Дискретная математика

том 23 выпуск 4 * 2011

УДК 519.7

Нижние оценки сложности булевых схем конечной глубины с произвольными элементами

© 2011 г. Д. Ю. Черухин

В работе рассмотрены схемы из функциональных элементов конечной глубины, в которых в качестве элементов допускаются произвольные булевы функции от любого числа переменных. Предложен метод получения нелинейных нижних оценок сложности, применимый, в частности, к оператору циклической свертки. Полученные нижние оценки для схем глубины $d \geqslant 2$ имеют вид $\Omega(n\lambda_{d-1}(n))$. В частности, при $d = 2, 3, 4$ они имеют вид $\Omega(n^{3/2})$, $\Omega(n \log n)$ и $\Omega(n \log \log n)$ соответственно; при $d \geqslant 5$ функция $\lambda_{d-1}(n)$ имеет медленный рост. Указанные нижние оценки являются наибольшими из известных при всех четных d, а также при $d = 3$. При $d = 2, 3$ эти оценки были получены в предыдущих работах автора.

Работа посвящена проблематике получения высоких нижних оценок сложности для явно заданных функций. В настоящее время не существует методов, позволяющих доказывать нелинейные нижние оценки сложности для булевых схем из функциональных элементов без ограничений. Поэтому многие исследователи рассматривают схемы с ограничениями.

В данной работе рассматриваются классы схем конечной глубины, в которых в качестве элементов допускаются произвольные булевы функции от любого числа переменных. Эту модель можно охарактеризовать как модель средней силы, сопоставимую с формулами в полном базисе. Схемы ограниченной глубины представляют особый интерес в связи с редукцией Вэльянта [1], сводящей схемы логарифмической глубины к схемам глубины 2. Тем самым, из достаточно большой нижней оценки для схем глубины 2 следует нелинейная оценка для схем логарифмической глубины.

В рассматриваемой модели, для каждой фиксированной глубины d, известны нелинейные нижние оценки сложности. Первая из них (для глубины $d = 2$) была получена в [2]. В [3] были получены оценки для всех $d \geqslant 3$. До результатов автора наибольшие известные нижние оценки имели следующий вид: для $d = 2$ справедлива оценка $\Omega(n \log^2 n/(\log \log n))$ (см. [4]), заметим, что здесь и далее запись $f = \Omega(g)$ означает $g = O(f)$, запись $f = \Theta(g)$ означает, что $f = O(g)$ и $g = O(f)$, $\log n$ означает $\log_2 n$; для $d \geqslant 3$ справедлива оценка $\Omega(n\lambda_d(n))$ [3, 5, 6]. При этом $\lambda_1(n) = \Theta(n^{1/2})$, $\lambda_2(n) = \Theta(\log n)$, $\lambda_3(n) = \Theta(\log \log n)$; при $d \geqslant 4$ функция $\lambda_d(n)$ является медленно растущей, то есть растет медленнее любого конечного числа логарифмов от n.

В [7, 8, 9, 10] были получены оценки, не превосходящие названных выше, однако применимые к более простым или более значимым функциям: операторам сдвига и умножения матриц.

В [11] автором была доказана оценка $\Omega(n^{3/2})$ для схем глубины 2; в [12] — оценка $\Omega(n \log n)$ для схем глубины 3. В настоящей работе для любой глубины d получена оценка $\Omega(n\lambda_{d-1}(n))$, которая включает в себя предыдущие оценки в качестве частных случаев. Эта оценка превосходит ранее известные оценки при всех четных d, а также при $d = 3$.

Нижняя оценка сложности доказывается для оператора циклической свертки, для которого используемый метод дает наибольшие возможные оценки. В доказательстве схема любой конечной глубины фактически сводится к схеме глубины 2, для которой применяется метод доказательства, являющийся развитием метода из [11]. Для указанного сведения используется теоретико-графовая техника, разработанная в [3, 6, 9], а именно, непосредственно используется лемма из [9].

