Нижние оценки сложности булевых схем конечной глубины с произвольными элементами

Д. Ю. Черухин

Механико-математический ф-т МГУ
Ленинские горы, Москва, 119992

Аннотация В работе рассмотрены схемы из функциональных элементов произвольной конечной глубины. Получены нелинейные нижние оценки сложности оператора циклической свёртки над произвольным полем. При этом, если поле K конечно, то в качестве функциональных элементов допускаются любые функции $K^n \to K$ от любого числа переменных (например, в случае поля $\mathbb{Z}_2$ допускаются любые булевы функции). Если же поле бесконечно, то в качестве элементов допускаются только многочлены.

Полученные нижние оценки для схем глубины $d \geqslant 2$ имеют вид $\Omega(n\lambda_{d-1}(n))$. В частности, для $d = 2, 3, 4$ 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. В случае булевых схем указанные нижние оценки являются наибольшими из известных для всех чётных d, а также для $d = 3$. При $d = 2, 3$ эти оценки были получены в предыдущих работах автора.

Key words: Булева функция, схема из функциональных элементов, сложность, глубина, нижняя оценка, циклическая свёртка

1 Введение

Работа относится к проблематике получения высоких нижних оценок сложности для явно заданных функций. Целью данной проблематики является исследование феномена сложности. Для этого предлагаются различные методы, позволяющие доказать, что та или иная функция сложна. Из этих методов делается заключение относительно того, что представляют из себя сложные функции, в чём "кроется" их сложность.

В настоящее время не существует методов, позволяющих доказывать нелинейные нижние оценки сложности для булевых схем из функциональных элементов без ограничений. Другими словами, нельзя доказать, что та или иная конкретная функция требует для своего вычисления нелинейное (относительно размера входных данных) число шагов. Поэтому многие исследователи ("нащупывая" подходы к этой проблеме) рассматривают схемы с ограничениями.

В зависимости от характера ограничений и доказанных для данной модели нижних оценок можно говорить о силе или слабости модели. В данной работе рассматриваются классы схем конечной глубины (т. е. имеется ограничение на глубину), 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. Напомним, что формулы также являются схемами с ограничениями; ограничение состоит в том, что выход каждого элемента может в дальнейшем использоваться только один раз.

В нашей модели, для каждой фиксированной глубины d, известны нелинейные нижние оценки сложности. Первая из них (для глубины $d = 2$) была получена в работе [2]. Затем, в работе [3] были рассмотрены схемы нечётной глубины, начиная с 5, в [6] — схемы глубины 3, в [5] — схемы чётной глубины, начиная с 4.

Дальнейшая работа в данной модели может состоять в получении как можно более высоких нижних оценок и расширении множества функций, для которых они применимы. Кроме того, интерес представляет связь схем ограниченной глубины со схемами без ограничений (или с меньшими ограничениями). К этому направлению относится известная редукция Вэльянта [1], сводящая 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, которые применяются на практике.

До наших результатов по данной теме наибольшие доказанные нижние оценки имели следующий вид: для $d = 2$ — $\Omega\left(\frac{n \log^2 n}{\log\log n}\right)$ [9], для $d \geqslant 3$ — $\Omega(n\lambda_d(n))$ [3, 6, 5]. При этом $\lambda_1(n) = \Theta(n^{1/2})$, $\lambda_2(n) = \Theta(\log n)$, $\lambda_3(n) = \Theta(\log\log n)$, а при $d \geqslant 4$ функция $\lambda_d(n)$ является медленно растущей, т. е. растёт медленнее любого конечного числа логарифмом от n. В работах [4, 7, 10, 14] были получены оценки, не превосходящие названных, однако применимые к более простым или более значимым функциям: операторам сдвига и умножения матриц.

В работе [11] нами была доказана оценка $\Omega(n^{3/2})$ для схем глубины 2, а в работе [13] — 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 наши предыдущие оценки. Эта оценка превосходит ранее известные при всех чётных d, а также при $d = 3$. В частности, новые оценки, полученные в данной работе, превосходят ранее известные при всех чётных $d \geqslant 4$. Мы доказываем нижнюю оценку для оператора циклического сдвига, для которого наш метод даёт наибольшие возможные оценки.

