ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Март—апрель 2009. Том 16, № 2. С. 3–15

УДК 519.114

АРИФМЕТИЧЕСКОЕ ЗАМЫКАНИЕ ДВУМЕРНЫХ СЛОВ ТЁПЛИЦА

Ц. Ч.-Д. Батуева

Аннотация. Арифметическое замыкание слова — это множество всех подслов его арифметических подпоследовательностей. В работе исследуются арифметические замыкания двумерных слов Тёплица. Найдена точная формула арифметической сложности для некоторого класса слов Тёплица.

Ключевые слова: сложность, арифметическая сложность, двумерные слова, слова Тёплица.

Введение

Классическая комбинаторная, или подсловная сложность бесконечного слова, — это число его подслов длины n. Различные её свойства изучены в работе [6]. Также были введены различные её модификации, основанные на подсчёте слов в бесконечных словах: например, палиндромная сложность [1], модифицированная сложность Накасимы и др. [10], шаблонная сложность, введённая Рестиво и Салеми [11], максимальная шаблонная сложность Камаэ и Замбони [8] и др.

Арифметическая сложность бесконечного слова, введённая в 2000 г. в работе [3], также является модификацией классической сложности. Она подсчитывает совокупное число всех подслов в арифметических подпоследовательностях слова. Очевидно, что арифметическая сложность ограничена снизу комбинаторной сложностью, а сверху — мощностью алфавита в степени n. Когда комбинаторная сложность растёт линейно, тогда арифметическая сложность может быть как экспоненциальной [3, 7], так и линейной [2, 3].

Слова Тёплица, их комбинаторная сложность и периодичность исследованы, например, в [5, 9]. Двумерные слова, например, слова с низкой комбинаторной сложностью, рассматривались в работе [4].

В данной работе вводятся понятия арифметического замыкания и арифметической сложности для двумерных слов. Найдено арифметическое замыкание для класса двумерных слов Тёплица, и для некоторых из них вычислена точная арифметическая сложность.

1. Определения и обозначения

Пусть $\Sigma = \{0, 1\}$ — заданный алфавит. Множество всех прямоугольных слов над алфавитом Σ обозначим через Σ^{**}.

Операция *инверсии* заменяет в слове u символы 1 на 0 и 0 на 1. Полученное слово обозначается через $\overline{u}$.

Введём для одномерных слов следующие определения и обозначения. Слово v называется *подсловом* слова u, если $u = x^1 v x^2$ для некоторых слов x^1 и x^2, возможно пустых. *Наращивание* слова $u = u_1 u_2 \dots u_n$ словом v означает вставку слова v между каждыми двумя соседними символами слова u и обозначается через $u Ш v = u_1 v u_2 v \dots v u_n$. Слово, состоящее из n подряд идущих символов a, обозначим через a^n.

Отображение $\varphi : \Sigma^{**} \to \Sigma^{**}$ называется *морфизмом*, если для него выполнено следующее свойство:

$$\varphi \begin{pmatrix} \alpha_{(l,1)} & \dots & \alpha_{(l,n)} \\ & \dots & \\ \alpha_{(1,1)} & \dots & \alpha_{(1,n)} \end{pmatrix} = \begin{matrix} \varphi(\alpha_{(l,1)}) & \dots & \varphi(\alpha_{(l,n)}) \\ & \dots & \\ \varphi(\alpha_{(1,1)}) & \dots & \varphi(\alpha_{(1,n)}) \end{matrix}$$

для любого конечного прямоугольного слова α. Очевидно, что морфизм φ можно задать значениями $\varphi(a) \in \Sigma^{**}$ для $a \in \Sigma$ (образы символов обязаны иметь одинаковый размер). В дальнейшем будут рассматриваться морфизмы, образы которых имеют размер $m \times k$.

В данной работе под двумерным бесконечным словом подразумевается замощение алфавитом правой верхней четверти плоскости:

$$\begin{matrix} & & \dots & \\ w_{(3,1)} & w_{(3,2)} & w_{(3,3)} & \dots \\ w_{(2,1)} & w_{(2,2)} & w_{(2,3)} & \dots \\ w_{(1,1)} & w_{(1,2)} & w_{(1,3)} & \dots \end{matrix}$$

Неподвижной точкой морфизма φ называется двумерное бесконечное слово w, удовлетворяющее равенству $\varphi(w) = w$.