Напомним некоторые понятия, относящиеся к сетям и схемам (из функциональных элементов). Сетью назовем конечный ориентированный граф без ориентированных циклов. Вершина называется входом сети, если в нее не входит ни одного ребра; вершина называется выходом сети, если из нее не выходит ни одного ребра. Пусть имеется сеть с n входами и m выходами и пусть ее входам поставлены в соответствие переменные $x_1, \ldots, x_n$. Кроме того, пусть каждой вершине v, отличной от входов, поставлена в соответствие булева функция g_v, причем указано взаимно однозначное соответствие между ее аргументами и ребрами, входящими в вершину v. Наконец, пусть выходы сети упорядочены (обозначим их $z_1, \ldots, z_m$). Тогда сеть назовем схемой.

Глубиной сети называется наибольшая длина ориентированного пути в ней. Будем рассматривать классы схем, глубина которых не больше заданной константы d. Под сложностью схемы мы понимаем число ребер в ней. Поскольку мы оцениваем сложность с точностью до порядка, и в то же время глубина сети ограничена константой, то можно считать, что вершины сети разбиты на ярусы, занумерованные числами $0, 1, \ldots, d$. А именно, можно считать, что входы находятся на нулевом ярусе, выходы — на последнем ярусе (с номером d), и каждое ребро ведет (для некоторого i) из вершины i-го яруса в вершину $(i + 1)$-го яруса. Любую схему можно привести к указанному виду, вводя вспомогательные вершины; при этом сложность схемы увеличится не более, чем в d раз.

В каждой вершине v схемы вычисляется некоторая функция $f_v\colon \{0,1\}^n \to \{0,1\}$ (зависящая от входных переменных $x_1, \ldots, x_n$), которую несложно определить по индукции. Из функций, вычисляемых на выходах схемы, составим оператор $F\colon \{0,1\}^n \to \{0,1\}^m$, положив $F = (f_{z_1}, \ldots, f_{z_m})$. При этом скажем, что схема вычисляет оператор F.

Зададим оператор циклической свертки $H_n\colon \{0,1\}^{2n} \to \{0,1\}^n$, $H_n = (h_1, \ldots, h_n)$, зависящий от переменных $x_1, \ldots, x_n$, $y_1, \ldots, y_n$. А именно, для любого j положим

$$h_j(x_1, \ldots, x_n,\ y_1, \ldots, y_n) = x_1 y_j \oplus x_2 y_{j+1} \oplus \ldots \oplus x_n y_{j-1}.$$

Основной результат данной работы (теорема 1) состоит в том, что при $d \geqslant 2$ сложность любой схемы, которая имеет глубину d и вычисляет оператор H_n, есть $\Omega(n\lambda_{d-1}(n))$.

Пусть функции $f, f_1, \ldots, f_m$ зависят от одного и того же набора переменных $y = (y_1, \ldots, y_n)$. Скажем, что функция f выразима через функции $f_1, \ldots, f_m$, если существует такая булева функция Φ от m переменных, что выполнено тождество

$$f(y) \equiv \Phi(f_1(y), \ldots, f_m(y)).$$

Назовем слабой энтропией произвольный функционал $\mathbf{E}(\cdot)$, определенный на множествах булевых функций от переменных $(y_1, \ldots, y_n)$, принимающий действительные неотрицательные значения и обладающий следующими свойствами:

(а) для любой функции f выполнено неравенство $\mathbf{E}(\{f\}) \leqslant 1$;

(б) если $1 \leqslant i_1 < \ldots < i_k \leqslant n$, то

$$\mathbf{E}(\{y_{i_1}, \ldots, y_{i_k}\}) = k;$$

(в) если каждая функция из множества F выразима через множество G, то

$$\mathbf{E}(F) \leqslant \mathbf{E}(G);$$

(г) для любых множеств F и G выполнено неравенство

$$\mathbf{E}(F) + \mathbf{E}(G) \geqslant \mathbf{E}(F \cup G)$$

(свойство субаддитивности).

