ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Октябрь—декабрь 2005. Серия 1. Том 12, № 4, 40–50

УДК 519.714

ОБ АДДИТИВНОЙ СЛОЖНОСТИ ЧАСТИЧНО КОММУТАТИВНЫХ СЛОВ*)

Ю. В. Мерекин

При получении нижних оценок аддитивной сложности частично коммутативных слов предлагается использовать модифицированный суффиксный метод, введённый автором для оценки сложности слов, порождаемых схемой конкатенации. Для частично коммутативной обобщённой последовательности Туэ–Морса получена оценка аддитивной сложности.

Введение

Пусть $\Sigma = \{a_1, \ldots, a_q\}$, $q \geqslant 3$, — алфавит, W — слово в алфавите Σ. Пусть символы алфавита упорядочены: $a_1 < \ldots < a_q$. Для алфавита Σ укажем некоторое множество G пар различных букв. Положим по определению, что для любого слова W при коммутации двух соседних букв, указанных в G, получается слово, равное слову W. Если множество G пусто, то слова $W_1 = a_{i_1} \ldots a_{i_k}$ и $W_2 = a_{j_1} \ldots a_{j_k}$ равны только тогда, когда $a_{i_t} = a_{j_t}, 1 \leqslant t \leqslant k$. Если в G содержатся все пары букв, то в любом слове W любая перестановка букв даёт слово, равное исходному, т. е. все слова, имеющие одинаковый буквенный состав, равны между собой. Такие слова называем *коммутативными* и обозначаем через $\hat{W}, \hat{U}, \ldots$ Далее в процессе построения слов, имея коммутативное слово, будем считать, что имеются все равные ему коммутативные слова.

В случае, когда в G содержатся не все пары букв, слово называем *частично коммутативным* и обозначаем через $\hat{W}(G), \hat{U}(G), \ldots$ Как и в случае коммутативных слов, получив частично коммутативное слово, будем считать, что имеются все равные ему частично коммутативные слова. В алгебре частично коммутативные слова возникают при изучении полугрупп и групп, а G состоит из определяющих соотношений (см., например, [3, 4]).

Следуя определениям из [1], определим понятие схемы конкатенации и сложность частично коммутативных слов.

---

*)Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 05–01–00364).

Последовательность частично коммутативных слов $a_1, \ldots, a_q, \hat{X}(G)$, $\hat{Y}(G)$, $\ldots, \hat{Z}(G)$, которую обозначим через $\hat{S}(G)$, назовём *схемой конкатенации* частично коммутативного слова $\hat{Z}(G)$, если для любого слова $\hat{W}(G)$ из $\hat{S}(G)$, начиная с $\hat{X}(G)$, в $\hat{S}(G)$ имеются такие слова $\hat{U}(G)$ и $\hat{V}(G)$ (возможно, $\hat{U}(G) = \hat{V}(G)$), предшествующие слову $\hat{W}(G)$, что $\hat{W}(G) = \hat{U}(G)\hat{V}(G)$.

Под *сложностью* $L(\hat{S}(G))$ схемы $\hat{S}(G)$ конкатенации частично коммутативного слова $\hat{Z}(G)$ понимается число слов в последовательности $\hat{X}(G)$, $\hat{Y}(G), \ldots, \hat{Z}(G)$. Пусть $L(\hat{Z}(G)) = \min L(\hat{S}(G))$, где минимум берётся по всем схемам конкатенации слова $\hat{Z}(G)$. Величину $L(\hat{Z}(G))$ назовём *аддитивной сложностью* частично коммутативного слова $\hat{Z}(G)$. В дальнейшем изложении под сложностью слов везде понимается аддитивная сложность. Схему $\hat{S}(G))$ назовём *оптимальной*, если $L(\hat{Z}(G)) = L(\hat{S}(G))$.

§ 1. Нижняя оценка сложности частично коммутативных слов

Рассмотрим сложность частично коммутативных слов. Для получения нижних оценок их сложности модифицируем предложенный автором в [1] суффиксный метод. В [2] метод распространён на слова, порождаемые схемами композиции, и получена оценка их сложности. Ниже аналогичные доказательства распространены на суффиксный метод получения нижних оценок сложности частично коммутативных слов. Суффиксный метод использует следующее представление слов.

