

НИЖНИЕ ОЦЕНКИ СЛОЖНОСТИ СИМВОЛЬНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ, ОПРЕДЕЛЯЕМЫХ СИММЕТРИЧЕСКИМИ БУЛЕВЫМИ ФУНКЦИЯМИ*)

Ю. В. Мерекин

Для символьных последовательностей, определяемых классом симметрических булевых функций, в классе схем конкатенации слов получена нижняя оценка сложности, которая для почти всех слов асимптотически совпадает с известной верхней оценкой, а для слов, определяемых симметрическими булевыми функциями, характеристическими последовательностями которых являются последовательности де Брейна, оценка принимает асимптотически максимальное значение.

Мы продолжаем серию работ [1–4] по оценке сложности итеративной процедуры построения слов, использующей на каждом шаге операцию конкатенации двух слов, полученных на предыдущих шагах. В настоящей статье для класса слов, определяемых симметрическими булевыми функциями, получена нижняя оценка сложности, асимптотически совпадающая с верхней оценкой из [3].

Рассматриваются слова в алфавите $\{0, 1\}$. *Длиной* $|W|$ слова W называется число входящих в него символов. Операция *конкатенации* слов U и V определяется как запись слова V за словом U и обозначается через $U \bullet V$. В некоторых случаях знак \bullet опускается. Слово V называется *подсловом* слова W и обозначается через $V \sqsubseteq W$, если для некоторых (возможно, пустых) слов X и Y справедливо равенство $W = X \bullet V \bullet Y$.

Последовательность слов $0, 1, X, Y, \dots, Z$ называется *схемой конкатенации* слова Z и обозначается через S , если для любого слова W из этой последовательности, начиная со слова X , в ней имеются такие слова U, V (возможно, $U = V$), предшествующие слову W , что $W = U \bullet V$. Под *сложностью* $L(S)$ схемы S конкатенации слова Z понимается число слов в последовательности X, Y, \dots, Z . Пусть $L(Z) = \min L(S)$, где минимум

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 99–01–00531).

берется по всевозможным схемам конкатенации слова Z . Величина $L(Z)$ называется *мультипликативной сложностью* слова Z .

При получении нижних оценок сложности используются специальные представления слов. Пусть слово W представлено в виде $W = UxV$, где $x \in \{0, 1\}$. Если V является либо символом, отсутствующим в слове Ux , либо $V \subseteq Ux$ и $xV \not\subseteq U$, то слово V называется *максимальным суффиксом* слова W (однобуквенное слово является максимальным суффиксом). Представление слова Z в виде $Z = Y_1 \bullet Y_2 \bullet \dots \bullet Y_r$ называется *суффиксным представлением*, если длина слова Y_1 равна единице, а всякое слово Y_i , $1 \leq i \leq r$, является максимальным суффиксом слова $Y_1 \bullet Y_2 \bullet \dots \bullet Y_i$. Очевидно, что суффиксное представление любого слова единственно. Число операций конкатенации в суффиксном представлении слова Z называется *суффиксной сложностью* слова Z и обозначается через $L^*(Z)$.

Пусть слово W представлено в виде $W = UV$, где $|V| > 1$. Если $V \not\subseteq U$, то слово V называется *расширенным суффиксом* слова W (при пустом U слово V является расширенным суффиксом). Представление слова Z в виде $Z = Y_1 \bullet Y_2 \bullet \dots \bullet Y_u$ называется *расширенным представлением*, если

(а) каждый из суффиксов Y_1, Y_2, \dots, Y_u является либо максимальным, либо расширенным;

(б) среди Y_1, Y_2, \dots, Y_u содержится хотя бы один расширенный суффикс.

Расширенное представление слова, вообще говоря, не единственно. Число операций конкатенации в i -м расширенном представлении (предполагается, что все такие представления перенумерованы) назовем *сложностью i -го расширенного представления* и обозначим через $L_i^{**}(Z)$.

В [1, 4] доказано, что для произвольного слова Z и любом i выполняются неравенства

$$L(Z) \geq L^*(Z) \geq L_i^{**}(Z). \quad (1)$$

Булева функция называется *симметрической*, если она принимает одно и то же значение на всех наборах с одинаковым числом единиц. Очевидно, что каждая симметрическая булева функция $f(x_1, \dots, x_n)$ может быть задана такой последовательностью $(\sigma_0, \sigma_1, \dots, \sigma_n)$, что σ_i есть значение функции $f(x_1, \dots, x_n)$ на любом наборе значений переменных, содержащем i единиц и $n - i$ нулей. Такую последовательность $(\sigma_0, \sigma_1, \dots, \sigma_n)$ назовем *характеристической* для $f(x_1, \dots, x_n)$ и обозначим через $\hat{\sigma} = \hat{\sigma}(f)$.