Приведем два примера слабой энтропии. Пусть даны функции $f_1, \ldots, f_m$. На множестве $\{0,1\}^n$ введем отношение эквивалентности: два набора эквивалентны, если любая из функций f_i принимает на них одинаковые значения. Пусть N — число классов эквивалентности. Тогда положим

$$\mathbf{E}(\{f_1, \ldots, f_m\}) = \log N. \tag{1}$$

Проверим справедливость условий (а)–(г). Поскольку функция принимает не более двух значений, то выполнено (а). Условие (б) выполнено, так как набор из k независимых переменных принимает все 2^k возможных значений. Далее, если все функции из множества F выразимы через функции из G, то при переходе от G к F эквивалентные наборы переходят в эквивалентные, а значит, число классов эквивалентности не увеличивается. Поэтому выполнено (в). Наконец, пусть множеству F соответствует N классов эквивалентности, а множеству G — M классов. Тогда множеству $F \cup G$ может соответствовать не более, чем NM классов, так как каждый класс множества F может разделиться не более, чем на M частей, соответствующих классам множества G. Поэтому выполнено (г).

Второй пример — энтропия Шеннона. Пусть $K_1, \ldots, K_N$ — все классы эквивалентности рассматриваемого отношения, $p_i = |K_i|/2^n$. Тогда положим

$$\mathbf{E}'(\{f_1, \ldots, f_m\}) = \sum_{i=1}^{N} p_i \log \frac{1}{p_i}. \tag{2}$$

Отметим, что энтропия (1) (в отличие от (2)) не удовлетворяет условию сильной субаддитивности

$$\mathbf{E}(F) + \mathbf{E}(G) \geqslant \mathbf{E}(F \cup G) + \mathbf{E}(F \cap G).$$

Этим объясняется введение термина слабая энтропия для такого рода функционалов.

Далее, будем считать, что фиксирован некоторый функционал слабой энтропии $\mathbf{E}(\cdot)$. Пусть дан булев оператор $F\colon \{0,1\}^{r+n} \to \{0,1\}^m$, $F = (f_1, \ldots, f_m)$, зависящий от переменных $x_1, \ldots, x_r, y_1, \ldots, y_n$. Положим $X = \{x_1, \ldots, x_r\}$. Далее, пусть $D \subseteq \{0,1\}^r$. Если $f\colon \{0,1\}^{r+n} \to \{0,1\}$ и $\alpha \in D$, то через f^α будем обозначать функцию $f^\alpha\colon \{0,1\}^n \to \{0,1\}$, полученную из f подстановкой набора констант α вместо соответствующих переменных из множества X.

Рассмотрим произвольное подмножество $T \subseteq X$, $T = \{x_{i_1}, \ldots, x_{i_k}\}$, $1 \leqslant i_1 < \ldots < i_k \leqslant r$. Проекцией $\alpha \in D$, $\alpha = (\alpha_1, \ldots, \alpha_r)$, на множество T назовем набор $(\alpha_{i_1}, \ldots, \alpha_{i_k})$, проекцию обозначим через $\mathrm{pr}_T(\alpha)$. Если $T = \varnothing$, то считаем, что проекция $\mathrm{pr}_T(\alpha)$ равна

пустому слову Λ. Положим $\mathrm{pr}_T(D) = \{\mathrm{pr}_T(\alpha) \mid \alpha \in D\}$. Другими словами, $\mathrm{pr}_T(D)$ — проекция множества D на T.

Рассмотрим произвольную схему из функциональных элементов S, вычисляющую оператор F. Для любой вершины v схемы S и любого подмножества $T \subseteq X$ обозначим через $T(v)$ множество всех переменных из T таких, что из соответствующих входов можно попасть в вершину v, двигаясь по ориентированным путям. Пусть в схеме S выделено множество вершин V, обладающее следующим свойством: каждый ориентированный путь, ведущий из некоторого входа в некоторый выход, содержит вершину из V. Другими словами, множество V разделяет входы и выходы.