Длиной $|\hat{W}(G)|$ частично коммутативного слова $\hat{W}(G)$ назовём число символов в $\hat{W}(G)$. Частично коммутативное слово $\hat{X}(G)$ назовём *подсловом* частично коммутативного слова $\hat{W}(G)$ и обозначим через $\hat{X}(G) \sqsubseteq \hat{W}(G)$, если при некоторых (возможно, пустых) $\hat{U}(G)$ и $\hat{V}(G)$ справедливо равенство $\hat{W}(G) = \hat{U}(G) \circ \hat{X}(G) \circ \hat{V}(G)$. При пустом $\hat{U}(G)$ подслово $\hat{X}(G)$ назовём *префиксом*, а при пустом $\hat{V}(G)$ — *суффиксом* частично коммутативного слова $\hat{W}(G)$.

Из результатов работы [4] о делимости частично коммутативных слов вытекает следующее утверждение.

Лемма 1. *Пусть $\hat{Y}_1(G), \ldots, \hat{Y}_m(G)$, $m \geqslant 2$, — некоторые суффиксы частично коммутативного слова $\hat{W}(G)$. Тогда существует единственный суффикс $\hat{X}(G)$ минимальной длины, в котором в качестве подслов содержатся все суффиксы $\hat{Y}_1(G), \ldots, \hat{Y}_m(G)$.*

Представление частично коммутативного слова $\hat{W}(G)$ вида $\hat{W}(G) = \hat{U}(G) \circ \hat{Y}(G) \circ \hat{V}(G) \circ \hat{Y}(G)$, где $\hat{U}(G)$ и $\hat{V}(G)$, возможно, пусты, назовём *представлением, копирующим суффикс* $\hat{Y}(G)$.

Пусть для частично коммутативного слова $\hat{W}(G)$ существует представление, копирующее суффикс $\hat{Y}(G)$, и не существует представления, которое копирует любой суффикс $\hat{Z}(G)$ такой, что $\hat{Z}(G) \sqsupset \hat{Y}(G)$. Тогда суффикс $\hat{Y}(G)$ назовём *тупиковым суффиксом* частично коммутативного слова $\hat{W}(G)$. Пусть для частично коммутативного слова $\hat{W}(G)$ имеется несколько тупиковых суффиксов $\hat{Y}_1(G), \ldots, \hat{Y}_m(G)$, $m \geqslant 2$. Тогда имеется такой суффикс $\hat{X}(G)$ минимальной длины, что $\hat{X}(G) \sqsupseteq \hat{Y}_1(G), \ldots, \hat{X}(G) \sqsupseteq \hat{Y}_m(G)$. Суффикс $\hat{X}(G)$ назовём *объединённым суффиксом* частично коммутативного слова $\hat{W}(G)$. Например, для $\Sigma = \{1, 2, 3\}$, $G = \{(1,3)\}$ и частично коммутативного слова $\hat{W}(G) = 1223231$ представления $1 \circ 22323 \circ 1$ и $122 \circ 3 \circ 21 \circ 3$ копируют тупиковые суффиксы $\hat{Y}_1(G) = 1$ и $\hat{Y}_2(G) = 3$, а объединённым суффиксом является $\hat{X}(G) = 31$.

Из леммы 1 следует, что объединённый суффикс всегда единственный.

Если суффикс слова $\hat{W}(G)$ является символом и входит в $\hat{W}(G)$ один раз, то назовём его *символьным суффиксом*. Частично коммутативное слово может иметь несколько символьных суффиксов. В этом случае они попарно коммутируют (лемма 3.15 из [4]).

Представление частично коммутативного слова $\hat{W}(G)$ в виде

$$\hat{W}(G) = \hat{X}_r(G) \circ \ldots \circ \hat{X}_1(G) \tag{1}$$

назовём *суффиксным представлением*, если каждое слово $\hat{X}_i(G)$, $1 \leqslant i \leqslant r$, является либо символьным, либо объединённым, либо тупиковым суффиксом частично коммутативного слова $\hat{X}_r(G) \circ \ldots \circ \hat{X}_i(G)$. Частично коммутативное слово может иметь более одного типа суффиксов. В этом случае при построении суффиксного представления установим приоритетность типов суффиксов: наибольший символьный суффикс (если символьный суффикс не единственный), объединённый суффикс, тупиковый суффикс. В результате получаем единственное суффиксное представление слова $\hat{W}(G)$. Например, для $\Sigma = \{1, 2, 3\}$, $G = \{(1,3)\}$ и частично коммутативного слова $\hat{W}(G) = 1223233121$ имеем суффиксное представление $\hat{W}(G) = 1 \circ 2 \circ 2 \circ 3 \circ 2 \circ 31 \circ 321$.