Арифметической подпоследовательностью двумерного бесконечного слова w называется слово $w_{(x,y)} w_{(x+z,y+l)} \dots w_{(x+nz,y+nl)} \dots$ с началом в точке с координатами (x, y) и шагом (z, l). В дальнейшем такая арифметическая подпоследовательность будет обозначаться как $S(x, y, z, l)$.

Подслово некоторой арифметической подпоследовательности двумерного бесконечного слова w называется *арифметическим подсловом* слова w. Множество всех арифметических подслов двумерного бесконечного слова w называется *арифметическим замыканием* слова w. *Арифметической сложностью* двумерного бесконечного слова w называется функция $f(n)$, равная числу его различных арифметических подслов длины n.

В данной работе исследуются неподвижные точки $w(m,k)$ морфизмов следующего вида:

$$\varphi(m,k): 0 \longrightarrow \begin{matrix} 0 & \dots & 0 & 1 \\ & \dots & & \\ 0 & \dots & 0 & 0 \end{matrix}, \quad 1 \longrightarrow \begin{matrix} 0 & \dots & 0 & 0 \\ & \dots & & \\ 0 & \dots & 0 & 0 \end{matrix},$$

где прямоугольные слова имеют размер $m \times k$.

Неподвижные точки таких морфизмов могут быть построены методом Тёплица (он будет описан ниже). Поэтому морфизмы такого вида мы будем называть *морфизмами Тёплица.*

Обозначим через $F(m,k)$ множество всех арифметических подслов неподвижной точки $w(m,k)$, т. е. её арифметическое замыкание.

В дальнейшем под *позицией* будем понимать пару координат.

Позиция (i,j) называется *n-ключевой*, если m^n делит i и k^n делит j, но m^{n+1} не делит i или k^{n+1} не делит j. Символ $w_{(i,j)}$ в этом случае также называется *n-ключевым.* Позиция называется *ключевой*, если она является n-ключевой для некоторого $n > 0$.

Пример. Рассмотрим два способа построения неподвижных точек на примере морфизма $\varphi(2,2)$.

Первый способ основан на последовательном вычислении неподвижной точки из 0: $0 \to \varphi(0) \to \varphi(\varphi(0)) \to \dots$,

$$0 \quad \to \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \quad \to \begin{matrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix} \quad \to \begin{matrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \quad \to \dots.$$

Свойство 1. Заметим, что символы, находящиеся в ключевых позициях каждого слова $\varphi^{n+1}(0)$, образуют слово, являющееся инверсией к слову $\varphi^n(0)$. Следовательно, в слове $\varphi^n(0)$ в любой позиции (ms_1, ks_2), где $s_1, s_2 \in \mathbb{N}$, находится символ, являющийся инверсией к символу, находящемуся в позиции (s_1, s_2). Во всех неключевых позициях слова $\varphi^n(0)$ находится символ 0.

Второй способ построения неподвижной точки морфизма $\varphi(m,k)$ — это конструкция Тёплица. Она базируется на свойстве 1. Сначала заполняются нулями все неключевые позиции. Затем все n-ключевые позиции,

где n нечётно, заполняются единицами, и все n-ключевые позиции, где n чётно, заполняются нулями:

$$
\begin{array}{ccccccccc}
 & & & \cdots & & & & & \\
0 & \diamond & 0 & \diamond & 0 & \diamond & 0 & \diamond & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & \diamond & 0 & \diamond & 0 & \diamond & 0 & \diamond & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & \diamond & 0 & \diamond & 0 & \diamond & 0 & \diamond & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots
\end{array}
\rightarrow
\begin{array}{ccccccccc}
 & & & & \cdots & & & & \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 0 & \diamond & 0 & 1 & 0 & \diamond & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots
\end{array}
$$

В пределе, как нетрудно видеть, получится неподвижная точка $w(m,k)$ морфизма $\varphi(m,k)$.

2. Связь арифметических подпоследовательностей в неподвижных точках

Для того чтобы найти строение слов в арифметическом замыкании, необходимо для начала разобраться в строении арифметических подпоследовательностей, в которые они входят. Для этого рассмотрим арифметическую подпоследовательность $S(x,y,z,t)$ при различных параметрах x,y,z,t.