Лемма 1. *Справедливо неравенство*

$$\sum_{v \in V} |\mathrm{pr}_{X(v)} D| \geqslant \mathbf{E}(\{f_j^\alpha \mid \alpha \in D,\ j = 1, \ldots, m\}). \tag{3}$$

Доказательство. Напомним, что через f_v обозначается функция, вычисляемая в вершине v. Так как множество V разделяет входы и выходы схемы, любая функция f_j, $j = 1, \ldots, m$, выражается через функции вида f_v, $v \in V$. Тогда для произвольного α функция f_j^α выражается через функции f_v^α, $v \in V$. Поэтому в силу свойства (в) слабой энтропии

$$\mathbf{E}(\{f_j^\alpha \mid \alpha \in D, j = 1, \ldots, m\}) \leqslant \mathbf{E}(\{f_v^\alpha \mid \alpha \in D, v \in V\}). \tag{4}$$

В силу свойств (г) и (а),

$$\mathbf{E}(\{f_v^\alpha \mid \alpha \in D,\ v \in V\}) \leqslant \sum_{v \in V} \mathbf{E}(\{f_v^\alpha \mid \alpha \in D\}) \leqslant \sum_{v \in V} |\{f_v^\alpha \mid \alpha \in D\}|. \tag{5}$$

Наконец, заметим, что функция f_v существенно зависит только от тех из переменных $x_1, \ldots, x_r$, которые входят в множество $X(v)$. Поэтому, если два набора α и β из D совпадают на множестве координат $X(v)$ (то есть, если $\mathrm{pr}_{X(v)}(\alpha) = \mathrm{pr}_{X(v)}(\beta)$), то $f_v^\alpha = f_v^\beta$. Следовательно, число различных функций среди функций вида f_v^α, $\alpha \in D$, не больше числа различных проекций наборов $\alpha \in D$ на множество $X(v)$. Последнее число есть мощность проекции множества D на $X(v)$. Таким образом,

$$|\{f_v^\alpha \mid \alpha \in D\}| \leqslant |\mathrm{pr}_{X(v)} D|. \tag{6}$$

Из (4)–(6) следует (3). Лемма доказана.

Выведем из леммы 1 следствие, относящееся к оператору циклической свертки. Пусть $r = n$, $F = H_n$, k, l — натуральные числа и $kl \leqslant n$. Для удобства будем считать, что имеется счетное множество переменных $x_1, x_2, \ldots$, причем переменные, индексы которых имеют один и тот же остаток при делении на n, отождествлены, то есть $x_i = x_{i+tn}$, для любых натуральных i, t. Аналогичные тождества введем для переменных $y_1, y_2, \ldots$ и выходных переменных $z_1, z_2, \ldots$ Введем обозначения

$$X_p^k = \{x_p, x_{p+1}, \ldots, x_{p+k-1}\}, \qquad Z_q^{k,l} = \{z_q, z_{q+k}, \ldots, z_{q+kl-k}\}.$$

Заметим, что если индекс i последовательно принимает значения $p, p+1, \ldots, p+k-1$, а индекс j (независимо от i) — значения $q, q+k, \ldots, q+kl-k$, то сумма $i+j$ при этом принимает kl последовательных натуральных значений. Учитывая тот факт, что $kl \leqslant n$, получим, что все переменные вида y_{i+j-1} различны.

Обозначим через $\alpha(i)$ единичный набор длины n (то есть набор из нулей и единиц, содержащий ровно одну единицу), в котором единица стоит на месте, соответствующем переменной x_i. Положим

$$D = \{\alpha(i) \mid x_i \in X_p^k\}.$$

Заметим, что $h_j^{\alpha(i)} = y_{i+j-1}$, поэтому множество $\{h_j^{\alpha(i)} \mid \alpha(i) \in D,\ z_j \in Z_q^{k,l}\}$ состоит из kl различных переменных, а значит, в силу свойства (б) слабой энтропии,

$$\mathbf{E}(\{h_j^{\alpha(i)} \mid \alpha(i) \in D,\ z_j \in Z_q^{k,l}\}) = kl. \tag{7}$$

Применив лемму 1 к подоператору $(h_q, h_{q+k}, \ldots, h_{q+kl-k})$, получим, что

$$\sum_{v \in V} |\mathrm{pr}_{X(v)}\, D| \geqslant \mathbf{E}(\{h_j^{\alpha(i)} \mid \alpha(i) \in D,\ z_j \in Z_q^{k,l}\}). \tag{8}$$