В суффиксном представлении (1) каждый суффикс $\hat{X}_i(G)$, $1 \leqslant i \leqslant r$, назовём *максимальным*. Заметим, что при пустом G получаем суффиксное представление слова W, ранее введённое в [1].

Число операций конкатенации в суффиксном представлении назовём *сложностью суффиксного представления* и обозначим через $L^*(\hat{W}(G))$.

Пусть слово $\hat{X}(G)$ является максимальным суффиксом слова $\hat{W}(G)$. Любой суффикс $\hat{Z}(G)$, $\hat{Z}(G) \sqsupset \hat{X}(G)$, назовём *расширенным суффиксом* частично коммутативного слова $\hat{W}(G)$. Например, для $\Sigma = \{1, 2, 3\}$,

$G = \{(1,3)\}$ и частично коммутативного слова $\hat{W}(G) = 1223231$ суффикс $\hat{X}(G) = 31$ максимальный, а суффиксы $231, 3231, 23231$, 223231 и 1223231 расширенные.

Представление частично коммутативного слова $\hat{W}(G)$ в виде

$$\hat{W}(G) = \hat{Y}_h(G) \circ \ldots \circ \hat{Y}_1(G)$$

назовём *расширенным суффиксным представлением*, если

(i) каждый суффикс $\hat{Y}_h(G), \ldots, \hat{Y}_1(G)$ является либо максимальным, либо расширенным;

(ii) среди $\hat{Y}_h(G), \ldots, \hat{Y}_1(G)$ имеется хотя бы один расширенный суффикс.

Например, для $\Sigma = \{1,2,3\}$, $G = \{(1,3)\}$ и частично коммутативного слова $\hat{W}(G) = 1223233121$ представление $\hat{W}(G) = 12 \circ 2 \circ 3 \circ 231 \circ 321$ (суффиксы 12 и 231 расширенные) является расширенным суффиксным представлением.

Расширенное суффиксное представление частично коммутативного слова, вообще говоря, не единственно. Число операций конкатенации в расширенном представлении назовём *сложностью расширенного суффиксного представления* и обозначим через $L^{**}(\hat{W}(G))$. Справедливы следующие факты.

Утверждение 1. *Для любых частично коммутативных слов $\hat{U}(G)$ и $\hat{V}(G)$ справедливо неравенство $L^*(\hat{U}(G)\hat{V}(G)) \geqslant L^*(\hat{U}(G))$.*

Утверждение 2. *Для произвольного частично коммутативного слова $\hat{W}(G)$ и любого его расширенного суффиксного представления справедливо неравенство $L^*(\hat{W}(G)) \geqslant L^{**}(\hat{W}(G))$.*

Убедимся в справедливости следующего факта.

Лемма 2. *Для любого частично коммутативного слова $\hat{W}(G)$ существует оптимальная схема конкатенации, содержащая такие частично коммутативные слова $\hat{U}(G)$, $\hat{V}(G)$, что $\hat{W}(G) = \hat{U}(G)\hat{V}(G)$ и $\hat{V}(G)$ является либо подсловом слова $\hat{U}(G)$, либо символьным суффиксом.*

Доказательство. Пусть $\hat{S}(G)$ — некоторая оптимальная схема конкатенации частично коммутативного слова $\hat{W}(G)$, содержащая такие частично коммутативные слова $\hat{U}(G)$ и $\hat{V}(G)$, что $\hat{W}(G) = \hat{U}(G)\hat{V}(G)$. Возможно, что