Лемма 1. *Пусть m, k — натуральные числа не меньше 2, тогда справедливо равенство*

$$S\left(mn, kt, mr+i_1, ks+i_2\right) = \overline{S\left(n, t, lr+\frac{li_1}{m}, ls+\frac{li_2}{k}\right)} \text{Ш} 0^{l-1}, \qquad (1)$$

где $i_1 \in \{0,\ldots,m-1\}$, $i_2 \in \{0,\ldots,k-1\}$, а l равно наименьшему общему кратному чисел $\frac{m}{(m,i_1)}$ и $\frac{k}{(k,i_2)}$.

Доказательство. Рассмотрим сначала случай, когда $i_1 = i_2 = 0$. Все символы подпоследовательности $S\left(mn, kt, mr, ks\right)$ являются ключевыми, так как i-символ последовательности, где $i \geqslant 0$, находится в позиции $(mn+mri, kt+ksi)$. По свойству 1 этот символ является инверсией символа в позиции $(n+ri, t+si)$. Таким образом, подпоследовательность $S\left(mn, kt, mr+i_1, ks+i_2\right)$ получена применением инверсии к $S(n,t,r,s)$, т. е.

$$S(mn,kt,mr,ks) = \overline{S(n,t,r,s)}.$$

Теперь рассмотрим общий случай, когда последовательность имеет вид $S(mn, kt, mr+i_1, ks+i_2)$, где $i_1 \in \{0,\ldots,m-1\}$, а $i_2 \in \{0,\ldots,k-1\}$.

Первый символ последовательности ключевой. Следующий ключевой символ будет стоять в последовательности на месте $l+1$, где l — минимальное натуральное число такое, что m делит li_1 и k делит li_2 (поскольку первая и вторая координаты позиции символа последовательности должны делиться соответственно на m и k).

Заметим, что минимальное число x такое, что z делит xy, равно $z/(z,y)$ для натуральных x,y и $z \neq 0$. Таким образом, l равно наименьшему общему кратному чисел $m/(m,i_1)$ и $k/(k,i_2)$.

Подпоследовательность R, состоящая из ключевых символов последовательности $S(mn,kt,mr+i_1,ks+i_2)$, будет иметь вид

$$S(mn,kt,l(mr+i_1),l(ks+i_2)),$$

где $l(mr+i_1) = mr'$, $l(ks+i_2) = ks'$ для некоторых r', s', так как l выбран таким, что m делит li_1 и k делит li_2. Как было показано выше, последовательность R равна инверсии последовательности $S(n,t,l(mr+i_1)/m,l(ks+i_2)/k)$, так как она состоит только из ключевых символов.

Осталось заметить, что между соседними ключевыми символами последовательности $S(mn,kt,mr+i_1,ks+i_2)$ стоят $l-1$ нулей, и таким образом

$$S\left(mn,kt,mr+i_1,ks+i_2\right) = \overline{S\left(n,t,lr+\frac{li_1}{m},ls+\frac{li_2}{k}\right)}\text{Ш}0^{l-1}.$$

Лемма 1 доказана.

Если последовательность начинается с неключевого символа, то, откинув все первые неключевые символы, получим последовательность, для которой справедлива лемма 1, т. е. $S\left(n,t,r,s\right) = 0^i S\left(mn',kt',r,s\right)$. Если же в ней нет ключевых символов, то она по свойству 1 является полностью нулевой.

Теперь можно перейти к вычислению арифметической сложности.

3. Арифметическая сложность при простом $m = k$

Пусть m равно k и является простым числом. Найдём арифметическую сложность двумерного бесконечного слова $w(m,m)$ для всех простых значений m.

Выражение: слово *u получено наращиванием k нулями*, означает, что слово u является подсловом слова вида $v\text{Ш}0^k$.

Лемма 2. *Слово принадлежит арифметическому замыканию двумерного бесконечного слова $w(m,m)$, где m — простое число, тогда и*

только тогда, когда оно является подсловом слова, полученного за несколько шагов поочередного применения к слову 0^n *инверсии и наращивания* $m-1$ *нулями, начиная с инверсии, где* n *— произвольное натуральное число.*

Доказательство. Пусть u — произвольное арифметическое подслово двумерного бесконечного слова $w(m,m)$ и u является подсловом некоторой арифметической последовательности $S(x,y,mr+i_1,ms+i_2)$, где $i_1,i_2\in\{0,\dots,m-1\}$. Если последовательность не содержит ключевых позиций, то слово нулевое. Утверждение леммы справедливо.