В нашем доказательстве схема любой конечной глубины фактически сводится к схеме глубины 2, для которой применяется метод доказательства, сходный с методом работы [11]. Этот метод в настоящей работе сформулирован в виде леммы 1, которая допускает применение к схемам любой глубины, однако в ней учитывается только один слой вершин, расположенных между входами и выходами, что фактически означает сведение к схемам глубины 2.

При таком неявном сведении нам необходимо оценить число рёбер нижнего яруса, т. е. число путей в исходной схеме, расположенных между выделенным слоем вершин и выходами. Здесь мы используем теоретико-графовую технику, разработанную в [3, 5, 10]; а именно, мы непосредственно ссылаемся на лемму из работы [10].

Наш метод для схем глубины 2 был впоследствии выражен в терминах энтропии функций в заметке [14]. Мы сохраняем общую энтропийную направленность метода, однако переводим её в плоскость аксиоматического задания функционала энтропии (как это было сделано в работе [12] применительно к другой задаче из области нижних оценок сложности). Такой подход позволяет нам одновременно доказывать оценки для разных функциональных систем: конечных функций и алгебраических функций над произвольным полем.

А именно, мы рассматриваем оператор циклической свёртки над произвольным полем. При этом, если поле конечно, то в схеме допускаются произвольные элементы с данной областью определения (например, для поля $\mathbb{Z}_2$ — произвольные булевы функции). Для бесконечного поля в качестве элементов допускаются только многочлены. При этом, ввиду билинейности реализуемого оператора, схема фактически становится билинейной, т. е. арифметической. Отметим, что в отличие от булевых схем, для арифметических схем известны гораздо большие нижние оценки сложности (см., например, [10]). Мы не претендуем на прогресс в данной области, а приводим соответствующие результаты в качестве расширения границ применимости нашего метода.

2 Функциональные системы и схемы

Пусть K — множество (возможно, снабжённое алгебраической структурой). Рассмотрим множество $\mathcal{F}$, состоящее из некоторых функций вида $K^n \to K$, где n может принимать целые неорицательные значения. Множество $\mathcal{F}$ называется *функциональной системой*, если оно замкнуто относительно суперпозиции, перестановки и отождествления переменных. Другими словами, если функции $f(x_1, \dots, x_n)$ и $g(x_1, \dots, x_k)$ принадлежат множеству $\mathcal{F}$, то для любых индексов $i_1, \dots, i_{n+k-1}$ функция

$$f(g(x_{i_1}, \dots, x_{i_k}), x_{i_{k+1}}, \dots, x_{i_{n+k-1}})$$

принадлежит $\mathcal{F}$.

Мы будем рассматривать следующие функциональные системы:

(I) K — конечное множество, $|K| \geqslant 2$ и $\mathcal{F}$ состоит из всевозможных функций вида $K^n \to K$;

(II) K — поле и $\mathcal{F}$ состоит из всех линейных функций над K;

(III) K — бесконечное поле и $\mathcal{F}$ состоит из всех многочленов (многих переменных) над K.

Заметим, что если поле K бесконечно и два многочлена над K совпадают численно (т. е. совпадают их значения в каждой точке), то эти многочлены совпадают также и синтаксически (т. е. совпадают их коэффициенты). Для конечного поля это неверно. Например, в поле $\mathbb{Z}_2$ многочлены x и x^2 совпадают численно, но не синтаксически. Поэтому в функциональную систему (III) мы включили требование бесконечности поля.

