ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—март 1996. Том 3, № 1, 52–56

УДК 519.714

НИЖНЯЯ ОЦЕНКА СЛОЖНОСТИ ДЛЯ СХЕМ КОНКАТЕНАЦИИ СЛОВ*)

Ю. В. Мерекин

Приводится метод получения нижних оценок сложности для схем конкатенации слов. В частности, для последовательности де Брейна получена нижняя оценка вида $l/\log_2 l$, где l — длина слова. Доказывается, что сложность линейной булевой функции $x_1 \oplus \cdots \oplus x_k$ в этом классе схем равна $2k-1$.

Введение

Одним из обобщений известной задачи о быстром возведении числа в заданную степень, или вычислении функции a^n, является задача о быстром получении заданного слова из букв алфавита и, быть может, некоторого данного множества слов с помощью определенных операций, которые естественны для исследуемой модели вычисления и достаточно просты в реализации.

Мы рассматриваем простейший вариант задачи синтеза слова — получение двоичного слова из букв 0 и 1 с помощью операции конкатенации двух слов, когда разрешается многократное использование уже построенных слов, подобно тому, как это делается в моделях синтеза схем из функциональных элементов, когда реализуются булевы функции [1, 2]. Вначале мы даем метод получения нижних оценок сложности для схем конкатенации, который использует специальное представление слов, а затем находим мультипликативную сложность двоичного слова, являющегося начальным отрезком длины 2^k последовательности Туэ — Морса. Эта последовательность не содержит трех равных последовательных подслов и обладает другими свойствами, интерес к которым возникает в различных математических исследованиях [3–6]. Начальный отрезок длины 2^k последовательности Туэ — Морса совпадает со столбцом значений линейной булевой функции $x_1 \oplus x_2 \oplus \ldots \oplus x_k$, называемой счетчиком четности, сложность реализации которого известна для

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 93-01-01484).

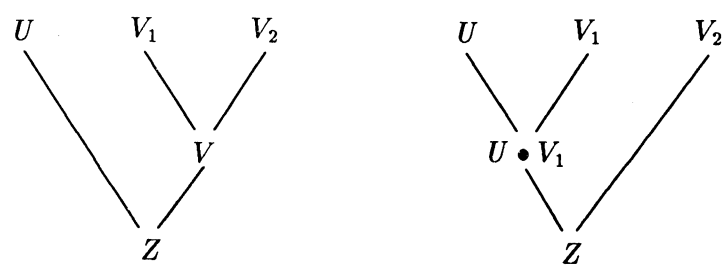
различных классов управляющих систем [7–9]. Тем самым мы дополняем результаты нахождением сложности счетчика четности в классе схем конкатенации слов.

Рассматриваются слова в алфавите $\{0,1\}$. *Длиной* слова W называется число входящих в него символов. Операция *конкатенации* слов U и V определяется как последовательная запись этих слов и обозначается через $U \bullet V$. Слово V называется *подсловом* слова W, если для некоторых (возможно, пустых) слов X и Y справедливо равенство $W = X \bullet V \bullet Y$. Последовательность слов $0, 1, X, Y, \ldots, Z$ называется *схемой конкатенации* слова Z и обозначается через S, если для любого слова W из этой последовательности, начиная со слова X, в ней существуют такие слова U, V (возможно, $U = V$), предшествующие слову W, что $W = U \bullet V$. Обозначим через $L(S)$ число слов в последовательности $X, Y, \ldots, Z$ и назовем *сложностью* схемы S. Пусть $L(Z) = \min L(S)$, где минимум берется по всевозможным схемам конкатенации слова Z. Эту величину назовем *мультипликативной сложностью* слова Z. Схему S назовем *оптимальной*, если $L(Z) = L(S)$.

§ 1. Нижняя оценка мультипликативной сложности слов

Лемма 1. *Для любого слова Z существует оптимальная схема конкатенации, содержащая слова U, V, Z такие, что $Z = U \bullet V$, и слово V является либо подсловом слова U, либо символом, отсутствующим в слове U.*

Доказательство. Пусть S — некоторая оптимальная схема конкатенации слова Z, содержащая слова U, V, Z такие, что $Z = U \bullet V$, и слово V не является подсловом слова U. В этом случае схема $0, 1, X, Y, \ldots, U$ конкатенации слова U не содержит слова V, а схема S содержит слова V_1, V_2 такие, что $V = V_1 \bullet V_2$. Заменим в схеме S слово V на слово $U \bullet V_1$.