Если последовательность $S(x,y,mr+i_1,ms+i_2)$ содержит ключевые позиции, то в этом случае найдётся начинающаяся с ключевой позиции последовательность $S(mn,mt,mr+i_1,ms+i_2)$, содержащая слово u в качестве подслова и имеющая такой же шаг, что и исходная последовательность $S(x,y,mr+i_1,ms+i_2)$. Чтобы убедиться в существовании такой последовательности, заметим, что для любого k и любых $X,Y>0$ в слове $w(m,m)$ существует подслово $\varphi^k(0)$ такое, что позиция (z,d) левого нижнего символа этого подслова удовлетворяет условиям $z\geqslant X$, $d\geqslant Y$. Пусть k — достаточно большое натуральное число такое, что рассмотренное арифметическое подслово u целиком содержится в $\varphi^k(0)$. Найдём вхождение подслова $\varphi^k(0)$ в $w(m,m)$, позиция (z,d) левого нижнего символа которого удовлетворяет условиям $z\geqslant m(mr+i_1)$, $d\geqslant m(ms+i_2)$. Это вхождение содержит слово u в качестве подслова некоторой арифметической последовательности с шагом $(mr+i_1,ms+i_2)$, проходящей через ключевые позиции. Пусть (z',d') — позиция первого символа данного подслова. Заметим, что $z'\geqslant m(mr+i_1), d'\geqslant m(ms+i_2)$, поэтому в качестве начальной позиции (mn,mt) искомой последовательности $S(mn,mt,mr+i_1,ms+i_2)$ можно рассмотреть некоторую позицию $(z'-j(mr+i_1),d'-j(ms+i_2))$, где $0\leqslant j<m$.

Если слово нулевое, то всё доказано.

Пусть слово u ненулевое. Рассмотрим два случая, когда $i_1=i_2=0$ и $i_1+i_2>0$.

Случай 1: $i_1=i_2=0$. Последовательность имеет вид $S(mn,mt,mr,ms)$. По лемме 1 справедливо равенство:

$$S\left(mn,mt,mr,ms\right)=\overline{S\left(n,t,r,s\right)}. \tag{2}$$

Следовательно, инверсия слова u принадлежит арифметическому замыканию слова $w(m,m)$.

Если в правой последовательности (n,t) — неключевая позиция, то возьмём новую последовательность $S(n-ri,t-si,r,s)$, где i такое мини-

мальное натуральное число, что позиция $(n-ri, t-si)$ является ключевой. Если (r,s) — ключевая позиция, то снова применим случай 1, иначе перейдём ко второму случаю.

Случай 2: $i_1, i_2 \in \{0,\dots,m-1\}$ и $i_1+i_2>0$. Без ограничения общности допустим, что $i_2>0$. Тогда $\frac{m}{(m,i_2)}=m$. Следовательно, $l=m$. По лемме 1 справедливо равенство

$$S\,(mn,mt,mr+i_1,ms+i_2)=\overline{S\,(n,t,mr+i_1,ms+i_2)}\text{Ш}0^{m-1}. \qquad (3)$$

Определим слово v как такое подслово последовательности $S(n,t,mr+i_1,ms+i_2)$ минимальной длины, что слово u является подсловом слова $\overline{v}\text{Ш}0^{m-1}$. Следовательно, слово u получено из слова v с помощью инверсии и наращивания $m-1$ нулями.

Если (n,t) — ключевая позиция, то последовательность $S\,(n,t,mr+i_1,ms+i_2)$, где $i_1,i_2\in\{0,\dots,m-1\}$ и $i_1+i_2>0$, будет относиться ко второму случаю. Иначе, как выше, возьмём новую последовательность, и она будет относиться ко второму случаю. Заметим, что последовательность, полученная во втором случае, никогда не будет относиться к первому случаю, так как её шаг совпадает с предыдущей последовательностью.

Таким образом, пока обе компоненты шага арифметической подпоследовательности, в которой встретилось слово u, делятся на m, можно переходить к арифметической подпоследовательности с шагом в m раз меньше, что соответствует случаю 1. В результате получим слово u_1, равное u или $\bar{u}$, которое встречается в некоторой подпоследовательности, одна из компонент шага которой не делится на m. Теперь можно применять формулу (3), получая каждый раз слово $u_{i+1}=v$ из слова u_i, пока не получим слово, состоящее из одних нулей. Покажем, что рано или поздно процесс завершится.