Введём понятие *схемы* (*из функциональных элементов*) над системой $\mathcal{F}$. *Сетью* назовём конечный ориентированный граф без ориентированных циклов. Вершина называется *входом* сети, если в неё не входит ни одного ребра; вершина называется *выходом* сети, если из неё не выходит ни одного ребра. Пусть имеется сеть с n входами и m выходами. Пусть её входам поставлены в соответствие переменные $x_1, \ldots, x_n$. Кроме того, пусть каждой вершине v, отличной от входов, поставлена в соответствие некоторая функция g_v из системы $\mathcal{F}$, при этом указано взаимнооднозначное соответствие между аргументами функции g_v и рёбрами, входящими в вершину v. Наконец, пусть выходы сети упорядочены (обозначим их $z_1, \ldots, z_m$). Тогда сеть называется схемой над $\mathcal{F}$.

Обычно на схему накладываются те или иные ограничения. Одним из них является константное ограничение на число аргументов у функций g_v (схемы такого типа являются сильной моделью вычислений). Другим ограничением, которое мы будем использовать в данной работе, является ограничение на глубину сети. Глубиной сети называется наибольшая длина ориентированного пути в сети. Мы будем рассматривать классы схем, глубина которых не больше заданной константы d. Под сложностью схемы мы понимаем число рёбер в ней.

Поскольку мы оцениваем сложность с точностью до порядка, а глубина сети ограничена константой, то можем считать, что вершины сети разбиты на ярусы. Ярусы сети мы будем нумеровать числами $0, 1, \ldots, d$, при этом можно считать, что входы находятся на 0-м ярусе, выходы — на последнем ярусе (с номером d) и каждое ребро ведёт (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 $f_v\colon K^n \to K$, зависящая от входных переменных $x_1, \ldots, x_n$, которую несложно определить по индукции. Заметим, что функция f_v выражается в виде суперпозиции функций вида $g_{v'}$, а поскольку последние принадлежат системе $\mathcal{F}$, то и функция f_v принадлежит $\mathcal{F}$. Из функций, вычисляемых на выходах схемы, составим оператор $F\colon K^n \to K^m$, а именно, $F = (f_{z_1}, \ldots, f_{z_m})$. Мы говорим, что схема *вычисляет* оператор F.

Введём оператор *циклической свёртки* H_n, для которого мы будем доказывать нижние оценки. Пусть K — произвольное поле. Если K конечно, то в качестве $\mathcal{F}$ возьмём функциональную систему (I), иначе — в качестве $\mathcal{F}$ возьмём систему (III). Будем считать, что оператор $H_n\colon K^{2n} \to K^n$ зависит от переменных $x_1, \ldots, x_n,\ y_1, \ldots, y_n$. Положим $H_n = (h_1, \ldots, h_n)$, где

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

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 $\Omega(n\lambda_{d-1}(n))$; функция $\lambda_d(n)$ вводится в разделе 5.

3 Выразимость функций и энтропия

Пусть функции $f, f_1, \ldots, f_m$ принадлежат функциональной системе $\mathcal{F}$ и зависят от одного и того же набора переменных $\tilde{y} = (y_1, \ldots, y_n)$. Скажем, что функция f *выразима* через функции $f_1, \ldots, f_m$, если существует такая функция $\Phi \in \mathcal{F}$, зависящая от m переменных, что выполнено тождество

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

Введём понятие *энтропии* множества функций $f_1, \ldots, f_m$, зависящих от набора переменных $\tilde{y}$; энтропию будем обозначать в виде $\mathcal{E}(\{f_1, \ldots, f_m\})$, а также в виде $\mathcal{E}(f_1, \ldots, f_m)$. Энтропией назовём любой функционал, который определён на всех конечных множествах функций вида $\{f_1, \ldots, f_m\}$, принимает действительные неотрицательные значения и удовлетворяет следующим условиям:

(a) энтропия множества, состоящего из одной функции, не больше 1;

(b) энтропия множества, состоящего из k различных переменных (взятых из набора $y_1, \ldots, y_n$), равна k;

(c) если каждая функция из множества F выразима через множество G, то $\mathcal{E}(F) \leqslant \mathcal{E}(G)$;

(d) субаддитивность: для любых множеств F и G

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

Приведём примеры энтропии для рассматриваемых нами функциональных систем.

Начнём с системы (II). В качестве энтропии множества линейных функций естественно рассматривать его ранг над полем K. Условия (a)-(d), очевидно, выполнены.