В результате получим такую схему конкатенации слова Z, что $Z = (U \bullet V_1) \bullet V_2$ и слово V_2 короче слова V. Полученная схема так же оптимальна, поскольку число слов в ней не изменилось (см. рисунок).

Если суффикс V_2 слова Z не удовлетворяет условиям леммы 1, то описанная выше процедура повторяется до тех пор, пока суффикс либо не окажется подсловом предшествующей ему части слова Z, либо не превратится в однобуквенное слово. Лемма 1 доказана.

Слово X называется *максимальным суффиксом* слова Z, если Z представимо в виде $Z = Y \bullet X$, где слово X является либо подсловом слова Y, либо символом, отсутствующим в слове Y, причем длина слова X не может быть увеличена.

Представление слова Z в виде $Z = Y_1 \bullet Y_2 \bullet \cdots \bullet Y_m$ назовем *суффиксным представлением*, если длина слова Y_1 равна единице, а всякое слово Y_i, $1 < i \leqslant m$, является максимальным суффиксом слова $Y_1 \bullet Y_2 \bullet \cdots \bullet Y_i$. Очевидно, что суффиксное представление любого слова единственно. Число операций конкатенации в суффиксном представлении слова Z назовем *суффиксной сложностью* слова Z и обозначим через $L^*(Z)$.

Лемма 2. *Для всякого слова Z выполняется неравенство $L(Z) \geqslant L^*(Z)$.*

Доказательство. Предположим, что лемма не верна. Из множества слов, для которых лемма 2 не верна, выберем слово с минимальной мультипликативной сложностью. Пусть это будет слово Z и $L(Z) < L^*(Z)$. Согласно лемме 1 среди оптимальных схем конкатенации слова Z имеется схема, содержащая слова U, V, Z такие, что $Z = U \bullet V$, и слово V является либо подсловом слова U, либо символом, отсутствующим в слове U. Пусть $Z = Y \bullet X$, где X — максимальный суффикс слова Z. Установим ряд неравенств.

1. Очевидно, что $L(Z) \geqslant L(U) + 1$.

2. Из минимальности выбранной по предположению схемы имеем $L(U) \geqslant L^*(U)$.

3. Слово Y является подсловом слова U и $L^*(U) \geqslant L^*(Y)$.

4. По определению суффиксного представления слова Z имеем $L^*(Y) = L^*(Z) - 1$.

Объединяя неравенства 1–4, получаем

$$L(Z) \geqslant L(U) + 1 \geqslant L^*(U) + 1 \geqslant L^*(Y) + 1 = L^*(Z),$$

что противоречит принятому предположению. Лемма 2 доказана.

Очевидно, что трудоемкость вычисления суффиксной сложности и, следовательно, получения нижней оценки мультипликативной сложности слова Z не превышает значения l^2 для слова длины l.

Заметим, что

1) аналогичный результат получается при замене в представлении слова Z суффиксов на префиксы;

2) обе леммы легко обобщаются на алфавит произвольной мощности.

Лемма 2 позволяет получать высокие нижние оценки мультипликативной сложности слов. Известно, что для любого $n > 1$ существует последовательность де Брейна длины $2^n + n - 1$, в которой каждое подслово длины n встречается один раз (см. [10, с. 128]). Следовательно, в ее суффиксном представлении длина каждого максимального суффикса не превосходит $n - 1$, т. е. суффиксная сложность этого слова, а по лемме 2 и мультипликативная сложность, не менее $2^n/n$. Таким образом, последовательности де Брейна имеют асимптотически наибольшую мультипликативную сложность [11].

§ 2. Мультипликативная сложность счетчика четности

Рассмотрим двоичные слова Z_k длины 2^k, совпадающие со столбцом значений булевой функции $x_1 \oplus x_2 \oplus \cdots \oplus x_k$, называемой счетчиком четности. Слова $Z_0 = 0$, $Z_1 = 01$, $Z_2 = 0110$, $Z_3 = 01101001, \ldots$ являются начальными отрезками бесконечной последовательности, называемой последовательностью Туэ — Морса (см., например, [12, с. 23]). Эта последовательность определяется индукцией по k. Пусть

1) $\varphi(0) = 01$, $\varphi(1) = 10$;

2) $Z_k = \varphi(Z_{k-1}) = \varphi^k(0) = Z_{k-1} \bullet \varphi^{k-2}(1) \bullet \varphi^{k-2}(0)$.