Утверждение 1. *Последовательность слов $u_1,u_2,\dots$ является конечной.*

Доказательство. Последовательность u_i обладает свойством

$$|u_{i+1}|\leqslant 2+|u_i|/m.$$

Допустим, что все u_i — ненулевые слова. Тогда, начиная с некоторого $i\geqslant N$, длина u_i равна двум. Следовательно, u_N равно одному из слов 11, 10 или 01. Два последних слова через несколько шагов придут к слову 00, так как нет арифметических последовательностей только с одной 1 или одним 0. Рассмотрим слово 11, оно получено с помощью первого или второго случая. В первом случае u_{N-1} равно слову 00, что невозможно по

предположению. Во втором случае u_{N-1} — подслово слова $\overline{1}0^m\overline{1} = 0^{m+2}$. Получили противоречие. Утверждение 1 доказано.

Докажем обратное утверждение, что любое слово, полученное поочередным конечным применением к слову 0^n инверсии и наращивания $m-1$ нулями, начиная с инверсии, принадлежит арифметическому замыканию двумерного бесконечного слова $w(m,m)$.

Применим индукцию по числу применений инверсии и наращивания $m-1$ нулями. Очевидно, что слово 0^n принадлежит арифметическому замыканию слова $w(m,m)$. Лемма 2 доказана.

Пусть слово u получено поочередным применением к нулевому слову инверсии и наращивания $m-1$ нулями, заканчивающимся на наращивание $m-1$ нулями. И пусть это слово является подсловом последовательности $S(n,t,r,s)$, где позиция (r,s) — неключевая (такая последовательность существует). Тогда инверсия слова u будет подсловом последовательности $S(mn,mt,mr,ms)$ по равенству (2).

Применим наращивание $m-1$ нулями к $\overline{u}$, тогда полученное слово v будет подсловом последовательности $S(mn,mt,r,s)$ по равенству (3).

Рассмотрим отдельно случай $m=2$, так как при нём двумерное бесконечное слово $w(m,m)$ имеет некоторую особенность строения арифметических подслов.

Лемма 3. *Для $n \geqslant 2$ в арифметическом замыкании двумерного бесконечного слова $w(2,2)$ число всех подслов длины n слов вида vШ0 равно числу всех подслов длины n слов вида vШ1 и равно $\frac{f(n)}{2}+1$.*

Доказательство. В арифметическом замыкании слова $w(2,2)$ все слова длины не меньше 2 можно однозначно отнести к подсловам слов одного из четырёх классов: 0^n; 1^n; uШ0; vШ1, где u не является нулевым словом и v не является единичным словом. Исключение составляют слова $0101\dots01$, $1010\dots10$ при чётных и $1010\dots01$, $0101\dots10$ при нечётных длинах: они могут считаться полученными наращиванием как единицами, так и нулями. Остаётся только заметить, что число подслов слов вида uШ0 равно числу подслов слов вида vШ1. Лемма 3 доказана.

Лемма 4. *Арифметическая сложность двумерного бесконечного слова $w(2,2)$ при $n \geqslant 2$ и $r \in \{0,1\}$ имеет следующий рекуррентный вид:*

$$f(2n+r) = \begin{cases} 2f(n)+4 & \text{при } r=0; \\ f(n+1)+f(n)+4 & \text{при } r=1. \end{cases}$$

Доказательство. По лемме 2 каждое слово из арифметического замыкания $F(2,2)$ является подсловом слова, полученного за несколько

шагов поочередного применения к нулевому слову инверсии и наращивания одним нулём.

Если слово из арифметического замыкания $F(2,2)$ получено наращиванием одним нулём, то его подслово v нечётной длины имеет один из двух видов: $0u_1 0u_2 \dots u_n 0$ и $u_1 0u_2 \dots u_n 0u_{n+1}$, где u — подслово слова вида hШ1. Следовательно, число таких подслов длины $2n+1$ в точности равно числу подслов длины n и $n+1$ слов вида hШ1. По лемме 3 оно будет равно $\frac{f(n)}{2} + \frac{f(n+1)}{2} + 2$, так как $n \geqslant 2$.