Наконец заметим, что множество $\mathrm{pr}_{X(v)}\, D$ состоит из нулевого набора (в частности, пустого слова) и тех единичных наборов, в которых единица соответствует переменной, одновременно принадлежащей множествам $X(v)$ и X_p^k. Поэтому

$$|\mathrm{pr}_{X(v)}\, D| = |X(v) \cap X_p^k| + 1 = |X_p^k(v)| + 1. \tag{9}$$

Используя (7)–(9), получим следующее утверждение.

Следствие 1. *Пусть множество вершин V удовлетворяет условию: любой ориентированный путь, ведущий из некоторого входа схемы в некоторый выход из множества $Z_q^{k,l}$, пересекается с V. Тогда*

$$\sum_{v \in V} (|X_p^k(v)| + 1) \geqslant kl.$$

Рассмотрим произвольную функцию f натурального аргумента, принимающую целые неотрицательные значения. Пусть для любого $n \geqslant 2$ справедливо неравенство $f(n) < n$. Через $f^{(k)}$ обозначим k-ю степень функции f относительно суперпозиции, то есть $f^{(k)} = f \circ f \circ \ldots \circ f$, где f повторяется k раз. Определим функцию f^* следующим образом:

$$f^*(n) = \min\{k \mid f^{(k)}(n) \leqslant 1\}.$$

Наконец, введем последовательность функций $\lambda_d(n)$:

$$\lambda_1(n) = \lfloor \sqrt{n} \rfloor, \quad \lambda_2(n) = \lceil \log n \rceil, \quad \lambda_d(n) = \lambda_{d-2}^*(n),\ d = 3, 4, \ldots$$

Определения чисел $\lambda_d(n)$ взяты из работы [9], там же содержатся недостающие доказательства. В следующем утверждении приводятся свойства функций $\lambda_d(n)$. Из него следует, что полученные в настоящей работе оценки являются нелинейными (п. 1). Кроме того, утверждение дает вид оценки для $d = 4$, а именно, $\Omega(\log\log n)$ (п. 2), а также показывает, что при четных $d \geqslant 4$ полученная оценка $\Omega(n\lambda_{d-1}(n))$ по порядку больше ранее известной оценки $\Omega(n\lambda_d(n))$ (п. 3).

Предложение 1. (1) *Для любого d функция $\lambda_d(n)$ возрастает по n и стремится к бесконечности при $n \to \infty$.*

(2) $\lambda_3(n) = \Theta(\log\log n)$.

(3) *Если d четно или $d = 3$, то*

$$\lambda_{d-1}(n) = \Omega(\lambda_d(n)),$$

в противном случае

$$\lambda_{d-1}(n) = \Theta(\lambda_d(n)).$$

Доказательство. Заметим, что если функция f возрастает, то и функция f^* возрастает. Кроме того, если функция f неограничена, то и f^* неограничена. Действительно, если f^* ограничена и принимает максимальное значение в точке n, то в силу неограниченности функции f существует такая точка m, что $f(m) > n$, а тогда $f^*(m) > f^*(n)$, получаем противоречие. Поэтому утверждение п. 1 следует из того, что функции $\lambda_1(n)$ и $\lambda_2(n)$ возрастают и неограничены.

Равенство п. 2 очевидно: при извлечении квадратного корня из числа его двойной логарифм уменьшается на единицу.

Для доказательства второй части п. 3 достаточно показать, что $\lambda_4(n) = \Theta(\lambda_5(n))$, а это следует из равенств

$$\lambda_3(n) = \Theta(\log\log n) = \Theta((\log n)^{(2)}) = \Theta(\lambda_2^{(2)}(n)).$$

Наконец, докажем первое утверждение п. 3. Оно очевидно при $d = 2, 3$. Пусть d — четное число и $d \geqslant 4$. Покажем, что если f — возрастающая неограниченная функция и $f(n) = o(n)$, то $f^*(n) = o(f(n))$. Действительно, так как $f(n) = o(n)$ и $f(n) \to \infty$ при $n \to \infty$, то $f^{(2)}(n) = o(f(n))$. В то же время $f^*(n) \leqslant f^{(2)}(n) + 1$, так как при каждом применении функции f значение уменьшается как минимум на единицу. Поэтому $f^*(n) = o(f(n))$. В частности, $\lambda_{d-1}(n) = \Theta(\lambda_{d-2}(n)) = \Omega(\lambda_d(n))$. Предложение доказано.