Ниже мы используем введенное в [3] отображение ξ последовательности $\hat{\alpha}$ слов $\alpha_0, \alpha_1, \dots, \alpha_t$, $t \geq 1$, в последовательность $\hat{\beta}$ слов $\beta_0, \beta_1, \dots, \beta_{t-1}$, определяемое следующим образом: $\hat{\beta} = \xi(\hat{\alpha}) = \alpha_0 \bullet \alpha_1, \alpha_1 \bullet \alpha_2, \dots$,

$\alpha_{t-1} \bullet \alpha_t$. Если положить $\xi^0(\hat{\alpha}) = \hat{\alpha}$, $\xi^i(\hat{\alpha}) = \xi(\xi^{i-1}(\hat{\alpha}))$, при $i \geq 1$, то множество $\xi^t(\hat{\alpha})$ состоит из одного слова. Непосредственно из определения слова $\xi^t(\hat{\alpha})$ следуют

Предложение 1. При любом $t \geq 1$ справедливо равенство

$$\xi^t(\alpha_0, \dots, \alpha_t) = \xi^{t-1}(\alpha_0, \dots, \alpha_{t-1}) \bullet \xi^{t-1}(\alpha_1, \dots, \alpha_t).$$

Предложение 2. Для любой двоичной последовательности (x_0, \dots, x_i) , $i \geq 0$, x_0 является префиксом, x_i — суффиксом слова $\xi^i(x_0, \dots, x_i)$.

Ясно, что слово $\xi^n(\hat{\sigma})$ имеет длину 2^n . В [3] доказано, что оно совпадает со столбцом значений таблицы истинности (при лексикографическом порядке наборов значений аргументов) симметрической булевой функции $f(x_1, \dots, x_n)$ с характеристической последовательностью $\hat{\sigma} = (\sigma_0, \sigma_1, \dots, \sigma_n)$.

Слова

$$B_0 = \xi^i(\sigma_0, \dots, \sigma_i), B_1 = \xi^i(\sigma_1, \dots, \sigma_{i+1}), \dots, B_{n-i} = \xi^i(\sigma_{n-i}, \dots, \sigma_n)$$

длины 2^i , из которых состоит последовательность слов $\xi^i(\sigma_0, \dots, \sigma_n)$, $1 \leq i \leq n$, назовем i -блоками.

Так как слово $\xi^h(B_j, \dots, B_{j+h})$, $0 \leq j \leq n-i-h$, является конкатенацией i -блоков B_j, \dots, B_{j+h} , то его можно рассматривать как слово, порождаемое отображением ξ из последовательности (B_j, \dots, B_{j+h}) . В этом случае слово назовем (h, i) -блоком.

Если все i -блоки последовательности $\xi^i(\sigma_0, \dots, \sigma_n) = B_0, \dots, B_{n-i}$, $1 \leq i \leq n-2$, различны, то трехблочное слово $B_t B_{t+1} B_{t+1}$, $0 \leq t \leq n-i-1$, назовем *ядром* и обозначим через K_{t+1} .

Из построения слова $\xi^n(\hat{\sigma})$ следуют

Предложение 3. Если в последовательности $\hat{\sigma} = (\sigma_0, \sigma_1, \dots, \sigma_n)$ подпоследовательности $(\sigma_j, \dots, \sigma_{j+i})$, $0 \leq j \leq n-i$, различны, то в последовательности $\xi^i(\hat{\sigma})$ i -блоки B_0, \dots, B_{n-i} различны.

Предложение 4. Если в последовательности $\xi^i(\hat{\sigma})$ i -блоки B_0, \dots, B_{n-i} , $1 \leq i \leq n-1$, различны, то при любых j и h , где $j \geq 0$, $h \geq 1$ и $n-i \geq j+h$, i -блок B_{j+h} является суффиксом (h, i) -блока $\xi^h(B_j, \dots, B_{j+h}) = X B_{j+h}$ и все i -блоки слова X отличны от B_{j+h} .

Известно, что мультипликативная сложность различных слов, определяемых симметрическими булевыми функциями от n переменных, изменяется от линейной (например, для счетчика четности [1]) до квадратичной (например, для монотонных симметрических функций [4]) относительно n . Поэтому при оценке мультипликативной сложности произвольного слова одного параметра n недостаточно. Ниже мы дополнительно используем функцию $m(\hat{\sigma})$, которая принимает значение k ,

если все подпоследовательности $(\sigma_i, \dots, \sigma_{i+k})$, $0 \leq i \leq n - k$, последовательности $\hat{\sigma}$ различны, а среди подпоследовательностей $(\sigma_i, \dots, \sigma_{i+k-1})$, $0 \leq i \leq n - k + 1$, встречаются одинаковые подпоследовательности. Очевидно, что для любых $\hat{\sigma}$ значение k не может быть меньше $\log_2 n$. Для почти всех последовательностей $k < (2 + \varepsilon) \log_2 n$. Для произвольного слова $\xi^n(\hat{\sigma})$ в классе схем конкатенации слов получим нижнюю оценку мультипликативной сложности как функцию от n и $m(\hat{\sigma})$.