При инверсии слова $0101\dots010$ и $1010\dots101$ переходят друг в друга (это единственные слова длины $2n+1$, обладающие таким свойством), поэтому их инверсия не даёт новых слов. Следовательно, число слов длины $2n+1$ равно $2(\frac{f(n+1)}{2} + 1 + \frac{f(n)}{2} + 1 + 1) - 2 = f(n+1) + f(n) + 4$.

Аналогично доказывается равенство для слов чётной длины. Лемма 4 доказана.

Лемма 5. *Арифметическая сложность двумерного слова $w(2,2)$ равна $f(n) = 4n - 4$ для $n \geqslant 2$.*

Доказательство. При $1 \leqslant n \leqslant 4$ арифметическая сложность имеет следующие значения: $f(1) = 2$, $f(2) = 4$, $f(3) = 8$, $f(4) = 12$. Нетрудно показать, что для всех $n \geqslant 4$ справедливо равенство

$$f(n+1) - f(n) = f\left(\lfloor n/2 \rfloor + 1\right) - f\left(\lfloor n/2 \rfloor\right).$$

Тогда $f(n+1) - f(n) = 4$ при $n \geqslant 2$ и, следовательно, $f(n) = 4n - 4$. Лемма 5 доказана.

Перейдем к рассмотрению случая $m > 2$.

Лемма 6. *Для $n \geqslant 2$ в арифметическом замыкании двумерного бесконечного слова $w(m,m)$ число всех подслов длины n слов вида vШ0^{m-1} равно числу всех подслов длины n слов вида vШ1^{m-1} и равно $f(n)/2$.*

Доказательство. В арифметическом замыкании слова $w(m,m)$ все слова длины не меньше 2 можно однозначно отнести к подсловам слов одного из четырёх классов: 0^n, 1^n, uШ0^{m-1}, vШ1^{m-1}, где u не является нулевым словом и v не является единичным словом. Остаётся заметить, что число подслов слов вида uШ0^{m-1} равно числу подслов слов вида vШ1^{m-1}. Лемма 6 доказана.

Лемма 7. *Арифметическая сложность двумерного бесконечного слова $w(m,m)$ при $q \geqslant 3$ и $0 \leqslant l \leqslant m-1$ имеет следующий рекуррентный вид:*

$$f(qm + l) = f(q+1)l + f(q)(m-l) + 2.$$

Доказательство. Рассмотрим слово v длины $n = qm + l$ из арифметического замыкания $F(m, m)$. По лемме 2 оно является подсловом слова, полученного за несколько шагов поочередного применения к нулевому слову инверсии и наращивания $m - 1$ нулями.

Если слово v является подсловом слова, полученного наращиванием $m-1$ нулями, то оно имеет вид $0^r u_1 0^{m-1} \dots 0^{m-1} u_z 0^s$, где r и s — целые числа такие, что $0 \leqslant r, s \leqslant m - 1$, а слово u принадлежит арифметическому замыканию $F(m, m)$. Так как слово v имеет длину $qm + l$, то $qm+l = r+z+(m-1)(z-1)+s = (z-1)m+r+s+1$. Рассмотрим для r, s и z два случая.

1) $r + s = l - 1$ и $z = q + 1$, тогда $0 \leqslant r \leqslant l - 1$ и $s = l - r - 1 \geqslant 0$. Число возможных различных пар (r, s) в этом случае будет равно l. Слово u есть подслово слова, полученного за несколько шагов поочередного применения к нулевому слову инверсии и наращивания $m-1$ нулями, заканчивающегося инверсией. Следовательно, слово u имеет вид hШ1^{m-1}, а число таких слов по лемме 6 равно $\frac{f(q+1)}{2}$. Таким образом, умножив l на $\frac{f(q+1)}{2}$, получим число всех слов длины n, полученных наращиванием $m - 1$ нулями и удовлетворяющих равенству $r + s = l - 1$.

Пример: $\overbrace{100\dots01\underbrace{00\dots0}_{s=l-1}}^{n}$ при $r = 0$.

2) $r+s = l+m-1$ и $z = q$, тогда $l \leqslant r \leqslant m-1$ и $s = l+(m-1)-r \leqslant m-1$. Число возможных различных пар (r, s) в этом случае будет равно $m - l$. Аналогично предыдущему случаю число слов u длины q равно $\frac{f(q)}{2}$. Следовательно, умножив $m - l$ на $\frac{f(q)}{2}$, получим число всех слов длины n, полученных наращиванием $m - 1$ нулями и удовлетворяющих равенству $r + s = l + m - 1$.

