Нижние оценки сложности для булевых схем глубины 2 и 3 с произвольными элементами

Д. Ю. Черухин

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

Аннотация В работе рассмотрены схемы из функциональных элементов конечной глубины в базисе, состоящем из всех булевых функций. Для схем глубины 3 доказана нижняя оценка сложности вида $\Omega(n \log n)$. Ранее автором была получена нижняя оценка вида $\Omega(n^{3/2})$ для схем глубины 2; в настоящей работе приводится усовершенствованное доказательство этой оценки. Обе упомянутые оценки являются наибольшими из известных для схем своей глубины и применимы, например, к оператору циклической свёртки.

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

1 Введение

Проблема доказательства высоких нижних оценок сложности для индивидуальных, явно заданных булевых функций и операторов является одной из наиболее острых задач современности, стоящих на стыке математики и информатики. Значительный прогресс в данной области был достигнут только в слабых моделях вычислений (имеющих существенные вычислительные ограничения), таких, как монотонные схемы и схемы ограниченной глубины в специальных базисах. В этих классах схем были получены экспоненциальные нижние оценки сложности.

В то же время в традиционных моделях вычислений, имеющих структурные (но не вычислительные) ограничения, таких, как формулы в полных базисах, контактные и контактно-вентильные схемы, были доказаны не более, чем полиномиальные нижние оценки сложности. Классы схем такого типа можно отнести к средним по силе. Наконец, в модели, приближенной к реальным вычислениям, а именно, схемах из функциональных элементов в полном конечном базисе, известны только линейные нижние оценки. Последнюю модель можно классифицировать как сильную.

Модель, рассматриваемая в данной работе, а именно, схемы ограниченной глубины в произвольном базисе, может быть по вышеописанным критериям отнесена к средним по силе. Далее мы будем говорить исключительно об этой модели. Напомним, что под сложностью в данной модели понимается число рёбер.

В классе схем произвольной конечной глубины известны нелинейные нижние оценки сложности. Для схем глубины 2 наибольшая из известных нижних оценок имеет порядок $n^{3/2}$. Она была получена автором в работе [7]. Для схем большей глубины все известные нижние оценки являются почти линейными. Функцию мы называем почти линейной, если она отличается от линейной медленно растущим множителем. Для схем глубины 3 этот множитель имеет логарифмический вид (см. ниже), а для схем большей глубины — сверхлогарифмический.

В настоящей работе мы приводим слегка усовершенствованное доказательство нижней оценки из работы [7], а также доказываем новую оценку для схем глубины 3. Новая оценка является наибольшей из известных автору для схем глубины 3. А именно, ранее была известна оценка порядка $n \log \log n$, а в настоящей работе доказана оценка порядка $n \log n$. При доказательстве новой оценки для схем глубины 3 используется редукция к схемам глубины 2, для которых применяется метод доказательства нижней оценки, сходный с методом работы [7].

Следует отметить, что в рассматриваемой модели вычислений любая булева функция имеет линейную сложность, равную числу своих аргументов, даже для схем глубины 1 (в силу того, что любая булева функция входит в базис). Поэтому имеет смысл изучать сложность совместной реализации систем булевых функций, т. е. операторов. В то же время, максимальная сложность оператора с n входами и n выходами ограничена величиной n^2 (следовательно, экспоненциальных оценок сложности в данной модели в принципе нельзя получить).

В работе [7] и в настоящей работе рассматривается одна и та же последовательность операторов, а именно, операторы циклической свёртки. Стоит отметить, что аналогичные оценки могут быть легко получены и для других операторов мультипликативного типа, а именно, для операторов умножения матриц и многочленов над полем $\mathbb{Z}_2$. Более того, в случае схем глубины 2, нижняя оценка сложности умножения матриц размера $n \times n$ по порядку совпадает с верхней и по порядку равна n^3.