Следующая лемма (см. [9], лемма 1.1) позволяет сводить сеть конечной глубины, за исключением небольшого числа ее вершин, к сети глубины 1. Мы будем ее применять к части схемы, содержащей все ярусы ребер, кроме первого; при этом мы сведем схему к схеме глубины 2 (за исключением небольшого числа вершин) и затем применим лемму 1.

Лемма 2. *Если $0 < \varepsilon < 1/400$ и если сеть глубины d содержит больше, чем n вершин, и меньше, чем $\varepsilon n \lambda_d(n)$ ребер, то в множествах входов, выходов и всех вершин этой сети можно выделить такие подмножества I, J, W, соответственно, что*

a) $|I| \leqslant 5\varepsilon d n$, $|J| \leqslant 5\varepsilon d n$, $\sqrt{n} \leqslant |W| = o(n)$;

b) существует не более $\varepsilon n^2/|W|$ ориентированных путей, ведущих из входов сети к ее выходам и не содержащих вершин из множества $I \cup J \cup W$.

Отметим, что сходная теоретико-графовая техника ранее использовалась в работах [3, 6]. Заметим также, что лемма 2 верна, в частности, при $d = 1$, что можно непосредственно проверить.

Из лемм 1 и 2 следует основной результат работы.

Теорема 1. *Пусть $d \geqslant 2$. Тогда любая схема глубины d, вычисляющая оператор циклической свертки H_n, содержит $\Omega(n\lambda_{d-1}(n))$ ребер.*

Доказательство. Рассмотрим схему S минимальной сложности, вычисляющую оператор H_n. Пусть указанная сложность равна L. Выберем достаточно малое $\varepsilon > 0$ (которое мы определим в конце доказательства) и предположим, что $L < \varepsilon n \lambda_{d-1}(n)$. Рассмотрим сеть S_1, состоящую из всех ребер, неинцидентных входам схемы S (то есть ребер второго и последующих ярусов). Глубина сети S_1 равна $d-1$. Применим к S_1 лемму 2 и найдем соответствующие множества вершин I, J, W. Заметим, что вершины из множества I принадлежат первому ярусу схемы S. Положим $l = 4|W|$, $k = \lfloor n/l \rfloor$. Тогда $kl \leqslant n$ и (в силу ограничений на $|W|$) $kl \sim n$.

Воспользуемся ранее введенными обозначениями множеств X_p^k и $Z_q^{k,l}$. Для каждого входа x_i обозначим через $d(x_i)$ число ребер схемы S, исходящих из x_i. Среди всех возможных p выберем такое, для которого из множества X_p^k исходит наименьшее число ребер. Поскольку общее число ребер, исходящих из входов $x_1, \ldots, x_n$, не больше, чем L, то

$$\sum_{x_i \in X_p^k} d(x_i) \leqslant \frac{k}{n} L. \tag{10}$$

Обозначим через V_1 множество вершин первого яруса схемы S. Вспомним, что $X_p^k(v)$ есть множество тех входов из X_p^k, из которых можно попасть в вершину v, двигаясь по ориентированным путям. Если $v \in V_1$, то пути, входящие в v, состоят из одного ребра, а значит,

$$\sum_{x_i \in X_p^k} d(x_i) = \sum_{v \in V_1} |X_p^k(v)|. \tag{11}$$

Действительно, в правой и левой частях равенства (11) подсчитывается мощность одного и того же множества ребер, а именно, множества ребер, исходящих из множества X_p^k. Указанные ребра ведут, очевидно, в множество вершин V_1.