Пример: $\overbrace{\underbrace{00\dots0}_{r=m-1}100\dots01\underbrace{00\dots0}_{s=l}}^{n}$ при $r = m - 1$.

Сложим число слов в обоих случаях, прибавим единицу (учёт нулевого слова) и умножим всё на 2 (учёт инверсий), в результате получим

$$f(qm + l) = 2\left(\frac{f(q+1)}{2}l + \frac{f(q)}{2}(m - l) + 1\right).$$

Раскроем скобки

$$f(qm + l) = f(q + 1)l + f(q)(m - l) + 1.$$

Получили необходимое выражение. Лемма 7 доказана.

Следствие 1. *Для всех $q \geqslant 3$ и $0 \leqslant l \leqslant m-1$ справедливо равенство*

$$f(qm+l+1) - f(qm+l) = f(q+1) - f(q).$$

Лемма 8. *При $n \geqslant 3$ справедливо равенство*

$$f(n+1) - f(n) = \begin{cases} 4, & \text{если } n \in \bigcup\limits_{z\geqslant 1} [m^z, 2m^z), \\ 2, & \text{если } n \in \bigcup\limits_{z\geqslant 1} [2m^z, m^{z+1}). \end{cases}$$

Доказательство. Докажем утверждение леммы индукцией по n.

При $3 \leqslant n \leqslant m$ арифметическая сложность $f(n)$ равна $2n+2$, так как слова длины n в арифметическом замыкании суть в точности слова 0^n, 10^{n-1}, 010^{n-2}, …, $0^{n-1}1$ и их инверсии.

При $m+1 \leqslant n \leqslant 2m$ слова длины n в арифметическом замыкании — это в точности слова

$$0^n,\ 10^{n-1},\ 010^{n-2}, \ldots,\ 0^{n-1}1,\ 10^{m-1}10^{n-m-1},\ \ldots,\ 0^{n-m-1}10^{m-1}1$$

и их инверсии. Следовательно,

$$f(n) = 2(n+1+(n-m)) = 4n-2m+2.$$

По лемме 7 при $2m+1 \leqslant n \leqslant 3m$ арифметическая сложность $f(n)$ равна $2(3(3m-n)+1+4(n-2m)) = 2n+2m+2$. Используя эти значения, вычислим следующую функцию для $3 \leqslant n < 2m$:

$$f(n+1) - f(n) = \begin{cases} 2, & 3 \leqslant n < m; \\ 4, & m \leqslant n < 2m; \\ 2, & 2m \leqslant n < 3m. \end{cases}$$

По следствию 1 имеем $f(n+1) - f(n) = 2$ для $3m \leqslant n < m^2$.

Доказали, что необходимое равенство верно для $m \leqslant n \leqslant m^2$.

Допустим, что равенство верно для $n < i$, докажем его для $n = i > m^2$. В следствии 1 было получено, что $f(n+1) - f(n) = f(\lfloor \frac{n}{m} \rfloor + 1) - f(\lfloor \frac{n}{m} \rfloor)$, а по предположению индукции

$$f\left(\left\lfloor \frac{n}{m} \right\rfloor + 1\right) - f\left(\left\lfloor \frac{n}{m} \right\rfloor\right) = \begin{cases} 4, & \lfloor \frac{n}{m} \rfloor \in \bigcup\limits_{z\geqslant 1} [m^z, 2m^z); \\ 2, & \lfloor \frac{n}{m} \rfloor \in \bigcup\limits_{z\geqslant 1} [2m^z, m^{z+1}). \end{cases}$$

Следовательно,

$$f(n+1)-f(n)=\begin{cases} 4, & n\in\bigcup\limits_{z\geqslant 2}[m^z,2m^z);\\ 2, & n\in\bigcup\limits_{z\geqslant 2}[2m^z,m^{z+1}).\end{cases}$$

Лемма 8 доказана.

Теорема. При $n \geqslant 3$ верны неравенства

$$\frac{2m}{m-1}n-\frac{2}{m-1}\leqslant f(n)\leqslant\frac{3m-2}{m-1}n-\frac{2}{m-1}\leqslant 4n.$$

Доказательство. Докажем сначала левое неравенство.