Далее, система (III). Энтропию наиболее естественно определить как размерность алгебраического многообразия, порождённого множеством многочленов. Однако для наших целей подойдёт следующее, более простое определение. А именно, у каждого многочлена мы выделим линейную часть, и энтропию множества многочленов положим равной энтропии множества их линейных частей, которая определена в предыдущем примере. Условия (a), (b) и (d) следуют из предыдущего примера.

Условие (c) выполнено в силу бесконечности поля K. Действительно, при линейных операциях с многочленами, с их линейными частями происходят те же линейные операции. Осталось заметить, что линейная часть произведения многочленов также выражается в виде линейной комбинации линейных частей множителей. Здесь существенную роль играет то, что если два многочлена равны тождественно, то равны их коэффициенты. Значит, линейные слагаемые не могут получиться из нелинейных при дальнейшем преобразовании, следовательно, нелинейные можно изначально отбросить. Итак, если многочлен f выражается через многочлены $f_1, \dots, f_m$ с помощью многочлена Φ (который можно свести к последовательности линейных операций и умножения), то и линейная часть многочлена f выражается через линейные части многочленов $f_1, \dots, f_m$. Тогда условие (c) следует из предыдущего примера.

Наконец, перейдём к наиболее важной для нас системе (I). Пусть даны функции $f_1, \dots, f_m$. На множестве K^n введём следующее отношение эквивалентности: два набора из K^n эквивалентны, если любая из функций f_i принимает на них одинаковые значения. Пусть N — число классов эквивалентности. Тогда положим

$$\mathcal{E}(f_1, \dots, f_m) = \log_{|K|} N. \tag{1}$$

Поскольку функция принимает не более, чем $|K|$ значений, то выполнено (a). Условие (b) выполнено, так как набор из k независимых переменных принимает все $|K|^k$ 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 (c). 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 (d).

Отметим, что энтропия в виде (1), применительно к данной задаче получения нижних оценок сложности, была введена в работе [14]. Мы также использовали понятие энтропии (называя её количеством информации) для получения других нижних оценок сложности в работе [12]. Однако при этом мы использовали иное (более сложное, но в то же время, более эффективное) определение энтропии, а именно, энтропию по Шеннону. Продолжим предыдущий пример. Пусть $K_1, \dots, K_N$ — все классы эквивалентности у рассматриваемого отношения, $p_i = |K_i|/|K|^n$. Тогда положим

$$\mathcal{E}'(f_1, \dots, f_m) = \sum_{i=1}^{N} p_i \log_{|K|} \frac{1}{p_i}.$$

Энтропия $\mathcal{E}'$ более эффективна, так как для неё (в отличие от $\mathcal{E}$) выполнено усиление свойства (d), а именно, сильная субаддитивность:

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

Отметим, что сильная субаддитивность выполнена также в первых двух примерах, относящихся к функциональным системам (III) и (II). Это свойство использовалось автором в работе [12]; там же содержится доказательство свойств (a)-(d) и сильной субаддитивности. Впрочем, доказательство следует из классической теории информации, если функцию вида $K^n \to K$

рассматривать как дискретную случайную величину, определённую на множестве элементарных событий K^n.

В данной работе мы ограничимся использованием обычной субаддитивности, поэтому нам достаточно определения (1). В качестве комментария общего плана заметим, что существуют неравенства для энтропии Шеннона, которые не выводятся из сильной субаддитивности. Первый пример такого неравенства был приведён в работе [8].

4 Основная лемма

Приведённая в данном разделе лемма содержит ключевой момент в доказательстве основного результата работы. Подобные соображения были впервые применены автором в работе [11], затем они слегка улучшались в работах [13, 14]. Мы приводим лемму в ещё раз слегка улучшенной формулировке.