Ниже доказываются четыре леммы и теорема. Первая лемма — вспомогательная, а в остальных выявляются свойства ядра. Основное свойство ядра — одноразовое вхождение в определенное слово — используется в теореме для выделения расширенных суффиксов и, следовательно, для получения нижней оценки мультипликативной сложности.

Лемма 1. Пусть при $r \geq 2$ двоичные последовательности $\hat{a} = (a_0, \dots, a_r)$ и $\hat{d} = (d_0, \dots, d_r)$ таковы, что имеются двоичные слова X, Y такие, что $1 \leq |X| = |Y| < 2^{r-1}$ и $X\xi^r(\hat{a}) = \xi^r(\hat{d})Y$. Тогда существует $x \in \{0, 1\}$ такое, что

$$\xi^r(\hat{a}) = \underbrace{x \dots x}_{2^{r-1}} a_r, \quad \xi^r(\hat{d}) = d_0 \underbrace{x \dots x}_{2^{r-1}}.$$

Доказательство. Для доказательства утверждения леммы достаточно показать справедливость равенств

$$a_0 = a_1 = \dots = a_{r-1} = d_1 = d_2 = \dots = d_r. \quad (2)$$

Пусть условия леммы выполнены. Если $r = 2$, то $|X| = |Y| = 1$ и $Xa_0a_1a_1a_2 = d_0d_1d_1d_2Y$. Из равенства $a_0a_1a_1 = d_1d_1d_2$ легко следует справедливость леммы. Пусть $r > 2$. Слова $A = \xi^r(\hat{a})$ и $D = \xi^r(\hat{d})$ представим в виде $A = A_1A_2A_3$ и $D = D_1D_2D_3$, где $|A_1| = |D_3| = 2^{r-1}$, $|A_2| = |D_2|$ и $|A_3| = |Y| = |D_1| = |X|$.

Рассмотрим два случая: $|X| = |Y| = 2^{r-2}$ и $|X| = |Y| \neq 2^{r-2}$.

Случай 1. Пусть $|X| = |Y| = 2^{r-2}$. Дважды применив предложение 1 к каждому слову $A = \xi^r(a_0, \dots, a_r)$ и $D = \xi^r(d_0, \dots, d_r)$, получаем

$$\begin{aligned} XA &= X \bullet A_1 \bullet A_2 \bullet A_3 \\ &= X \bullet \xi^{r-2}(a_0, \dots, a_{r-2}) \xi^{r-2}(a_1, \dots, a_{r-1}) \bullet \xi^{r-2}(a_1, \dots, a_{r-1}) \bullet A_3, \end{aligned}$$

$$\begin{aligned} DY &= D_1 \bullet D_2 \bullet D_3 \bullet Y \\ &= D_1 \bullet \xi^{r-2}(d_1, \dots, d_{r-1}) \bullet \xi^{r-2}(d_1, \dots, d_{r-1}) \xi^{r-2}(d_2, \dots, d_r) \bullet Y. \end{aligned}$$

Из равенства $A_1A_2 = D_2D_3$, где

$$\begin{aligned} A_1A_2 &= \xi^{r-2}(a_0, \dots, a_{r-2}) \xi^{r-2}(a_1, \dots, a_{r-1}) \xi^{r-2}(a_1, \dots, a_{r-1}), \\ D_2D_3 &= \xi^{r-2}(d_1, \dots, d_{r-1}) \xi^{r-2}(d_1, \dots, d_{r-1}) \xi^{r-2}(d_2, \dots, d_r), \end{aligned}$$

следуют равенства

$$\begin{aligned}\xi^{r-2}(a_0, \dots, a_{r-2}) &= \xi^{r-2}(d_1, \dots, d_{r-1}), \\ \xi^{r-2}(a_1, \dots, a_{r-1}) &= \xi^{r-2}(d_1, \dots, d_{r-1}), \\ \xi^{r-2}(a_1, \dots, a_{r-1}) &= \xi^{r-2}(d_2, \dots, d_r)\end{aligned}$$

и, следовательно, равенство (2).

СЛУЧАЙ 2. Пусть $|X| = |Y| \neq 2^{r-2}$. Рассмотрим два подслучая: $1 \leq |X| = |Y| < 2^{r-2}$ и $2^{r-2} < |X| = |Y| < 2^{r-1}$.

Доказательства в обоих подслучаях аналогичны. Поэтому остановимся на втором из них. Пусть $2^{r-2} < |X| = |Y| < 2^{r-1}$. Доказательство проведем индукцией по r . При $r = 3$ имеем $|X| = 3$ и

$$\begin{aligned}XA &= X \bullet A_1 \bullet A_2 \bullet A_3 = X \bullet a_0 a_1 a_1 a_2 \bullet a_1 \bullet a_2 a_2 a_3, \\ DY &= D_1 \bullet D_2 \bullet D_3 \bullet Y = d_0 d_1 d_1 \bullet d_2 \bullet d_1 d_2 d_2 d_3 \bullet Y.\end{aligned}$$

Из равенства $a_0 a_1 a_1 a_2 a_1 = d_2 d_1 d_2 d_2 d_3$ следует, что $a_0 = a_1 = a_2 = d_1 = d_2 = d_3$.