2 Предыдущие результаты. Методы доказательства

В предыдущих работах, известных автору, доказательство имеет теоретико-графовую природу. А именно, выделяется некоторое теоретико-графовое свойство, которому должна удовлетворять любая схема, вычисляющая данный оператор. Далее доказывается, что сложность любой сети данной глубины, которая удовлетворяет этому свойству, не меньше требуемой величины.

Наиболее часто применяемым теоретико-графовым свойством является наличие суперконцентратора. Суперконцентратор — это сеть с n входами и n выходами, в которой любое подмножество входов связано семейством непересекающихся путей с любым подмножеством выходов той же мощности.

Например, любая схема для оператора свёртки должна быть суперконцентратором [1]. Для суперконцентраторов любой конечной глубины получены нелинейные нижние оценки числа рёбер, из которых следуют соответствующие нижние оценки сложности схем.

Первая нелинейная нижняя оценка для суперконцентратора конечной глубины (а именно, глубины 2) была получена в работе [2]. Затем последовала серия работ, в которых этот результат или улучшался, или рассматривались сети другой глубины. К настоящему времени известен порядок минимальной сложности суперконцентратора для любой конечной глубины. Сводка результатов на эту тему содержится в работе [5].

Отметим, что для глубины 2 минимальная сложность суперконцентратора по порядку равна $n\frac{\log^2 n}{\log\log n}$, а для глубины 3 она равна $n\log\log n$. До результатов автора эти значения являлись наибольшими известными нижними оценками сложности в классах схем глубины 2 и 3 соответственно. В работах [3, 6] использовались теоретико-графовые свойства, более слабые, чем наличие суперконцентратора. Соответственно, полученные нижние оценки были ещё меньше.

Мы не выделяем используемое нами теоретико-графовое свойство; по нашему мнению, оно не имеет самостоятельного значения, такого, как имеют суперконцентраторы. Идея доказательства для схем глубины 2 состоит в следующем: оператор зависит от двух наборов переменных: $\tilde{x}$ и $\tilde{y}$. Выделим группу входов из набора $\tilde{x}$ и группу выходов. На входы $\tilde{y}$ будем подавать разные значения и при этом, в силу выбора оператора, на выделенных выходах будут реализовываться разнообразные функции от выделенных входов (остальные входы мы не учитываем).

Это разнообразие окажется достаточно велико (в информационном смысле). Поэтому между выделенными входами и выходами должно находиться большое количество рёбер (чтобы в процессе вычисления мы могли учесть соответствующий объём управляющей информации, которая поступает со входов $\tilde{y}$). Поскольку глубина схемы равна 2, то эти рёбра инцидентны либо выделенным входам, либо выделенным выходам. Суммируя по всем группам входов и выходов, с учётом циклического сдвига, получим нижнюю оценку числа рёбер.

Случай глубины 3 сводится к случаю глубины 2 и это сведение можно изложить исключительно в графовых терминах. По-видимому, используя более изощрённую графовую технику можно улучшить имеющиеся нижние оценки для произвольной конечной глубины. В то же время представляет интерес поиск общего метода, частным случаем которого является метод автора для глубины 2.

3 Определения и формулировки результатов

Пусть n — натуральное число. *Булевой* функцией n переменных называется функция вида $f\colon \{0,1\}^n \to \{0,1\}$. Введём систему булевых функций $H = (h_1, \ldots, h_n)$, в которой каждая функция h_j зависит от $2n$ аргументов

$x_1, \ldots, x_n,\ y_1, \ldots, y_n$. По определению, положим

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

Оператор H называется *булевой циклической свёрткой*.

Напомним широко известное понятие булевой *схемы* из функциональных элементов. В данном случае нас интересуют схемы глубины d (где $d = 2$ или $d = 3$) с произвольными элементами, вычисляющие оператор H. Определим схему S интересующего нас вида как тройку объектов $(G, g, \prec)$, удовлетворяющую следующим условиям.