$$\hat{W}(G) = \hat{U}(G)\hat{V}(G) = \hat{V}(G)\hat{U}(G). \tag{2}$$

Если $\hat{V}(G)$ — символ, отсутствующий в $\hat{U}(G)$, или в случае равенств (2) $\hat{U}(G)$ — символ, отсутствующий в $\hat{V}(G)$, то лемма справедлива. Если в случае равенств (2) $\hat{U}(G)$ является подсловом слова $\hat{V}(G)$, то лемма справедлива. Поэтому остаётся рассмотреть случай, когда $\hat{V}(G)$ не является подсловом слова $\hat{U}(G)$. В этом случае подсхема $\hat{S}_1(G)$, $\hat{S}_1(G) \subset \hat{S}(G)$, конкатенации частично коммутативного слова $\hat{U}(G)$ не содержит $\hat{V}(G)$, а схема $\hat{S}(G)$ содержит такие $\hat{V}_1(G)$ и $\hat{V}_2(G)$, что $\hat{V}(G) = \hat{V}_1(G)\hat{V}_2(G)$ (возможно, что $\hat{V}(G) = \hat{V}_1(G)\hat{V}_2(G) = \hat{V}_2(G)\hat{V}_1(G)$). В схеме $\hat{S}(G)$ частично коммутативное слово $\hat{V}(G)$ заменим на $\hat{V}_3(G) = \hat{U}(G)\hat{V}_1(G)$ (возможно, что $\hat{V}_4(G) = \hat{U}(G)\hat{V}_2(G)$ в случае, когда $\hat{V}_1(G)\hat{V}_2(G) = \hat{V}_2(G)\hat{V}_1(G)$). В результате получим схему $\hat{S}_2(G)$ конкатенации частично коммутативного слова $\hat{W}(G)$, содержащую такие $\hat{V}_3(G)$ и $\hat{V}_2(G)$, что

$$\hat{W}(G) = \hat{V}_3(G)\hat{V}_2(G). \tag{3}$$

Возможно, что если $\hat{V}_1(G)\hat{V}_2(G) = \hat{V}_2(G)\hat{V}_1(G)$, то

$$\hat{W}(G) = \hat{V}_4(G)\hat{V}_1(G). \tag{4}$$

Суффиксы $\hat{V}_2(G)$ из (3) и $\hat{V}_1(G)$ из (4) короче суффикса $\hat{V}(G)$, а схема $\hat{S}_2(G)$ также оптимальна, поскольку число слов в ней равно числу слов в схеме $\hat{S}(G)$.

Если суффикс $\hat{V}_2(G)$ (или $\hat{V}_1(G)$) частично коммутативного слова $\hat{W}(G)$ не удовлетворяет утверждению леммы, то описанная выше процедура выделения более коротких суффиксов повторяется до появления суффикса, который удовлетворяет утверждению леммы. Если в слове $\hat{W}(G)$ существует более одного символьного суффикса, то на последнем шаге описанной процедуры в качестве символьного суффикса можно выбрать любой из них. Лемма 2 доказана.

Теорема 1. *Для любого частично коммутативного слова $\hat{W}(G)$ справедливы неравенства $L(\hat{W}(G)) \geqslant L^*(\hat{W}(G)) \geqslant L^{**}(\hat{W}(G))$.*

Доказательство. Неравенство $L^*(\hat{W}(G)) \geqslant L^{**}(\hat{W}(G))$ содержится в утверждении 2. Докажем справедливость неравенства

$$L(\hat{W}(G)) \geqslant L^*(\hat{W}(G)).$$

Предположим, что теорема неверна. Из множества частично коммутативных слов, для которых теорема неверна, выберем слово с минимальной сложностью. Пусть таким словом является $\hat{W}(G)$ и

$$L(\hat{W}(G)) < L^*(\hat{W}(G)). \tag{5}$$

Согласно лемме 2 среди оптимальных схем конкатенации частично коммутативного слова $\hat{W}(G)$ имеется схема, содержащая такие $\hat{U}(G)$,

$\hat{V}(G)$ и $\hat{W}(G)$, что $\hat{W}(G) = \hat{U}(G)\hat{V}(G)$ и $\hat{V}(G)$ является либо подсловом слова $\hat{U}(G)$, либо символом, отсутствующим в $\hat{U}(G)$. Пусть $\hat{W}(G) = \hat{Y}(G)\hat{X}(G)$, где $\hat{X}(G)$ — максимальный суффикс частично коммутативного слова $\hat{W}(G)$. Установим ряд соотношений. Очевидно, что

$$L(\hat{W}(G)) \geqslant L(\hat{U}(G)) + 1. \tag{6}$$

Из минимальности сложности слова $\hat{W}(G)$ при предположении (5) имеем

$$L(\hat{U}(G)) \geqslant L^*(\hat{U}(G)). \tag{7}$$

Если $\hat{V}(G)$ — символ, отсутствующий в $\hat{U}(G)$, то по определению слово $\hat{V}(G)$ является максимальным суффиксом частично коммутативного слова $\hat{W}(G)$. Следовательно,

$$\hat{U}(G) = \hat{Y}(G), \quad L(\hat{W}(G)) = L(\hat{U}(G)) + 1, \quad L^*(\hat{W}(G)) = L^*(\hat{U}(G)) + 1$$

и согласно (5) $L(\hat{U}(G)) < L^*(\hat{U}(G))$, что противоречит предположению о минимальной сложности слова $\hat{W}(G)$.