Предположим, что лемма верна при всех $i \leq r-1$. Используя предположение 1, слова $A = \xi^r(a_0, \dots, a_r)$ и $D = \xi^r(d_0, \dots, d_r)$, входящие в равенство $XA = DY$, представим в виде

$$\begin{aligned}XA &= X \bullet A_1 \bullet A_2 A_3 = X \bullet \xi^{r-1}(a_0, \dots, a_{r-1}) \bullet \xi^{r-1}(a_1, \dots, a_r), \\ DY &= D_1 D_2 \bullet D_3 \bullet Y = \xi^{r-1}(d_0, \dots, d_{r-1}) \bullet \xi^{r-1}(d_1, \dots, d_r) \bullet Y.\end{aligned}$$

Слова $A_1 A_2 = \xi^{r-1}(a_0, \dots, a_{r-1}) \bullet A_2$ и $D_2 D_3 = D_2 \bullet \xi^{r-1}(d_1, \dots, d_r)$ удовлетворяют условиям леммы. Тогда по индуктивному предположению имеем

$$a_1 = a_2 = \dots = a_{r-1} = d_1 = d_2 = \dots = d_{r-1}. \quad (3)$$

Из предложения 2 следует, что подслово $A_2 = A'_2 a_i$ слова $A_2 A_3 = \xi^{r-1}(a_1, \dots, a_r) = A_2 A'_2 a_r$ не содержит a_r и, следовательно, $a_i \in \{a_1, \dots, a_{r-1}\}$. В слове $D = \xi^r(d_0, \dots, d_r)$ буква d_r является суффиксом. Из равенства $XA_1 A_2 = D$ следует, что $a_i = d_r$ и $d_r \in \{a_1, \dots, a_{r-1}\}$. Поэтому цепочку равенств (3) можно продолжить, включив в нее букву d_r .

Аналогично, анализируя равенство $A = D_2 D_3 Y$, цепочку равенств (3) можно продолжить, включив в нее букву a_0 . В результате получим равенства (2). Лемма 1 доказана.

Лемма 2. Пусть $\hat{\sigma} = (\sigma_0, \dots, \sigma_n)$, m — такое натуральное число, что $2 \leq m < n$, и все m -блоки из последовательности $\xi^m(\hat{\sigma}) = B_0, \dots, B_{n-m}$ различны. Тогда для любых двоичных слов X, Y таких,

что $1 \leq |X| < 2^m$ и $|X| + |Y| = 2^m$, при любых $i, j, k \in \{0, \dots, n - m - 1\}$ и $l, p \in \{0, \dots, n - m\}$ справедливы неравенства

$$XB_i B_{i+1} B_{i+1} Y \neq B_j B_{j+1} B_k B_{k+1}, \quad XB_i B_{i+1} B_{i+1} Y \neq B_l B_j B_{j+1} B_p.$$

Доказательство. Предположим, что

$$XB_i B_{i+1} B_{i+1} Y = B_j B_{j+1} B_k B_{k+1}. \quad (4)$$

Рассмотрим три случая: $1 \leq |X| < 2^{m-1}$, $|X| = 2^{m-1}$, $2^{m-1} < |X| < 2^m$.

Пусть $1 \leq |X| < 2^{m-1}$. Из (4) следует равенство $XB_i = B_j B'_{j+1}$, где B'_{j+1} — префикс блока B_{j+1} и $|B'_{j+1}| = |X|$. Обозначив через $a^{(b)}$ слово длины b в алфавите $\{a\}$ и применив к равенству $XB_i = B_j B'_{j+1}$ лемму 1, получаем $B_i = x^{(2^{m-1})} \sigma_{i+m}$ и $B_j = \sigma_j x^{(2^{m-1})}$. Из $B_j = \xi^m(\sigma_j, \dots, \sigma_{j+m}) = \sigma_j x^{(2^{m-1})}$ с учетом предложения 2 получаем $B_{j+1} = \xi^m(\sigma_{j+1}, \dots, \sigma_{j+m+1}) = x^{(2^{m-1})} \sigma_{j+m+1}$. Следовательно, $B'_{j+1} = x^{(|X|)}$. Слово XB_i оканчивается символом σ_{i+m} , а слово $B_j B'_{j+1}$ — символом x . Поэтому $\sigma_{i+m} = x$ и $B_i = x^{(2^m)}$.

Аналогично из равенства суффиксов $B''_{i+1} B_{i+1} Y$ и $B_k B_{k+1}$ одинаковых слов (4), где B''_{i+1} — суффикс m -блока B_{i+1} длины $|X|$, следует, что m -блок B_{i+1} имеет вид $B_{i+1} = x^{(2^m)}$.

По доказанному $B_i = B_{i+1} = x^{(2^m)}$, а по условиям леммы все m -блоки B_0, \dots, B_{n-m} различны. Полученное противоречие доказывает лемму 2 для случая $1 \leq |X| < 2^{m-1}$.