1) G — конечный неориентированный граф без ориентированных циклов и длина максимального ориентированного пути в нём равна d.

2) В графе G имеется $2n$ *входов* и n *выходов*. Вершина является входом тогда и только тогда, когда в неё не входит ни одного ребра. Вершина является выходом тогда и только тогда, когда из неё не выходит ни одного ребра.

3) Отображение g каждой вершине v, отличной от входов, ставит в соответствие булеву функцию, которая *локально вычисляется* в этой вершине; обозначим эту функцию через $g[v]$. При этом число аргументов функции $g[v]$ равно числу рёбер, входящих в вершину v.

4) Отношение $\prec$ есть полный порядок на множестве вершин графа G. Порядок $\prec$ нужен для того, чтобы мы могли определить, каким образом рёбра, входящие в данную вершину v, соответствуют аргументам функции $g[v]$. А именно, ребра упорядочиваются в соответствии с порядком вершин, из которых они исходят. Кроме того, отношение $\prec$ упорядочивает множество входов и множество выходов.

Входам схемы, в соответствии с порядком $\prec$, припишем переменные $x_1, \ldots, x_n,\ y_1, \ldots, y_n$. Выходы схемы, в соответствии с тем же порядком, будем обозначать через $z_1, \ldots, z_n$.

5) Для каждого $j = 1, \ldots, n$ функция h_j вычисляется на выходе z_j. Определим это условие подробнее.

Для каждой вершины v определим функцию, которая *глобально вычисляется* в вершине v; обозначим эту функцию через $f[v]$. Функция $f[v]$, в отличие от $g[v]$, зависит от входных переменных схемы, т. е. от $x_1, \ldots, x_n$, $y_1, \ldots, y_n$. Определим функцию $f[v]$ по индукции. Если v — вход схемы, то положим, что $f[v]$ тождественно равна переменной, соответствующей этому входу.

Пусть теперь v не является входом и пусть $v_1, \ldots, v_k$ — все вершины, из которых выходят рёбра, входящие в v, причём $v_1 \prec \ldots \prec v_k$ (порядок $\prec$ обсуждался ранее). Тогда, предполагая, что функции $f[v_1], \ldots, f[v_k]$ уже определены, зададим функцию $f[v]$ тождеством

$$f[v](\tilde{x}) \equiv g[v](f[v_1](\tilde{x}), \ldots, f[v_k](\tilde{x})) \ , \qquad (2)$$

где $\tilde{x} = (x_1, \ldots, x_n,\ y_1, \ldots, y_n)$. Условие 5) состоит в том, что $h_j = f[z_j]$.

Число рёбер в графе G назовём *сложностью* схемы S и обозначим её через $L(S)$.

В работе доказаны следующие теоремы.

Theorem 1. *Если $d = 2$, то $L(S) = \Omega(n^{3/2})$.*

Theorem 2. *Если $d = 3$, то $L(S) = \Omega(n \log n)$.*

Дальнейший план изложения таков. В разделе 4 содержится лемма 1, в которой сосредоточена теоретико-сложностная составляющая доказательства обеих теорем. Затем приводятся следствия из основной леммы. Одним из следствий является теорема 1; другое следствие будет использовано при доказательстве теоремы 2. Раздел 5 содержит общематематическую лемму 2. Наконец, раздел 6 содержит доказательство теоремы 2, опирающееся на предшествующие разделы.

Сделаем одно замечание. Вершины схемы естественным образом разбиваются на *ярусы*. А именно, вершина относится к k-му ярусу, если длина наибольшего ориентированного пути, ведущего из некоторого входа в данную вершину, равна k. Например, входы являются вершинами 0-го яруса. Мы знаем, что число ярусов равно d. Без ограничения общности мы можем считать, что схема является *строго* d-ярусной, а именно, что все выходы лежат на последнем ярусе и каждое ребро соединяет вершины из соседних ярусов.