Нижняя граница $f(n)$. Рассмотрим отрезок $[m^z, m^{z+1}]$ для $z \geqslant 1$. Найдём разность значений функции $f(n)$ на краях этого отрезка, используя лемму 8:

$$f(m^{z+1})-f(m^z)=\sum_{n\in(m^z,m^{z+1}]}(f(n+1)-f(n))=2m^{z+1}. \tag{4}$$

Так как функция $f(n)$ на отрезке $[m^z, m^{z+1}]$ сначала растёт на 4, а затем на 2 значения вверх, то линейная функция φ_z, совпадающая на краях этого отрезка с функцией $f(n)$, будет нижней линейной границей $f(n)$ на этом отрезке.

Найдём функцию $\varphi_z(n) = a_z n + b_z$.

Из равенств $f(m^z) = a_z m^z + b_z$, $f(m^{z+1}) = a_z m^{z+1} + b_z$ и равенства (4) легко видно, что $a_z = \frac{2m}{m-1}$. Получается, что для любого $z \geqslant 1$ функция $\varphi_z(n) = \frac{2m}{m-1}n + b_z$.

Функции φ_{z+1} и φ_z в точке m^{z+1} равны $f(m^{z+1})$ для всех $z \geqslant 1$. Следовательно, $b_{z+1} = b_z$ для всех $z \geqslant 1$. А так как $\varphi_1(m) = f(m) = 2m + 2$, то $b_1 = -\frac{2}{m-1}$.

Получается, что $\varphi_z(n) = \frac{2m}{m-1}n - \frac{2}{m-1}$ при $n \in [m^z, m^{z+1}]$ и всех $z \geqslant 1$. Тогда нижняя граница для $f(n)$ при $n \geqslant 3$ будет иметь вид $\frac{2m}{m-1}n - \frac{2}{m-1}$.

Верхняя граница $f(n)$ доказывается аналогично. Теорема доказана.

Следствие 2. *Функция $f(n)$ растёт линейно: $f(n) = O(n)$.*

Выражаю большую благодарность своему научному руководителю Анне Эдуардовне Фрид за ценные советы при написании этой статьи.

ЛИТЕРАТУРА

1. **Allouche J.-P., Baake M., Cassaigne J., Damanik D.** Palindrom complexity // Theoret. Comput. Sci. — 2003. — Vol. 292. — P. 9–31.

2. **Avgustinovich S. V., Cassaigne J., Frid A. E.** Sequences of low arithmetical complexity // Theoret. Informatics Appl. — 2006. — Vol. 40, N 4. — P. 569–582.
3. **Avgustinovich S. V., Fon-Der-Flaass D. G., Frid A. E.** Arithmetical complexity of infinite words // Words, Languages & Combinatorics III. 2003. P. 51–62, Singapore. World Scientific Publishing. ICWLC 2000, Kyoto, Japan, 2000. March. P. 14–18.
4. **Cassaigne J.** Double sequences with complexity $mn+1$ // J. Autom. Lang. Comb. — 1999. — Vol. 4, N 3. — P. 153–170.
5. **Cassaigne J., Karhumäki J.** Toeplitz words, generalized periodicity and periodically iterated morphisms // European J. Combin. — 1997. — Vol. 18. — P. 497–510.
6. **Ferenczi S.** Complexity of sequences and dynamical systems // Discrete Math. — 1999. — Vol. 206. — P. 145—154.
7. **Frid A. E.** Arithmetical complexity of symmetric DOL words // Theoret. Comput. Sci. — 2003. — Vol. 306. — P. 535–542.
8. **Kamae T., Zamboni L.** Sequence entropy and the maximal pattern complexity of infinite words // Ergodic Theory Dynam. System. — 2002. — Vol. 22. — P. 1191–1199.
9. **Koskas M.** Complexités de suites de Toeplitz // Discrete Math. — 1998. — Vol. 183. — P. 161–183.
10. **Nakashima I., Tamura J.-I., Yasutomi S.-I.** *-Sturmian words and complexity // J. Theórie des Nombres de Bordeaux. — 2003. — Vol. 15. — P. 767–804.
11. **Restivo A., Salemi S.** Binary patterns in infinite binary words // Formal and Natural Computing. Berlin: Springer, 2002. — P. 107–118 (Lect. Notion in Comput. Sci.; Vol. 2300).

Батуева Цындыма Чимит-Доржиевна,
e-mail: cendema@ngs.ru

Статья поступила
13 марта 2008 г.
Переработанный вариант —
9 февраля 2009 г.