Для случая $2^{m-1} < |X| < 2^m$ доказательство аналогично.

Пусть $|X| = 2^{m-1}$. Используя предложение 1, зададим XB_i и $B_j B_{j+1}$ из (4) в виде

$$XB_i = X \bullet \xi^{m-1}(\sigma_i, \dots, \sigma_{i+m-1}) \bullet \xi^{m-1}(\sigma_{i+1}, \dots, \sigma_{i+m}) = X \bullet B'_i \bullet B''_i,$$

$$\begin{aligned} B_j B_{j+1} &= B'_j \bullet \xi^{m-1}(\sigma_{j+1}, \dots, \sigma_{j+m}) \bullet \xi^{m-1}(\sigma_{j+1}, \dots, \sigma_{j+m}) \bullet B''_{j+1} \\ &= B'_j \bullet B'_j \bullet B'_{j+1} \bullet B''_{j+1}, \end{aligned}$$

где $|X| = |B'_i| = |B''_i| = |B'_j| = |B''_j| = |B'_{j+1}| = |B''_{j+1}| = 2^{m-1}$.

Из равенства (4) следуют равенства $B'_i = B''_j$ и $B''_i = B'_{j+1}$, т. е.

$$\begin{aligned} \xi^{m-1}(\sigma_i, \dots, \sigma_{i+m-1}) &= \xi^{m-1}(\sigma_{j+1}, \dots, \sigma_{j+m}), \\ \xi^{m-1}(\sigma_{i+1}, \dots, \sigma_{i+m}) &= \xi^{m-1}(\sigma_{j+1}, \dots, \sigma_{j+m}). \end{aligned}$$

Поэтому $\xi^{m-1}(\sigma_i, \dots, \sigma_{i+m-1}) = \xi^{m-1}(\sigma_{i+1}, \dots, \sigma_{i+m})$, а следовательно, $(\sigma_i, \dots, \sigma_{i+m-1}) = (\sigma_{i+1}, \dots, \sigma_{i+m})$, т. е. $\sigma_i = \dots = \sigma_{i+m}$. Тогда m -блок B_i имеет вид $B_i = x^{(2^m)}$.

Аналогичные рассуждения для $B_{i+1} Y$ и $B_k B_{k+1}$ из (4) позволяют убедиться в том, что m -блок B_{i+1} имеет вид $B_{i+1} = x^{(2^m)}$.

По доказанному $B_i = B_{i+1} = x^{(2^m)}$, а по условиям леммы все m -блоки B_0, \dots, B_{n-m} различны. Полученное противоречие завершает доказательство леммы 2 при условии, что $X B_i B_{i+1} B_{i+1} Y \neq B_j B_{j+1} B_k B_{k+1}$.

В случае $X B_i B_{i+1} B_{i+1} Y \neq B_l B_j B_{j+1} B_p$ доказательство аналогично. Лемма 2 доказана.

Лемма 3. Пусть $\hat{\sigma} = (\sigma_0, \dots, \sigma_n)$, m — такое натуральное число, что $2 \leq m < n$, и все m -блоки из последовательности $\xi^m(\hat{\sigma}) = B_0, \dots, B_{n-m}$ различны. Тогда если при некотором i , $0 \leq i \leq n-m-1$, справедливо равенство $\xi^n(\hat{\sigma}) = X B_i B_{i+1} B_{i+1} Y$, то либо X — пустое слово, либо $|X|$ кратна 2^m .

Доказательство. Пусть $\xi^m(\hat{\sigma}) = B_0, \dots, B_{n-m}$. Тогда

$$\begin{aligned}\xi^{m+1}(\hat{\sigma}) &= \xi(B_0, \dots, B_{n-m}) = (B_0 B_1, B_1 B_2, \dots, B_{n-m-1} B_{n-m}), \\ \xi^n(\hat{\sigma}) &= \xi^{n-m-1}(B_0 B_1, B_1 B_2, \dots, B_{n-m-1} B_{n-m}).\end{aligned}$$

Следовательно, слово $\xi^n(\hat{\sigma})$ может быть получено последовательным приписыванием справа к слову $B_0 B_1$ слов вида $B_j B_{j+1}$, $0 < j \leq n-m-1$. Поэтому любое подслово слова $\xi^n(\hat{\sigma})$, состоящее из любых четырех последовательных m -блоков, имеет вид $B_j B_{j+1} B_k B_{k+1}$ либо $B_l B_j B_{j+1} B_p$. Согласно лемме 2 ядро K_{i+1} не является подсловом четырехблочного слова $B_j B_{j+1} B_k B_{k+1}$ или $B_l B_j B_{j+1} B_p$, если K_{i+1} имеет непустое пересечение со всеми четырьмя m -блоками. Поэтому ядро K_{i+1} всегда совпадает с тремя соседними m -блоками $A_j A_{j+1} A_{j+2}$, $1 \leq j \leq 2^{n-m} - 3$, из представления $\xi^n(\hat{\sigma}) = A_1 \dots A_{2^{n-m}}$. Лемма 3 доказана.

