

Обозначим через $E(t_1)$ множество всех термов t_2 из $\mathcal{T}(n,n-l)$ таких, что $C(B(t_1)) = C(B(t_2))$. Легко видеть, что

$$\mathcal{EPT}(l,n-l) = \{(t_1,t_2)\ :\ t_1\in\mathcal{T}(n,l),\ t_2\in\mathcal{T}(n,n-l),\ t_2\in E(t_1)\}. \tag{9}$$

Предположим, что $i\neq j$ и $i,j\leq n$. Определим отображение φ_{ij} из $\mathcal{T}(n,n-l)$ в $\mathcal{T}(n,n-l)$, которое переставляет переменные x_i и x_j в каждом терме. Пусть компонента C_1 зависит от m переменных, где $1\leq m\leq n-1$. Предположим, что C_1 зависит от переменной x_i и не зависит от переменной x_j. Обозначим через I все такие пары i,j. Нетрудно убедиться, что

(1) $E(t_1)\cap\varphi_{ij}(E(t_1)) = \varnothing$ для всех $i,j\in I$,
(2) более того $\varphi_{ij}(E(t_1))\cap\varphi_{kl}(E(t_1)) = \varnothing$ для различных $i,j\in I$ и $k,l\in I$,
(3) отображение φ_{ij} есть биекция между $E(t_1)$ и $\varphi_{ij}(E(t_1))$, поэтому $|E(t_1)| = |\varphi_{ij}(E(t_1))|$.

Из этого следует

$$E(t_1)\cup\bigcup_{(i,j)\in I}\varphi_{ij}(E(t_1))\subseteq\mathcal{T}(n,n-l)$$

и

$$|\mathcal{T}(n,n-l)|\geq|E(t_1)| + \sum_{(i,j)\in I}|\varphi_{ij}(E(t_1))| = (|I|+1)|E(t_1)|. \tag{10}$$

Заметим, что $|I|\geq\frac{n}{2}$. Действительно, рассмотрим два случая. Если $m\leq\frac{n}{2}$, то можно зафиксировать i такое, что C_1 зависит от x_i и имеется не менее $\frac{n}{2}$ вариантов вабора j так, что C_1 не зависит от x_j. Поэтому $|I|\geq\frac{n}{2}$ в этом случае. Если $m>\frac{n}{2}$, то можно зафиксировать j так, что C_1 не зависит от x_j и снова имеется не менее $\frac{n}{2}$ вариантов выбора i так, что C_1 зависит от x_i. Снова имеем $|I|\geq\frac{n}{2}$.

Теперь оценка (10) влечет

$$|E(t_1)|\leq\frac{|\mathcal{T}(n,n-l)|}{\left(\frac{n}{2}+1\right)}.$$

Это означает, что

$$\frac{|\mathcal{EPT}_{l,n-l}|}{n^n} = \frac{|\mathcal{T}(n,l)|\cdot|E(t_1)|}{|\mathcal{T}(n,l)|\cdot|\mathcal{T}(n,n-l)|}\leq\frac{1}{\frac{n}{2}+1}.$$

Учитывая (8) и (9), получаем

$$\rho_n\Big(\mathcal{EPT}\Big) = \frac{1}{n-2}\sum_{l=1}^{n-1}\frac{|\mathcal{EPT}(l,n-l)|}{n^n}\leq\frac{1}{n-2}\cdot\frac{n-2}{\frac{n}{2}+1} = \frac{1}{\frac{n}{2}+1}.$$

Это значит, что множество $\mathcal{EPT}$ перенебрежимо. □

Теорема 3. *Проблема проверки тождеств в моноиде Брандта B_n^1 генерически разрешима за полиномиальное время.*

Доказательство. Рассмотрим полиномиальный эффективно генерический алгоритм $\mathcal{A}$, который работает на входе $(t_1, t_2) \in \mathcal{PT}$ следующим образом.