Действительно, произвольную d-ярусную схему можно привести к строго d-ярусной, заменяя при необходимости рёбра цепочками рёбер и вводя промежуточные вершины. При этом сложность схемы увеличивается не более, чем в d раз, что приемлемо для нас, поскольку d — константа, а в теоремах 1 и 2 сложность схемы фигурирует с точностью до порядка. В дальнейшем будем считать, что рассматриваемые схемы являются строго d-ярусными.

4 Случай $d = 2$. Лемма 1 и её следствия

Пусть $f, f_1, \ldots, f_k$ — булевы функции, зависящие от одного и того же набора переменных $\tilde{y} = (y_1, \ldots, y_n)$. Скажем, что функция f *вычислима* через функции $f_1, \ldots, f_k$, если существует такая булева функция Φ от k переменных, что

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

Поскольку функцию Φ от k переменных мы можем выбрать 2^{2^k} способами, то число различных функций, вычислимых через $f_1, \ldots, f_k$, не превосходит 2^{2^k}. С другой стороны, если $f_1, \ldots, f_k$ — независимые переменные, например, $f_1 = y_1, \ldots, f_k = y_k$, то через $f_1, \ldots, f_k$ можно вычислить любую функцию от переменных $y_1, \ldots, y_k$. Следовательно, существует ровно 2^{2^k} функций, вычислимых через $f_1, \ldots, f_k$.

В этом разделе мы будем рассматривать случай $d = 2$. Мы уже упоминали о том, что вершины графа G можно разбить на ярусы. Его рёбра также разобьём на ярусы. А именно, ребро относится к i-му ярусу, если оно выходит из вершины $(i-1)$-го яруса (и входит, соответственно, в вершину

i-го яруса). Число рёбер i-го яруса обозначим через L_i. Некоторые вершины среднего (т. е. первого) яруса мы будем считать *особыми*. Каждая особая вершина должна быть соединена рёбрами со всеми входами. Однако, не каждая такая вершина может являться особой. Обозначим через L_2^* число рёбер второго яруса, не инцидентных особым вершинам.

Lemma 1. *Пусть k и l — натуральные числа и $kl \leqslant n$. Тогда*

$$kL_1 + lL_2^* \geqslant nkl \ .$$

Proof. Для каждой вершины v графа G введём вспомогательные функции $f_0[v], f_1[v], \ldots, f_n[v]$. Они являются подфункциями функции $f[v]$, причём каждая из функций $f_i[v]$ зависит только от переменных $y_1, \ldots, y_n$, а вместо переменных $x_1, \ldots, x_n$ подставлены те или иные константы. А именно, для получения функции $f_0[v]$ нужно подставить в $f[v]$ нули вместо переменных $x_1, \ldots, x_n$. Для получения функции $f_i[v]$ $(i = 1, \ldots, n)$ нужно подставить единицу вместо переменной x_i и нули вместо остальных переменных $x_1, \ldots, x_n$.

Пусть J — множество, состоящее из первых l натуральных чисел, дающих при делении на k остаток 1. Обозначим через $\mathcal{F}$ множество функций вида $f_i[z_j]$, где $1 \leqslant i \leqslant k$ и $j \in J$. По определению оператора H, функция $f_i[z_j]$ равна коэффициенту при переменной x_i в представлении (1), т. е. y_{i+j-1}. Если i пробегает значения от 1 до k, а j пробегает множество J, то $i + j - 1$ при этом пробегает значения от 1 до kl. Следовательно, множество $\mathcal{F}$ состоит из независимых переменных $y_1, \ldots, y_{kl}$.

Обозначим через X_i множество вершин среднего яруса, которые соединены рёбрами со входом x_i; через Z_j обозначим множество вершин среднего яруса, соединённых рёбрами с выходом z_j. Заметим, что в силу тождества (2), функция $f[z_j]$ вычислима через функции вида $f[v]$, где $v \in Z_j$. Тогда функция $f_i[z_j]$ вычислима через функции вида $f_i[v]$, $v \in Z_j$, так как при подстановке констант вместо переменных тождество сохраняется, а значит, и отношение вычислимости сохраняется.