Лемма 4. Пусть $\hat{\sigma} = (\sigma_0, \dots, \sigma_n)$, m — такое натуральное число, что $2 \leq m < n$, и все m -блоки из последовательности $\xi^m(\hat{\sigma}) = B_0, \dots, B_{n-m}$ различны. Тогда для любых натуральных чисел t и h , $2 \leq t \leq h \leq n-m$, ядро $K_{h-1} = B_{h-2} B_{h-1} B_{h-1}$ имеет единственное вхождение в слово $\xi^t(B_{h-t}, \dots, B_h) = X_t K_{h-1} B_h$.

Доказательство. Из m -блоков B_{h-t}, \dots, B_h построим последовательность $(2, m)$ -блоков

$$\xi^2(B_{h-t}, \dots, B_h) = \xi^2(B_{h-t}, B_{h-t+1}, B_{h-t+2}), \dots, \xi^2(B_{h-2}, B_{h-1}, B_h). \quad (5)$$

Затем из $(2, m)$ -блоков последовательности (5) построим (t, m) -блок

$$\begin{aligned}\xi^t(B_{h-t}, \dots, B_h) &= \xi^{t-2}(\xi^2(B_{h-t}, \dots, B_h)) \\ &= \xi^{t-2}(B_{h-t} B_{h-t+1} B_{h-t+1} B_{h-t+2}, \dots, B_{h-2} B_{h-1} B_{h-1} B_h). \quad (6)\end{aligned}$$

Согласно предложению 4 $(2, m)$ -блок

$$\xi^2(B_{h-2}, B_{h-1}, B_h) = B_{h-2} B_{h-1} B_{h-1} B_h \quad (7)$$

встречается в (t, m) -блоке (6) единственный раз и является суффиксом слова. Среди $(2, m)$ -блоков из последовательности (5) имеется единственный $(2, m)$ -блок (7), в котором содержится ядро $K_{h-1} = B_{h-2}B_{h-1}B_{h-1}$.

Покажем, что ядро K_{h-1} имеет единственное вхождение в (t, m) -блок (6). Предположим, что в (t, m) -блоке (6) существует второе вхождение ядра K_{h-1} . В этом случае начало ядра является суффиксом некоторого $(2, m)$ -блока $B_iB_{i+1}B_{i+1}B_{i+2}$ из (5), а конец — префиксом другого $(2, m)$ -блока $B_jB_{j+1}B_{j+1}B_{j+2}$ из (5). Возможны два случая:

(i) ядро K_{h-1} имеет непустое пересечение только с тремя m -блоками слова $B_{i+1}B_{i+2}B_jB_{j+1}$, т. е. $K_{h-1} = B_{i+1}B_{i+2}B_j$;

(ii) ядро K_{h-1} имеет непустое пересечение с четырьмя m -блоками слова $B_{i+1}B_{i+1}B_{i+2}B_jB_{j+1}B_{j+1}$.

В первом случае $K_{h-1} = B_{h-2}B_{h-1}B_{h-1} = B_{i+1}B_{i+2}B_j$. Следовательно, $B_{h-1} = B_j$ и среди $(2, m)$ -блоков (5) должен существовать $(2, m)$ -блок $\xi^2(B_j, B_{j+1}, B_{j+2}) = \xi^2(B_{h-1}, B_h, B_{h+1})$. Однако среди $(2, m)$ -блоков (5), где представлены все возможные $(2, m)$ -блоки (t, m) -блока (6), такой $(2, m)$ -блок отсутствует.

Существование второго вхождения ядра K_{h-1} при непустом пересечении с четырьмя m -блоками противоречит лемме 3. Лемма 4 доказана.

Через $m(\hat{\sigma})$ обозначим такое минимальное k , что все подпоследовательности $(\sigma_i, \sigma_{i+1}, \dots, \sigma_{i+k})$, $0 \leq i \leq n - k$, последовательности $\hat{\sigma} = (\sigma_0, \dots, \sigma_n)$ различны.

Теорема. Пусть $\hat{\sigma} = (\sigma_0, \dots, \sigma_n)$ — такая двоичная последовательность, что $2 \leq m(\hat{\sigma}) \leq n - 3$. Тогда

$$L(\xi^n(\hat{\sigma})) \geq \binom{n - m(\hat{\sigma}) - 1}{2}.$$

Доказательство. Построим расширенное суффиксное представление слова $\xi^n(\hat{\sigma})$ и найдем сложность этого представления. Для этого проведем пошаговое выделение расширенных суффиксов слова $\xi^{n-k}(B_0, \dots, B_{n-k})$. На каждом шаге выделим расширенные суффиксы, составляющие правую половину слова.