Далее выберем множество выходов $Z_q^{k,l}$. А именно, среди всех таких множеств выберем то, которое одновременно удовлетворяет двум условиям:

(i) $|Z_q^{k,l} \cap J| \leqslant (2l/n) 5\varepsilon d n$;

(ii) существует не более чем $(2l/n)\varepsilon n^2/|W|$ ориентированных путей, ведущих из множества вершин V_1 в множество $Z_q^{k,l}$ и не содержащих вершин из множества $I \cup J \cup W$.

Поскольку $|J| \leqslant 5\varepsilon d n$, то доля множеств $Z_q^{k,l}$, для которых условие (i) не выполнено, меньше, чем $1/2$. Аналогично, поскольку всего существует не более, чем $\varepsilon n^2/|W|$ указанных путей (ведущих к произвольным выходам), то доля множеств $Z_q^{k,l}$, не удовлетворяющих условию (ii), меньше, чем $1/2$. Следовательно, найдется множество, удовлетворяющее обоим условиям.

Обозначим через V_1' множество тех вершин из V_1, из которых можно попасть в вершины множества $Z_q^{k,l}$, двигаясь по ориентированным путям, не проходящим через вершины из $I \cup J \cup W$. Поскольку число путей не меньше числа вершин, в которых эти пути начинаются, то в силу условия (ii)

$$|V_1'| \leqslant 2\frac{l}{n}\varepsilon n^2/|W|. \tag{12}$$

Наконец, определим множество V, фигурирующее в следствии 1:

$$V = V_1' \cup I \cup (Z_q^{k,l} \cap J) \cup W.$$

Проверим, что это множество V удовлетворяет условию следствия 1. Пусть ориентированный путь Π начинается в некотором входе схемы S и оканчивается в вершине из множества $Z_q^{k,l}$. Если Π пересекается с множеством $I \cup J \cup W$, то он пересекается с множеством $I \cup (Z_q^{k,l} \cap J) \cup W$, входящим в V; действительно, J состоит только из выходов, а путь Π не может содержать двух разных выходов. Теперь предположим, что Π не пересекается с множеством $I \cup J \cup W$. В любом случае он пересекается с множеством V_1, а значит, с его подмножеством V_1', так как только из вершин этого подмножества можно попасть в множество $Z_q^{k,l}$, двигаясь по ориентированному пути, не проходящему через $I \cup J \cup W$. Итак, в любом случае путь Π пересекается с множеством V.

Применив следствие 1, получим, что

$$\sum_{v \in V} (|X_p^k(v)| + 1) \geqslant kl. \tag{13}$$

Далее, оценим левую часть неравенства (13). При этом будем использовать неравенства (10)–(12), условие (a) леммы 2, условие (i), неравенство $|X_p^k(v)| \leqslant k$ (выполненное для любой вершины v), а также неравенство $\lambda_{d-1}(n) \leqslant \sqrt{n}$. В результате получим, что

$$\begin{aligned}
\sum_{v \in V} (|X_p^k(v)| + 1) &\leqslant \sum_{v \in V_1' \cup I} |X_p^k(v)| + |V_1'| + |I| + \sum_{v \in (J \cap Z_q^{k,l}) \cup W} (|X_p^k(v)| + 1) \\
&\leqslant \sum_{v \in V_1} |X_p^k(v)| + |V_1'| + |I| + (k+1)(|J \cap Z_q^{k,l}| + |W|) \\
&\leqslant \frac{k}{n} L + \frac{2l}{n} \varepsilon \frac{n^2}{|W|} + 5\varepsilon d n + 2k\Big(\frac{2l}{n} 5\varepsilon d n + \frac{l}{4}\Big) \\
&\leqslant \varepsilon k \lambda_{d-1}(n) + 8\varepsilon n + 5\varepsilon d n + 20\varepsilon d k l + \frac{1}{2} 5kl \\
&\leqslant \varepsilon k \sqrt{n} + 35\varepsilon d n + n/2.
\end{aligned}$$

Поскольку $|W| \geqslant \sqrt{n}$, то $l \geqslant \sqrt{n}/4$, а тогда $k \leqslant 4\sqrt{n}(1 + o(1))$. Следовательно, при достаточно малом ε оцениваемая сумма будет асимптотически меньше, чем n, а в силу (13) она не меньше, чем $kl \sim n$. Получаем противоречие. Теорема доказана.