Если $\hat{V}(G) \sqsubseteq \hat{U}(G)$, то $\hat{V}(G) \sqsubseteq \hat{X}(G)$, $\hat{Y}(G) \sqsubseteq \hat{U}(G)$ и согласно утверждению 1

$$L^*(\hat{U}(G)) \geqslant L^*(\hat{Y}(G)). \tag{8}$$

По определению суффиксного представления частично коммутативного слова $\hat{W}(G)$ имеем

$$L^*(\hat{Y}(G)) = L^*(\hat{W}(G)) - 1. \tag{9}$$

Объединяя (6)–(9), получаем

$$L(\hat{W}(G)) \geqslant L(\hat{U}(G)) + 1 \geqslant L^*(\hat{U}(G)) + 1 \geqslant L^*(\hat{Y}(G)) + 1 = L^*(\hat{W}(G)),$$

что противоречит принятому предположению. Теорема 1 доказана.

Из теоремы 1 непосредственно вытекает

Следствие 1. *Пусть* $\hat{W}(G) = \hat{U}(G) \circ \hat{X}_l(G) \circ \ldots \circ \hat{X}_1(G)$, *где* $\hat{X}_l(G)$, $\ldots$, $\hat{X}_1(G)$ — *максимальные (расширенные) суффиксы. Тогда* $L(\hat{W}(G)) \geqslant L(\hat{U}(G)) + l$.

§ 2. Пример

Рассмотрим обобщённую последовательность Туэ–Морса [5] в алфавите $\Sigma = \{1, 2, 3\}$, которая определяется реккурентной схемой

$$(a) \quad \varphi^0(i) = \quad i \quad \text{п} \quad i = 1, 2, 3;$$

$$(b) \quad \varphi^n(i) = \begin{cases} \varphi^{n-1}(i)\varphi^{n-1}(i+1) & \text{п} \quad i = 1, 2 \quad \text{и} \quad n \geqslant 1; \\ \varphi^{n-1}(i)\varphi^{n-1}(i-2) & \text{п} \quad i = 3 \quad \text{и} \quad n \geqslant 1. \end{cases}$$

Тогда $\varphi^2(1) = 1223$, $\varphi^3(1) = 12232331$, $\varphi^4(1) = 1223233123313112$.

Слова

$$B_1 = \varphi^2(1) = 1223, \quad B_2 = \varphi^2(2) = 2331, \quad B_3 = \varphi^2(3) = 3112$$

назовём *блоками*. Нетрудно видеть, что любое слово $\hat{\varphi}(\cdot)(G)$ представимо конкатенацией блоков $B_i, i = 1, 2, 3$.

Если в множестве G нет пары, содержащей символ x, $x \in \Sigma$, то в любом частично коммутативном слове $\hat{W}(G)$ все вхождения символа x *неподвижны*. Поэтому в слове $\hat{\varphi}^n(\cdot)(G)$ неподвижным является символ 2.

Для нахождения аддитивной сложности частично коммутативного слова $\hat{\varphi}^n(1)(G)$ нам потребуются вспомогательные утверждения (леммы 3–6), из которых первые три относятся к обобщённой последовательности Туэ–Морса, не являющейся частично коммутативной.

Лемма 3. *Пусть $\varphi^n(1)$ — обобщённая последовательность Туэ–Морса в алфавите $\Sigma = \{1, 2, 3\}$. Тогда при любом $n \geqslant 3$ последовательность $\varphi^n(1)$ имеет суффиксное представление вида*

$$\varphi^n(1) = 1{\circ}2{\circ}2{\circ}3{\circ}\varphi^1(2){\circ}\varphi^0(3){\circ}\varphi^0(1){\circ}\ldots{\circ}\varphi^{n-2}(2){\circ}\varphi^{n-3}(3){\circ}\varphi^{n-3}(1). \quad (10)$$

Доказательство. Из определения обобщённой последовательности Туэ–Морса следуют равенства

$$\varphi^n(1) = \varphi^{n-1}(1) \circ \varphi^{n-1}(2) = \varphi^{n-1}(1) \circ \varphi^{n-2}(2) \circ \varphi^{n-2}(3)$$
$$= \varphi^{n-1}(1) \circ \varphi^{n-2}(2) \circ \varphi^{n-3}(3) \circ \varphi^{n-3}(1).$$

Повторив аналогичные преобразования для префиксов $\varphi^{n-1}(1), \ldots, \varphi^1(1)$, получаем представление (10). Докажем, что оно суффиксное.