На первом шаге применим предложение 1 сначала к слову $\xi^{n-k}(B_0, \dots, B_{n-k})$, а затем к каждому суффиксу $\xi^{n-k-j}(B_j, \dots, B_{n-k})$, $1 \leq j \leq n - k - 3$, из очередного представления. В результате получим

$$\begin{aligned} \xi^{n-k}(B_0, \dots, B_{n-k}) &= \xi^{n-k-1}(B_0, \dots, B_{n-k-1}) \bullet \xi^{n-k-1}(B_1, \dots, B_{n-k}) \\ &= \xi^{n-k-1}(B_0, \dots, B_{n-k-1}) \bullet \xi^{n-k-2}(B_1, \dots, B_{n-k-1}) \bullet \xi^{n-k-2}(B_2, \dots, B_{n-k}) \\ &= \dots = \xi^{n-k-1}(B_0, \dots, B_{n-k-1}) \bullet \xi^{n-k-2}(B_1, \dots, B_{n-k-1}) \bullet \dots \bullet \\ &\quad \xi^2(B_{n-k-3}, \dots, B_{n-k-1}) \bullet \xi^2(B_{n-k-2}, \dots, B_{n-k}). \end{aligned} \quad (8)$$

Согласно предложению 4 каждое (кроме последнего) из выделенных в (8) слов представимо в виде $\xi^t(B_{n-k-1-t}, \dots, B_{n-k-1}) = Y_t B_{n-k-1}$, $2 \leq t \leq n-k-1$, где все k -блоки слова Y_t отличны от B_{n-k-1} . Согласно лемме 4 ядро K_{n-k-2} имеет единственное вхождение в слово

$$\xi^t(B_{n-k-1-t}, \dots, B_{n-k-1}) = X_t K_{n-k-2} B_{n-k-1}, \quad 2 \leq t \leq n-k-1. \quad (9)$$

По лемме 4 ядро K_{n-k-1} из суффикса

$$\xi^2(B_{n-k-2}, \dots, B_{n-k}) = K_{n-k-1} B_{n-k} \quad (10)$$

имеет единственное вхождение в слово $\xi^{n-k}(B_0, \dots, B_{n-k})$.

Покажем, что в слове $\xi^{n-k}(B_0, \dots, B_{n-k})$ ядро K_{n-k-2} имеет вхождение только в суффиксы слов (9) из разложения (8). Предположим, что ядро K_{n-k-2} имеет непустое пересечение одновременно с двумя соседними словами $X_{i+1} K_{n-k-2} B_{n-k-1}$ и $X_i K_{n-k-2} B_{n-k-1}$, $2 \leq i \leq n-k-2$, из (8). Возможны два случая:

(i) ядро K_{n-k-2} имеет непустое пересечение только с тремя k -блоками;

(ii) ядро K_{n-k-2} имеет непустое пересечение с четырьмя k -блоками.

Первый случай исключается, так как k -блок B_{n-k-1} не является одним из трех k -блоков ядра K_{n-k-2} . Второй случай противоречит лемме 3. Следовательно, в словах (9) и (10) представлены все ядра K_{n-k-2} и K_{n-k-1} слова $\xi^n(\hat{\sigma})$.

Учитывая (9) и (10), разложенное слово $\xi^{n-k}(B_0, \dots, B_{n-k})$ из (8) перепишем в виде

$$(X_{n-k-1} K_{n-k-2} B_{n-k-1}) \bullet (X_{n-k-2} K_{n-k-2} B_{n-k-1}) \bullet \dots \bullet (X_3 K_{n-k-2} B_{n-k-1}) \bullet (X_2 K_{n-k-2} B_{n-k-1}) \bullet (K_{n-k-1} B_{n-k}), \quad (11)$$

где X_2 пусто. Перегруппируем под слова в выражении (11) следующим образом:

$$(X_{n-k-1} K_{n-k-2}) \bullet (B_{n-k-1} X_{n-k-2} K_{n-k-2}) \bullet \dots \bullet (B_{n-k-1} X_3 K_{n-k-2}) \bullet (B_{n-k-1} X_2 K_{n-k-2}) \bullet (B_{n-k-1} K_{n-k-1} B_{n-k}). \quad (12)$$

Покажем, что слова в последних $n-k-2$ скобках, которые разделены знаком \bullet , являются расширенными суффиксами. Так как ядро K_{n-k-1} единственно, то $B_{n-k-1} K_{n-k-1} B_{n-k}$ — расширенный суффикс. Отметим некоторые особенности выделенных в (12) слов (за исключением последнего):

(а) среди k -блоков слова $X_{n-k-1} K_{n-k-2}$ отсутствует k -блок B_{n-k-1} , а в каждое слово $B_{n-k-1} X_t K_{n-k-2}$, $2 \leq t \leq n-k-2$, этот k -блок входит один раз;

(b) все вхождения ядра K_{n-k-2} в слово $\xi^n(\hat{\sigma})$ представлены в (12) и в каждое из выделенных слов $B_{n-k-1}X_tK_{n-k-2}$, $2 \leq t \leq n-k-2$, такое ядро входит единственный раз;

(c) число k -блоков слова $X_{n-k-1}K_{n-k-2}$ равно $2^{n-k-1} - 1$, а в каждом слове $B_{n-k-1}X_tK_{n-k-2}$, $2 \leq t \leq n-k-2$, число k -блоков равно 2^t .