(1) Строит графы $B(t_1)$ и $B(t_2)$. Это делается за полиномиальное время.
(2) Если $C(B(t_1)) \neq C(B(t_2))$, то выдает ответ «НЕТ». В этом случае термы t_1 и t_2 не эквивалентны над полугруппой B_n по теореме Сейфа и Сабо [18]. Тем более они не эквивалентны над моноидом B_n^1.
(3) Если $C(B(t_1)) = C(B(t_2))$, то выдает ответ «?». Этот случай пренебрежим по лемме 4.

Пренебрежимость множества входов, на которых алгоритм $\mathcal{A}$ выдает ответ «?», следует из леммы 4. □

References

[1] J. Almeida, M. V. Volkov, S. V. Goldberg. *Complexity of the identity checking problem for finite semigroups*, Journal of Mathematical Sciences, **158:5** (2009), 605–614.

[2] S. Burris, J. Lawrence. *The equivalence problem for finite rings*, Journal of Symbolic Computations, **15:1** (1993), 67–71.

[3] S. Burris, J. Lawrence. *Results on the equivalence problem for finite groups*, Algebra Universalis, **52:4** (2005), 495–500.

[4] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. 1*, Wiley and Sons, (1966) 528 p.

[5] D. Hirschfeldt. *Some Questions in Computable Mathematics*, Computability and Complexity, (2017), 22–55.

[6] G. Horvath, J. Lawrence, L. Merai, C. Szabo. *The complexity of the equivalence problem for nonsolvable groups*, Bulletin of the London Mathematical Society, **39:3** (2007), 433–438.

[7] G. Horvath, C. Szabo. *The complexity of checking identities over finite groups*, International Journal of Algebra and Computation, **16:5** (2006), 931–939.

[8] H. B. Hunt, R. E. Stearns. *The complexity of the equivalence for commutative rings*, Journal of Symbolic Computations, **10:5** (1990), 411–436.

[9] C. Jockusch, P. Schupp. *Generic computability, Turing degrees, and asymptotic density*, Journal of the London Mathematical Society, **85:2** (2012), 472–490.

[10] I. Kapovich, A. Myasnikov, P. Schupp, V. Shpilrain. *Generic-case complexity and decision problems in group theory*, Journal of Algebra, **264** (2003), 665–694.

[11] A. Kisielewicz. *Complexity of semigroup identity checking*, International Journal of Algebra and Computation, **14:4** (2004), 455–464.

[12] O. G. Kharlampovich, M. V. Sapir. *Algorithmic problems in varieties*, International Journal of Algebra and Computation, **5:4-5** (1995), 379–602.

[13] O. Klima. *Complexity issues of checking identities in finite monoids*, Semigroup Forum, **79** (2009), 435–442.

[14] A. Meyer. *An open problem on creative sets*, Recursive Function Theory Newsletter, **4** (1973), 15–16.

[15] S. V. Pleshcheva, V. Verteshi. *Complexity of the identity checking in one finite 0-simple semigroup*, Izvestiya Uralskogo gosudarstvennogo universiteta, **43** (2006), 72–102. (in Russian)

[16] L. N. Shevrin, M. V. Volkov. *Identities of semigroups*, Soviet Math. (Iz. VUZ), **29:11** (1985), 1–64.

[17] S. Seif. *The Perkins semigroup has co-NP-complete term-equivalence problem*, International Journal of Algebra and Computation, **15:2** (2005), 317–326.

[18] S. Seif, C. Szabo. *Computational complexity of checking identities in 0-simple semigroups and matrix semigroups over finite fields*, Semigroup Forum, **72:2** (2006), 207–222.

[19] C. Szabo, V. Vertesi. *The complexity of checking identities for finite matrix rings*, Algebra Universalis, **51:4** (2004), 439–445.

Alexander Nikolaevich Rybalov
Sobolev Institute of Mathematics,
prospekt Koptyuga 4,
Novosibirsk, 630090, Russia.
Email address: alexander.rybalov@gmail.com