Из определения следует, что слово Z_k представимо в виде

$$Z_k = 0 \bullet 1 \bullet \varphi^0(1) \bullet \varphi^0(0) \bullet \varphi^1(1) \bullet \varphi^1(0) \bullet \ldots \bullet \varphi^{k-2}(1) \bullet \varphi^{k-2}(0) \qquad (1)$$

и существует схема конкатенации слова Z_k:

$$0, 1, \varphi^1(0), \varphi^1(1), \varphi^2(0), \varphi^2(1), \ldots, \varphi^{k-1}(0), \varphi^{k-1}(1), \varphi^k(0). \qquad (2)$$

Теорема. *При любом натуральном k*

$$L(Z_k) = 2k - 1.$$

Доказательство. Рассмотрим схему (2). Так как ее сложность равна $2k - 1$, то $L(Z_k) \leqslant 2k - 1$.

Докажем, что представление (1) слова Z_k суффиксное. Для этого достаточно показать, что каждый приведенный суффикс является максимальным. Предположим противное для некоторого суффикса $\varphi^i(\cdot)$, $0 \leqslant i \leqslant k - 2$. Сделав замены $01 \to 0$ и $10 \to 1$ в слове Z_k, мы легко убедимся в верности предположения и для некоторого вдвое более короткого суффикса. Продолжив «спуск», придем к суффиксу единичной длины, что и опровергает предположение. Следовательно, представление (1) слова Z_k суффиксное и $L^*(Z_k) = 2k - 1$. По лемме 2 имеем $L(Z_k) \geqslant 2k - 1$. Теорема доказана.

Автор выражает благодарность С. В. Августиновичу и А. А. Евдокимову за ряд полезных советов и участникам семинара «Символьные последовательности и языки с запретами» за обсуждение поставленной задачи.

ЛИТЕРАТУРА

1. **Лупанов О. Б.** Асимптотические оценки сложности управляющих систем. М.: Изд-во Моск. гос. ун-та, 1984.
2. **Гашков С. Б., Кочергин В. В.** Об аддитивных цепочках векторов, вентильных схемах и сложности вычисления степеней // Методы дискретного анализа в теории графов и сложности: Сб. науч. тр. Новосибирск: Ин-т математики СО РАН, 1992. Вып. 52. С. 22–40.
3. **Колотов А. Т.** Апериодические последовательности и функции роста алгебр // Алгебра и логика. 1981. Т. 20, № 2. С. 138–154.
4. **Евдокимов А. А.** Полные множества слов и их числовые характеристики // Методы дискретного анализа в исследовании экстремальных структур: Сб. науч. тр. Новосибирск: Ин-т математики СО АН СССР, 1983. Вып. 39. С. 7–19.
5. **Саломаа А.** Жемчужины теории формальных языков. М.: Мир, 1986.
6. **Morse M., Hedlund G.** Unending chess, symbolic dynamics, and a problem in semigroups // Duke Math. J. 1944. V. 11, N 1. P. 1–7.
7. **Cardot C.** Quelques résultats sur l'application de l'algèbre de Boole à la synthèse des circuits à relais // Annales des Télécommun. 1952. T. 7, N 2. P. 75–84.
8. **Храпченко В. М.** О сложности реализации линейной функции в классе π-схем // Мат. заметки. 1971. Т. 9, № 1. С. 35–40.
9. **Редькин Н. П.** Доказательство минимальности схем из функциональных элементов // Проблемы кибернетики: Сб. науч. тр. М.: Наука, 1970. Вып. 23. С. 83–101.
10. **Холл М.** Комбинаторика. М.: Мир, 1970.
11. **Strassen V.** Berechnungen in patriellen Algebren endlichen Typs // Computing. 1973. V. 11. P. 181–196.
12. **Lotaire M.** Combinatorics on words. Reading, Mass.: Addison-Wisley Publ. Co., 1983. (Encyclopedia of Mathematics and its Applications; V. 17).

Адрес автора:

Россия,
630090 Новосибирск,
Университетский пр., 4,
Институт математики
им. С. Л. Соболева СО РАН

Статья поступила
5 января 1996 г.