Если все k -блоки B_{n-k-1} являются префиксами, ядра K_{n-k-2} — суффиксами, длина каждого очередного выделенного слова меньше длины предыдущего, а длины всех таких слов кратны 2^k , то (12) задает расширенное суффиксное представление слова $\xi^n(\hat{\sigma})$. Таким образом, на первом шаге в слове $\xi^n(\hat{\sigma})$ выделено $n-k-2$ расширенных суффиксов.

Дальнейшее пошаговое разложение слов $\xi^s(B_0, \dots, B_s)$, $3 \leq s \leq n-k-1$, аналогично. На каждом шаге все суффиксы, кроме последнего, имеют вид (12). В последнем суффиксе отсутствует k -блок B_s , так как на предыдущем шаге этот k -блок используется в крайнем левом суффиксе $B_sX_{s-1}K_{s-1}$. Поэтому крайний правый суффикс слова $\xi^s(B_0, \dots, B_s)$ имеет вид $B_{s-1}K_{s-1}$, оставаясь (в силу единственности ядра K_{s-1}) расширенным суффиксом. Число расширенных суффиксов, полученных на очередном шаге, сокращается на единицу по сравнению с предыдущим шагом. В результате после $n-k-2$ шагов разложения выделяется $(n-k-2) + (n-k-3) + \dots + 1 = \binom{n-k-1}{2}$ расширенных суффиксов и остается неразложенным префикс $B_0B_1B_1$, который по определению тоже является расширенным суффиксом. Поэтому $L_i^{**}(\xi^n(\hat{\sigma})) = \binom{n-m_2(\hat{\sigma})-1}{2}$. Отсюда из (1) следует, что $L(\xi^n(\hat{\sigma})) \geq L_i^{**}(\xi^n(\hat{\sigma}))$. Теорема доказана.

Полученная нижняя оценка мультипликативной сложности слов, определяемых симметрическими булевыми функциями, для почти всех слов асимптотически совпадает с верхней оценкой, полученной в [3] для класса всех рассматриваемых слов

$$\binom{n - (2 + \varepsilon) \log_2 n - 1}{2} \leq L(\xi^n(\hat{\sigma})) \leq \binom{n + 1}{2}.$$

Если дополнительно учитывать операции, необходимые для формирования k -блоков B_0, \dots, B_{n-k} , то возможно увеличение нижней оценки сложности. Покажем возможности такого подхода на примере монотонных симметрических функций. Пусть значения символов левой половины последовательности $\hat{\sigma}$ равны нулю, а правой половины — единице. Так как все слова в последовательности $\xi^{n/2}(\hat{\sigma})$ различны, то по теореме имеем $L(\xi^n(\hat{\sigma})) \geq \binom{n/2-1}{2} > n^2(1 - o(1))/8$. Аналогичная оценка из [4], при получении которой учитывались особенности монотонных функций, равна $n^2(1 - o(1))/4$.

Более существенное повышение нижней оценки сложности за счет вклада от префикса $B_0B_1B_1$ (который в теореме рассматривается как

расширенный суффикс) достигается при значительном отличии $m(\hat{\sigma})$ от $\log_2 n$. Например, при $m(\hat{\sigma}) \sim n$ согласно теореме нижняя оценка сложности может быть равна константе. В этом случае значительный вклад в оценку сложности вносят k -блоки B_0 и B_1 . Например, для счетчика четности в результате построения суффиксного представления достигается линейная от n оценка сложности [1].

Ранее А. А. Евдокимов высказал предположение, что среди слов, определяемых симметрическими булевыми функциями, характеристическими последовательностями которых являются последовательности де Брейна [5, с. 128], содержатся слова с максимальной мультипликативной сложностью. Полученная оценка не дает ответа на эту гипотезу, но для любой последовательности де Брейна дает нижнюю оценку $L(\xi^n(\hat{\sigma})) \geq \binom{n - \log_2 n - 1}{2} > n^2(1 - o(1))/2$.

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 сложности символьных последовательностей, определяемых линейными булевыми функциями // Сибирский журнал индустриальной математики. 1998. Т. 1, № 1. С. 145–147.
3. Мерекин Ю. В. Верхние оценки сложности символьных последовательностей, порождаемых симметрическими булевыми функциями // Дискрет. анализ и исслед. операций. Сер. 1. 1998. Т. 5, № 3. С. 38–43.
4. Мерекин Ю. В. Нижние оценки мультипликативной сложности символьных последовательностей, определяемых монотонными симметрическими булевыми функциями // Дискрет. анализ и исслед. операций. Сер. 1. 1999. Т. 6, № 3. С. 3–9.
5. Холл М. Комбинаторика. М.: Мир, 1970.

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
16 декабря 1999 г.