Предположим, что в (10) суффикс $\varphi^i(\cdot)$, $1 \leqslant i \leqslant n-2$, не является максимальным. Сделав замены $12 \to 1$, $23 \to 2, 31 \to 3$, получим вдвое более короткий суффикс, для которого также справедливо это предположение. Продолжив «спуск», придём к суффиксу единичной длины, что противоречит предположению. Следовательно, представление (10) слова $\varphi^n(1)$ суффиксное. Лемма 3 доказана.

Лемма 4. *Пусть представление $\varphi^n(1) = X_r \circ \ldots \circ X_1$ слова $\varphi^n(1)$, $n \geqslant 6$, суффиксное. Тогда для каждого максимального суффикса $X_i = \varphi^k(\cdot)$, $k \geqslant 3$, в любом представлении вида*

$$X_r \ldots X_i = U X_i V X_i, \quad 1 \leqslant i < r,$$

слова U и V являются блочными (возможно, одно из них пустое).

Доказательство. Суффиксное представление (10) слова $W = \varphi^n(1)$, $n \geqslant 6$, преобразуем к виду

$$\begin{aligned}\varphi^n(1) &= \varphi^5(1) \circ X_j \circ \ldots \circ X_1\\ &= \varphi^5(1) \circ \varphi^4(2) \circ \varphi^3(3) \circ \varphi^3(1) \circ \ldots \circ \varphi^{n-2}(2) \circ \varphi^{n-3}(3) \circ \varphi^{n-3}(1), \quad (11)\end{aligned}$$

где $j = 3(n-5)$. Для выделенных максимальных суффиксов $X_j, \ldots, X_1$ согласно определению максимального суффикса имеем

$$\varphi^5(1)X_j \ldots X_i = UX_iVX_i, \quad j \geqslant i \geqslant 1, \qquad (12)$$

где слова U и V, возможно, пустые. К обеим частям равенства (12) припишем справа блочное слово $A = X_{i-1} \ldots X_1$. В результате получим равенство $\varphi^n(1) = UX_iVX_iA$. Покажем, что слова U и VX_iA блочные.

Любое блочное слово $\varphi^k(\cdot)$, $k > 3$, имеет представление вида $\varphi^k(\cdot) = \varphi^3(\cdot)W$, где слово W блочное. Поэтому достаточно доказать лемму при $k = 3$. В этом случае существуют три представления

$$\varphi^3(1) = 12232331, \ \varphi^3(2) = 23313112, \ \varphi^3(3) = 31121223.$$

В каждом представлении имеется ровно два блока. Один блок является префиксом, а другой — суффиксом. Следовательно, любое 2-блочное слово $\varphi^3(\cdot)$, входящее в блочное слово $\varphi^n(\cdot)$, $n > k$, не может иметь непустое пересечение с каждым из трёх соседних блоков слова $\varphi^n(\cdot)$. Поэтому слова U и VX_iA блочные, а слова X_i и A блочные по условию. Следовательно, слово V блочное. Лемма 4 доказана.

В представлении $X = Ys$ блочного слова X суффикс s назовём *простым*, если символ 2 является суффиксом слова Y и не содержится в s.