Пусть функциональная система $\mathcal{F}$ содержит константы 0 и 1. Рассмотрим схему над системой $\mathcal{F}$, которая вычисляет некоторый оператор $F\colon K^{2n} \to K^n$, зависящий от переменных $x_1, \ldots, x_n,\ y_1, \ldots, y_n$, причём $F = (f_1, \ldots, f_n)$. Обозначим через f_j^i функцию, полученную из f_j подстановкой констант вместо переменных $x_1, \ldots, x_n$; а именно, вместо x_i подставляется 1, а вместо остальных переменных — нули.

Пусть I — множество, состоящее из некоторых входов схемы из набора $x_1, \ldots, x_n$, J — множество, состоящее из некоторых выходов схемы (напомним, выходы мы обозначаем через $z_1, \ldots, z_n$). Для каждой вершины v обозначим через $I(v)$ множество тех входов из I, из которых можно попасть в вершину v по ориентированных путям.

Lemma 1. *Пусть множество вершин V удовлетворяет условию: любой ориентированный путь, ведущий из некоторого входа схемы в некоторый выход из множества J, пересекается с V. Тогда для любого функционала энтропии $\mathcal{E}(\cdot)$ выполнено*

$$\sum_{v \in V} (|I(v)| + 1) \geqslant \mathcal{E}(\{f_j^i \mid x_i \in I,\ z_j \in J\}).$$

Proof. Напомним, что через f_v обозначается функция, вычисляемая в вершине v; она зависит от переменных $x_1, \ldots, x_n,\ y_1, \ldots, y_n$. Обозначим через f_v^i функцию, полученную из f_v подстановкой единицы вместо x_i и нулей вместо остальных переменных $x_1, \ldots, x_n$; через f_v^0 обозначим функцию, полученную из f_v подстановкой нулей вместо всех переменных $x_1, \ldots, x_n$. Так как все пути, ведущие от входов схемы к выходам из J пересекаются с множеством V, то любая функция $f_j,\ z_j \in J$, выражается через функции вида $f_v,\ v \in V$. Тогда для произвольного i функция f_j^i выражается через функции $f_v^i,\ v \in V$. Поэтому в силу свойства (c) энтропии

$$\mathcal{E}(\{f_j^i \mid x_i \in I,\ z_j \in J\}) \leqslant \mathcal{E}(\{f_v^i \mid x_i \in I,\ v \in V\}). \tag{2}$$

В силу свойств (d) и (a) имеем

$$\mathcal{E}(\{f_v^i \mid x_i \in I,\ v \in V\}) \leqslant \sum_{v\in V} \mathcal{E}(\{f_v^i \mid x_i \in I\}) \leqslant \sum_{v\in V} |\{f_v^i \mid x_i \in I\}|. \qquad (3)$$

Наконец, заметим, что если $x_i \notin I(v)$, то функция f_v не зависит от переменной x_i, а значит, $f_v^i = f_v^0$. Следовательно,

$$\{f_v^i \mid x_i \in I\} = \{f_v^i \mid x_i \in I(v)\} \cup \{f_v^0\},$$

а значит,

$$|\{f_v^i \mid x_i \in I\}| \leqslant |I(v)| + 1. \qquad (4)$$

Лемма следует из выкладок (2)–(4). □

Теперь выведем интересующее нас следствие для оператора циклической свёртки. Рассмотрим поле K, соответствующую ему функциональную систему $\mathcal{F}$, и схему над $\mathcal{F}$, вычисляющую оператор циклической свёртки $H_n = (h_1, \dots, h_n)$. Пусть $kl \leqslant n$. Обозначим через I_p^k множество, состоящее из k подряд идущих входов (из набора $x_1, \dots, x_n$), начиная со входа x_p, т. е. $I_p^k = \{x_p, x_{p+1}, \dots, x_{p+k-1}\}$. При этом мы считаем, что после входа x_n идёт вход x_1 (циклический порядок). Через $J_q^{k,l}$ обозначим множество, состоящее из l выходов, идущих друг за другом с шагом k (с соблюдением циклического порядка), начиная с выхода z_q, т. е. $J_q^{k,l} = \{z_q, z_{q+k}, \dots, z_{q+kl-k}\}$.