Обозначим множество функций вида $f_i[v]$, где $1 \leqslant i \leqslant k$, $v \in Z_j$ и $j \in J$, через $\mathcal{G}$. Мы видим, что все функции из множества $\mathcal{F}$ вычислимы через функции из $\mathcal{G}$. В то же время, $\mathcal{F}$ состоит из kl независимых переменных. Поэтому число функций, вычислимых через множество $\mathcal{F}$, равно $2^{2^{kl}}$. Но каждая функция, вычислимая через $\mathcal{F}$, вычислима также через $\mathcal{G}$ (так как отношение вычислимости транзитивно). Поэтому через $\mathcal{G}$ вычислимы не менее, чем $2^{2^{kl}}$ функций, а значит,

$$|\mathcal{G}| \geqslant kl \ . \tag{3}$$

Мы подошли к основной моменту в доказательстве. Если вершина v среднего яруса не соединена со входом x_i, то функция $f[v]$ не зависит от переменной x_i, следовательно, $f_i[v] = f_0[v]$. Во всех случаях, где это возможно, заменим в множестве $\mathcal{G}$ функцию $f_i[v]$ на равную ей функцию $f_0[v]$. Тогда множество $\mathcal{G}$ содержит только те функции вида $f_i[v]$, для которых вершина v соединена со входом x_i, т. е. для которых $v \in X_i$. Кроме того, множество $\mathcal{G}$ содержит функции вида $f_0[v]$ такие, что $v \in Z_j$, $j \in J$, и вершина v не соединена хотя бы с одним из входов $x_1, \ldots, x_k$.

Вспомним, что каждая особая вершина соединена со всеми входами. Следовательно, множество $\mathcal{G}$ содержит только функции вида $f_i[v]$, где $v \in X_i$, $1 \leqslant i \leqslant k$, и функции вида $f_0[v]$, где v — неособая вершина, $v \in Z_j$ и $j \in J$. Обозначив через Z_j^* множество неособых вершин из Z_j, получим

$$|\mathcal{G}| \leqslant \sum_{i=1}^{k} |X_i| + \sum_{j \in J} |Z_j^*| \ .$$

Отсюда и из (3) следует

$$kl \leqslant \sum_{i=1}^{k} |X_i| + \sum_{j \in J} |Z_j^*| \ . \tag{4}$$

Заметим, что предшествующая часть доказательства леммы не изменится, если индекс i будет пробегать не множество $1, 2, \ldots, k$, а любое другое множество, полученное из данного циклическим сдвигом по модулю n. Аналогично, мы можем сдвигать область изменения индекса j, т. е. множество J. Для удобства сдвиги индексов i и j будем производить синхронно, т. е. на одно и то же число позиций. В результате получим n различных сдвигов и для каждого из них будет справедливо неравенство, аналогичное (4). Сложив все эти неравенства, получим

$$nkl \leqslant k \sum_{i=1}^{n} |X_i| + l \sum_{j=1}^{n} |Z_j^*| \ . \tag{5}$$

Для завершения доказательства осталось заметить, что первая сумма в (5) равна L_1, а вторая — L_2^*. □

Corollary 1. $L(S) = \Omega(n^{3/2})$.

Proof. Применим лемму к числам $k = l = [\sqrt{n}]$. Получим

$$[\sqrt{n}] L(S) = [\sqrt{n}](L_1 + L_2) \geqslant [\sqrt{n}](L_1 + L_2^*) \geqslant n \cdot [\sqrt{n}]^2 \ . \quad \square$$

Итак, мы доказали теорему 1.