Лемма 5. *Пусть слово $\varphi^k(\cdot)$, $k \geqslant 3$, задано в виде $\varphi^k(\cdot) = Ys$, где суффикс s простой. Тогда если слово Y входит в слово $\varphi^n(\cdot)$, $n \geqslant k$, то непосредственно за Y следует слово s.*

Доказательство. Для слова $\varphi^k(\cdot)$, $k \geqslant 3$, существуют три вида простых суффиксов:

$$\varphi^k(\cdot) = \left\{\begin{array}{llll} W_1\varphi^3(1) = & W_1 12232 \circ 331 & = & Y_1 \circ s_1, \quad |s_1| = 3,\\ W_2\varphi^3(2) = & W_2 21133112 & = & Y_2 \circ s_2, \quad |s_2| = 0,\\ W_3\varphi^3(3) = & W_3 3112122 \circ 3 & = & Y_3 \circ s_3, \quad |s_3| = 1.\end{array}\right.$$

Рассмотрим третий случай. В суффиксе 3112122 слова $Y_3 = W_3 3112122$ имеется только один блок 3112. Поэтому Y_3 входит только в (2^{k-2})-блочное слово $W_3 3112122 \circ 3$ и не входит в $(2^{k-2}+1)$-блочное слово вида $A_1 W_3 3112122 A_2$, где $1 \leqslant |A_1| \leqslant 3$, так как

(1) при $|A_1| = 1$ подслово 2122 слова $A_1W_33112122A_2$ не является блоком,

(2) при $|A_1| = 2$ подслово 1212 слова $A_1W_33112122A_2$ не является блоком,

(3) при $|A_1| = 3$ подслово 1121 слова $A_1W_33112122A_2$ не является блоком.

Следовательно, при $|s_3| = 1$ лемма доказана.

Для других простых суффиксов доказательство аналогично. Лемма 5 доказана.

Лемма 6. *Пусть в суффиксном представлении* $\varphi^n(1) = X_r \circ \ldots \circ X_1$ *слова* $\varphi^n(1)$, $n \geqslant 6$, *выделены все такие максимальные суффиксы* $X_j, \ldots, X_1$, *что* X_v, $1 \leqslant v \leqslant j$, *имеет вид* $X_v = \varphi^k(\cdot) = Y_v s_v$, *где* s_v — *простой суффикс. Тогда*

(1) *в каждом слове* $X_r \ldots X_{i+1}Y_i$, $1 \leqslant v \leqslant j$, *суффикс* Y_i *максимальный;*

(2) *в каждом частично коммутативном слове*

$$\hat{X}_r(G) \ldots \hat{X}_{i+2}(G)\hat{Y}_{i+1}(G)\hat{s}_{i+1}\hat{Y}_i(G), \quad 1 \leqslant i \leqslant j,$$

суффикс $\hat{s}_{i+1}\hat{Y}_i(G)$ *либо максимальный, либо расширенный.*

Доказательство. Согласно лемме 4 для максимального суффикса $X_i = Y_i s_i$ в любом представлении вида

$$X_r \ldots X_{i+1} \circ Y_i s_i = U \circ Y_i s_i \circ V \circ Y_i s_i, \quad j \geqslant i \geqslant 1, \tag{13}$$

слова U и V блочные.

Из леммы 5 следует, что слово Y_i входит в слово $X_r \ldots X_i$ только совместно со словом s_i. Поэтому в представлениях вида (13) показаны все слова Y_i, входящие в слово $X_r \ldots X_i$. Удалив из обеих частей равенств (13) простой суффикс s_i, получаем равенство

$$X_r \ldots X_{i+1} \circ Y_i = U \circ Y_i s_i \circ V \circ Y_i.$$

Следовательно, в слове $X_r \ldots X_{i+1}Y_i$ суффикс Y_i максимальный. Первое утверждение леммы 6 доказано.

Выделив в блочных словах U и V из (13) блоки-суффиксы $U = U_1 B_t$ и $V = V_1 B_u$, а в блоках B_t и B_u простые суффиксы $B_t = Y_t s_t$ и $B_u = Y_u s_u$, получаем

$$X_r \ldots X_{i+1}Y_i = U_1 Y_t s_t Y_i s_i V_1 Y_u s_u Y_i. \tag{14}$$

Из определения максимального суффикса имеем $B_t \neq B_u$ Следовательно, $s_t \neq s_u$. Рассматривая слова из (14) как частично коммутативные, получаем представление

$$\hat{X}_r(G)\ldots\hat{X}_{i+1}(G)\hat{Y}_i(G) = \hat{U}_1(G)\hat{Y}_t(G)\hat{s}_t\hat{Y}_i(G)\hat{s}_i\hat{V}_1(G)\hat{Y}_u(G)\hat{s}_u\hat{Y}_i(G),$$

в котором длина частично коммутативного максимального суффикса не превосходит $|\hat{s}_u\hat{Y}_i(G)|$, так как суффиксами частично коммутативных слов $\hat{Y}_t(G)$ и $\hat{Y}_u(G)$ являются неподвижные символы. Следовательно, суффикс $\hat{s}_u\hat{Y}_i(G)$ либо максимальный, либо расширенный. Лемма 6 доказана.