Если индекс i пробегает числа $p, p+1, \dots, p+k-1$, а индекс j — числа $q, q+k, \dots, q+kl-k$, то сумма $i+j$ при этом пробегает kl последовательных чисел. Заметим, что $h_j^i = y_{i+j-1}$. Поэтому множество $\{h_j^i \mid x_i \in I_p^k,\ z_j \in J_q^{k,l}\}$ состоит из kl различных переменных, а значит, в силу свойства (b) энтропии

$$\mathcal{E}(\{h_j^i \mid x_i \in I_p^k,\ z_j \in J_q^{k,l}\}) = kl.$$

Комбинируя это равенство с леммой, получим требуемое следствие:

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

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

5 Медленно растущие функции и графовая техника

Материал данного раздела, включая определения и формулировки результатов, взят нами из работы [10]; там же содержатся недостающие доказательства. Отметим, что сходная техника ранее использовалась в работах [3, 5].

Пусть f — функция натурального аргумента, принимающая целые неотрицательные значения, и для любого $n \geqslant 2$ выполнено $f(n) < n$. Через

$f^{(k)}$ обозначим k-ю степень функции f относительно суперпозиции, т. е. $f^{(k)} = f \circ f \circ \ldots \circ f$, где f повторяется k раз. Определим функцию f^* следующим образом:

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

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

$$\lambda_1(n) = \lfloor\sqrt{n}\rfloor, \quad \lambda_2(n) = \lceil\log_2 n\rceil, \quad \lambda_d(n) = \lambda^*_{d-2}(n),\ d = 3, 4, \ldots$$

В следующем утверждении приводятся свойства функций $\lambda_d(n)$. Из него следует, что полученные нами оценки являются нелинейными (см. п. 1). Кроме того, утверждение даёт точный вид оценки для $d = 4$, а именно, $\Omega(n \log_2 \log_2 n)$ (см. п. 2), а также показывает, что при чётных $d \geqslant 4$ полученная нами оценка $\Omega(n \log_{d-1}(n))$ по порядку больше ранее известной $\Omega(n \log_d(n))$ (см. п. 3).

Claim. 1) для любого d функция $\lambda_d(n)$ возрастает по n и стремится к бесконечности при $n \to \infty$;

2) $\lambda_3(n) = \Theta(n \log_2 \log_2 n)$;

3) если d чётно или $d = 3$, то $\lambda_{d-1}(n) = \Omega(\lambda_d(n))$, в противном случае, $\lambda_{d-1}(n) = \Theta(\lambda_d(n))$.

Proof. 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 $f^*(m) > f^*(n)$; противоречие. Поэтому п. 1) следует из того, что функции $\lambda_1(n)$ и $\lambda_2(n)$ возрастают и неограничены.

П. 2) очевиден: при извлечении квадратного корня из числа его двойной логарифм уменьшается на единицу. Для доказательства второго утверждения п. 3) достаточно доказать, что $\lambda_4(n) = \Theta(\lambda_5(n))$, а это следует из того, что $\lambda_3(n) = \Theta(\log_2 \log_2 n) = \Theta((\log_2 n)^{(2)}) = \Theta(\lambda_2^{(2)}(n))$.

Наконец, докажем первое утверждение п. 3). Оно очевидно при $d = 2, 3$. Пусть d — чётное 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, $\lambda_{d-1}(n) = \Theta(\lambda_{d-2}(n)) = \Omega(\lambda_d(n))$. $\square$

Следующая лемма позволяет сводить сеть конечной глубины, за исключением небольшого числа её вершин, к сети глубины 1. Мы будем её применять к части схемы, содержащей все ярусы рёбер, кроме первого; при этом мы сведём схему к схеме глубины 2 (за исключением небольшого числа вершин) и затем применим основную лемму.