Отметим, что описанный в данной работе метод получения нижних оценок сложности булевых функций без труда обобщается на случай произвольной функциональной системы, обладающей функционалом слабой энтропии. При этом полученные нижние оценки переносятся, например, на оператор циклической свертки над произвольным полем. Следует учесть, что если поле бесконечно, то в качестве функциональных элементов можно допускать только алгебраические функции. Тогда в качестве функционала энтропии достаточно взять размерность множества линейных частей функций.

Отметим также, что описанный метод применим (и приводит к аналогичным нижним оценкам) к некоторым другим мультипликативным операторам, а именно, к операторам умножения двоичных чисел и умножения многочленов над $\mathbf{Z}_2$. При $d = 2$ нижняя оценка $\Omega(n^{3/2})$ справедлива также для оператора умножения матриц (см. замечание в работе [11] и полное доказательство в [10]), однако при $d \geqslant 3$ описанный выше метод, по-видимому, неприменим к умножению матриц.

Далее, заметим, что в данной работе лемма 1 применялась к небольшим множествам D, а именно, множествам мощности $O(r)$. Представляет интерес ее применение к множествам большей мощности, например, к множеству $D = \{0,1\}^r$.

Наконец, отметим, что для энтропии Шеннона справедливы неравенства, которые не выводятся из (а)–(г) и сильной субаддитивности. Первый пример такого неравенства был приведен в работе [13]. В этой связи представляют интерес методы получения нижних оценок, использующие энтропийные неравенства, выходящие за рамки (а)–(г).

Список литературы

1. Valiant L. G., Graph-theoretic arguments in low-level complexity. *Lecture Notes Computer Sci.* (1977) **53**, 162–176.
2. Pippenger N., Superconcentrators of depth 2. *J. Comput. Syst. Sci.* (1982) **24**, 82–90.
3. Dolev D., Dwork C., Pippenger N., Wigderson A., Superconcentrators, generalizers and generalized connectors with limited depth. In: *Proc. ACM STOC'83*. ACM Press, New York, pp. 42–51.
4. Radhakrishnan J., Ta-Shma A., Bounds for dispersers, extractors and depth-two superconcentrators. *SIAM J. Discrete Math.* (2000) **13**, 2–24.
5. Alon N., Pudlák P., Superconcentrators of depth 2 and 3: odd levels help (rarely). *J. Comput. Syst. Sci.* (1994) **48**, 194–202.
6. Pudlák P., Communication in bounded depth circuits. *Combinatorica* (1994) **14**, 203–216.
7. Pudlák P., Savický P., On shifting networks. *Theoret. Comput. Sci.* (1993) **116**, 415–419.
8. Pudlák P., Rödl V., Sgall J., Boolean circuits, tensor ranks and communication complexity. *SIAM J. Comput.* (1997) **26**, 605–633.
9. Raz R., Shpilka A., Lower bounds for matrix product in bounded depth circuits with arbitrary gates. *SIAM J. Comput.* (2003) **32**, 488–513.
10. Jukna S., Entropy of operators or why matrix multiplication is hard for depth-two circuits. *Theory Comput. Syst.* (2010) **46**, 301–310.
11. Cherukhin D. Yu., Complexity estimation from below in the class of schemes of depth 2 without restrictions on basis. *Moscow Univ. Math. Bull.* (2005) **60**, №4, 42–44.
12. Cherukhin D. Yu., Lower bounds for depth-2 and depth-3 Boolean circuits with arbitrary gates. *Lecture Notes Comput. Sci.* (2008) **5010**, 122–133.
13. Zhang Z., Yeung R. W., On characterization of entropy function via information inequalities. *IEEE Trans. Inform. Theory* (1998) **44**, 1440–1450.

Статья поступила 25.02.2009.