Теорема 2. *Пусть* $\Sigma = \{1,2,3\}$ *и* $G = \{(1,3)\}$. *Тогда* $L(\hat{\varphi}^n(1)(G)) = 3n-3$ *при любом* $n \geqslant 6$.

Доказательство. Верхняя оценка следует из схемы

$$\begin{array}{llll}
\hat{S}(G) = 1,2,3, & \hat{\varphi}^1(1)(G), & \hat{\varphi}^1(2)(G), & \hat{\varphi}^1(3)(G), \\
 & \vdots & \vdots & \vdots, \\
 & \hat{\varphi}^{n-2}(1)(G), & \hat{\varphi}^{n-2}(2)(G), & \hat{\varphi}^{n-2}(3)(G), \\
 & \hat{\varphi}^{n-1}(1)(G), & \hat{\varphi}^{n-1}(2)(G), & \\
 & \hat{\varphi}^{n}(1)(G), & &
\end{array}$$

сложность которой равна $3n-3$.

Нижняя оценка получается в результате построения расширенного суффиксного представления частично коммутативного слова $\hat{\varphi}^n(1)(G)$. В представлении (11) выделены максимальные суффиксы вида $\varphi^k(\cdot)$, $k \geqslant 3$. Зададим слово $\hat{\varphi}^5(1)(G)$ в виде $\hat{\varphi}^5(1)(G) = \hat{W}(G)\hat{s}_{3(n-5)+1}$. К каждому максимальному суффиксу из (11) применим лемму 6. В результате получаем представление

$$\hat{\varphi}^n(1)(G) = \hat{W}(G) \circ \hat{s}_{3(n-5)+1}\hat{Y}_{3(n-5)}(G) \circ \ldots \circ \hat{s}_3\hat{Y}_2(G) \circ \hat{s}_2\hat{Y}_1(G)\hat{s}_1, \quad (15)$$

в котором каждое слово $\hat{s}_v\hat{Y}_{v-1}(G)$, $3 \leqslant v \leqslant 3(n-5)+1$, а также слово $\hat{s}_2\hat{Y}_1\hat{s}_1$, является либо максимальным, либо расширенным суффиксом. В слове $\hat{W}(G)$ выделим пять максимальных суффиксов:

$$\hat{W}(G) = 1223233121 \circ 31 \circ 3312 \circ 23313112 \circ 3112 \circ 122. \quad (16)$$

Для частично коммутативного слова 1223233121 из (16) существует оптимальная схема

$$\hat{S}_1(G) = 1,2,3,12,23,1223,122323,312,3121,1223233121, \quad (17)$$

сложность которой равна 7. Суммируя число максимальных суффиксов, выделенных в (15) и (16), с учётом сложности схемы (17) по следствию 1 имеем $L(\hat{\varphi}^n(1)(G)) \geqslant 3n - 3$. Теорема 2 доказана.

Автор благодарит Л. А. Бокутя за полезные обсуждения.

ЛИТЕРАТУРА

1. **Мерекин Ю. В.** Нижняя оценка сложности для схем конкатенации слов // Дискрет. анализ и исслед. операций. 1996. Т. 3, № 1. С. 52–56.
2. **Мерекин Ю. В.** О порождении слов с использованием операции композиции // Дискрет. анализ и исслед. операций. Сер. 1. 2003. Т. 10, № 4. С. 70–78.
3. **Bokut L. A., Shiao L.-S.** Gröbner-Shirshov bases for Coxeter groups // Commun. in Algebra. 2001. V. 29, N 9. P. 4305–4319.
4. **Esyp E. S., Kazatchkov I. V., Remeslennikov V. N.** Divisibility theory and complexity of algorithms for free partially commutative groups // Contemp. Math. 2005. V. 378. P. 317–346.
5. **Shallit J., Tromp J.** Subword complexity of generalized Thue–Morse word // Inform. Process. Lett. 1995. V. 54, N 6. P. 313–316.

Адрес автора:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
630090 Новосибирск,
Россия.
E-mail: merekin@math.nsc.ru

Статья поступила
21 марта 2005 г.