Lemma 2. (*[10], Lemma* 1.1) *Если* $0 < \varepsilon < 1/400$ *и если сеть глубины* d *содержит больше, чем* n *вершин и меньше, чем* $\varepsilon n \lambda_d(n)$ *рёбер, то в множествах входов, выходов и всех вершин этой сети можно выделить такие подмножества* I, J, W *соответственно, что:*

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

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

6 Основная теорема

Theorem 1. *Пусть* $d \geqslant 2$, K *— произвольное поле и* $\mathcal{F}$ *— соответствующая функциональная система вида (I) или (III). Тогда любая схема глубины* d *над системой* $\mathcal{F}$*, вычисляющая оператор циклической сврткки* H_n*, содержит* $\Omega(n\lambda_{d-1}(n))$ *рёбер.*

Proof. Пусть число рёбер в сети равно L. Выберем достаточно малое ε (какое, будет ясно в конце доказательства) и предположим, что $L < \varepsilon n\lambda_{d-1}(n)$. Рассмотрим сеть S, состоящую из всех рёбер, не инцидентных входам схемы (т. е. рёбер второго и последующих ярусов). Глубина сети S равна $d-1$. Применим к ней лемму 2 и найдём соответствующие множества вершин I, J, W. Заметим, что вершины из I лежат на первом ярусе исходной сети. Положим $l = 4|W|$, $k = \lfloor n/l \rfloor$. Тогда $kl \leqslant n$ и (в силу ограничений на $|W|$) $kl \sim n$.

Для каждого входа x_i обозначим через $d(x_i)$ число рёбер, исходящих из x_i. Среди всех возможных p выберем такое, для которого из множества I_p^k (см. раздел 4) исходит наименьшее число рёбер. Поскольку общее число рёбер, исходящих из $x_1, \ldots, x_n$ не больше, чем L, то

$$\sum_{x_i \in I_p^k} d(x_i) \leqslant \frac{k}{n} L. \tag{5}$$

Обозначим через V_1 множество вершин первого яруса исходной сети. Вспомним, что $I_p^k(v)$ есть множество тех входов из I_p^k, из которых можно прийти в вершину v по ориентированным путям. Если v лежит на первом ярусе, то пути, входящие в v, состоят из одного ребра, а значит,

$$\sum_{x_i \in I_p^k} d(x_i) = \sum_{v \in V_1} |I_p^k(v)|. \tag{6}$$

Действительно, справа и слева подсчитывается мощность одного и того же множества рёбер, а именно, рёбер, исходящих из множества I_p^k; эти рёбра ведут, очевидно, в множество вершин V_1 первого яруса.

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

a') $|J_q^{k,l} \cap J| \leqslant 2\frac{l}{n} \cdot 5\varepsilon dn$;

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

Поскольку $|J| \leqslant 5\varepsilon dn$, то доля множеств $J_q^{k,l}$, для которых условие a') не выполнено, меньше, чем $1/2$. Аналогично, поскольку всего существует не более, чем $\varepsilon n^2/|W|$ таких путей (ведущих к произвольным выходам), то доля множеств $J_q^{k,l}$, не удовлетворяющих условию b'), меньше, чем $1/2$. Значит, найдётся множество, удовлетворяющее обоим условиям.

Обозначим через V_1' множество тех вершин первого слоя, из которых можно попасть в множество $J_q^{k,l}$ по ориентированным путям, не проходящим через $I \cup J \cup W$. Поскольку число путей не меньше числа вершин, в которых эти пути начинаются, то в силу выбора множества $J_q^{k,l}$, получим

$$|V_1'| \leqslant 2\frac{l}{n} \cdot \varepsilon n^2/|W|. \tag{7}$$

Наконец, определим множество V, фигурирующее в следствии 1:

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

Проверим, что это множество удовлетворяет условию следствия 1. Пусть некоторый путь ведёт из входа схемы в множество выходов $J_q^{k,l}$. Если он пересекается с множеством $I \cup J \cup W$, то он пересекается с множеством $I \cup (J_q^{k,l} \cap J) \cup W$, входящим в V; действительно, J состоит только из выходов, а путь не может содержать двух разных выходов. Если путь не пересекается с множеством $I \cup J \cup W$, то он проходит через множество V_1 (вершин первого яруса), а значит, он проходит через множество V_1', так как только из этого множества можно попасть в $J_q^{k,l}$. Итак, в любом случае наш путь пересекается с множеством V.

Применив следствие 1, получим

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

Далее оценим левую часть неравенства, используя выкладки (5)–(7), условие a) леммы 2 и условие a'). Мы будем использовать также, что $|I_p^k(v)| \leqslant k$ для любой вершины v. Получим

$$\begin{aligned}
&\sum_{v \in V} (|I_p^k(v)| + 1) \leqslant \\
&\sum_{v \in V_1' \cup I} |I_p^k(v)| + |V_1'| + |I| + \sum_{v \in (J \cap J_q^{k,l}) \cup W} (|I_p^k(v)| + 1) \leqslant \\
&\sum_{v \in V_1} |I_p^k(v)| + |V_1'| + |I| + (k+1)(|J \cap J_q^{k,l}| + |W|) \leqslant \\
&\frac{k}{n}L + 2\frac{l}{n} \cdot \varepsilon \frac{n^2}{|W|} + 5\varepsilon dn + 2k\Big(2\frac{l}{n} \cdot 5\varepsilon dn + \frac{l}{4}\Big) \leqslant \\
&\varepsilon k \lambda_{d-1}(n) + 8\varepsilon n + 5\varepsilon dn + 20\varepsilon dkl + 0,5kl \leqslant \\
&\varepsilon k\sqrt{n} + 35\varepsilon dn.
\end{aligned}$$

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 она должна быть не меньше, чем $kl \sim n$. Получаем противоречие. Теорема доказана. $\square$

Благодарности Работа переведена на английский язык [автор перевода]. Работа выполнена при частичной финансовой поддержке Программы поддержки ведущих научных школ РФФИ (проект НШ-5400.2006.1).

Литература

1. Valiant L. G.: Graph-theoretic arguments in low-level complexity. LNCS **53** (1977) 162–176
2. Pippenger N.: Superconcentrators of depth 2. J. of Computer and System Sciences **24** (1982) 82–90
3. Dolev D., Dwork C., Pippenger N., Wigderson A.: Superconcentrators, generalizers and generalized connectors with limited depth. Proc. 15th ACM STOC (1983) 42–51
4. Pudlák P., Savický P.: On shifting networks. Theoretical Comput. Sci. **116** (1993) 415–419
5. Pudlák P.: Communication in Bounded Depth Circuits. Combinatorica **14:2** (1994) 203–216
6. Alon N., Pudlák P.: Superconcentrators of depth 2 and 3; odd levels help (rarely). J. of Computer and System Sciences **48** (1994) 194–202
7. Pudlák P., Rödl V., Sgall J.: Boolean circuits, tensor ranks and communication complexity. SIAM J. on Computing **26:3** (1997) 605–633
8. Zhang Z., Yeung R. W.: On Characterization of entropy function via information inequalities. IEEE Transactions on Information Theory. **44** (1998) 1440–1450
9. Radhakrishnan J., Ta-Shma A.: Bounds for dispersers, extractors and depth-two superconcentrators. SIAM J. of Discrete Mathematics **13:1** (2000) 2–24
10. Raz R., Shpilka A.: Lower Bounds for Matrix Product, in Bounded Depth Circuits with Arbitrary Gates. SIAM J. Comput. **32:2** (2003) 488–513
11. Cherukhin D. Yu.: The lower estimate of complexity in the class of schemes of depth 2 without restrictions on a basis. Moscow Univ. Math. Bull. **60:4** (2005) 42–44
12. Cherukhin D. Yu.: On complexity of informational networks. Manuscript, available at http://cherukhin.narod.ru/research/research.htm (2008), in Russian
13. Cherukhin D. Yu.: Lower Bounds for Depth-2 and Depth-3 Boolean Circuits with Arbitrary Gates. Proc. 3rd Int. Comput Sci Symposium in Russia (CSR-2008), to appear.
14. Jukna S.: Entropy of operators or why matrix multiplication is hard for small depth circuits. Electronic Colloquium on Computational Complexity, Report Nr. TR08-019, 2008