Corollary 2. *Если число особых вершин не больше, чем $\frac{n}{2}$, то $L_1 \cdot L_2^* \geqslant \frac{n^3}{16}$.*

Proof. Заметим, что число вершин среднего слоя не меньше, чем n. Это очевидно в силу необходимости передачи информации между входами и выходами. Более подробно: функции $f_1[z_1], \ldots, f_1[z_n]$ являются попарно различными переменными. В то же время, они вычислимы через функции вида

$f_1[v]$, где v — вершины среднего слоя. Значит, число вершин среднего слоя не меньше, чем n.

Тогда, в силу условия следствия, число неособых вершин среднего слоя не меньше, чем $\frac{n}{2}$. Из каждой вершины среднего слоя выходит хотя бы одно ребро (т. к. эти вершины не являются выходами), следовательно, $L_2^* \geqslant \frac{n}{2}$. Кроме того, из каждого входа схемы также выходит хотя бы одно ребро, а значит, $L_1 \geqslant n$.

Рассмотрим 3 случая: 1) $L_1 \geqslant \frac{n^2}{2}$; 2) $L_2^* \geqslant n^2$; 3) $L_1 < \frac{n^2}{2}$ и $L_2^* < n^2$.

Случай 1. Имеем $L_1 \cdot L_2^* \geqslant \frac{n^2}{2} \cdot \frac{n}{2} = \frac{n^3}{4}$.

Случай 2. Аналогично, $L_1 \cdot L_2^* \geqslant n \cdot n^2 = n^3$.

Случай 3. Выберем числа k и l:

$$k = \left[\left(n \cdot \frac{L_2^*}{L_1} \right)^{1/2} \right] \quad , \quad l = \left[\left(n \cdot \frac{L_1}{L_2^*} \right)^{1/2} \right] \ .$$

Из условия случая 3 следует, что $nL_1 \geqslant n^2 > L_2^*$ и $nL_2^* \geqslant \frac{n^2}{2} > L_1$. Значит, числа k и l являются натуральными, т. е. определены корректно. Кроме того, $kl \leqslant n$. Применяя лемму 1, и используя неравенство $[x] \geqslant \frac{x}{2}$, справедливое при $x \geqslant 1$, получим

$$n \leqslant \frac{L_1}{l} + \frac{L_2^*}{k} = \frac{L_1}{\left[\left(n \cdot \frac{L_1}{L_2^*} \right)^{1/2} \right]} + \frac{L_2^*}{\left[\left(n \cdot \frac{L_2^*}{L_1} \right)^{1/2} \right]} \leqslant$$

$$\leqslant \frac{2L_1}{\left(n \cdot \frac{L_1}{L_2^*} \right)^{1/2}} + \frac{2L_2^*}{\left(n \cdot \frac{L_2^*}{L_1} \right)^{1/2}} = 4 \left(\frac{L_1 L_2^*}{n} \right)^{1/2} \ .$$

Окончательно, $L_1 L_2^* \geqslant \frac{n^3}{16}$. $\square$

5 Общематематическая лемма 2

Следующая лемма является обобщением леммы 4 из работы [4]. Первоначальное доказательство автора было улучшено рецензентом.

Lemma 2. *Пусть $a_1, \ldots, a_n$, $b_1, \ldots, b_n$ — неотрицательные действительные числа, $a_1 \geqslant \ldots \geqslant a_n$, и*

$$\begin{aligned} a_1 + a_2 + \ldots + a_n \ &\geqslant \ b_1 + b_2 + \ldots + b_n \ , \\ a_2 + \ldots + a_n \ &\geqslant \ b_2 + \ldots + b_n \ , \\ &\ldots \\ a_n \ &\geqslant \ b_n \ . \end{aligned} \tag{6}$$

Далее, пусть $\varphi\colon \mathbb{R}_+ \to \mathbb{R}_+$ — возрастающая вогнутая функция. Тогда

$$\varphi(a_1) + \ldots + \varphi(a_n) \geqslant \varphi(b_1) + \ldots + \varphi(b_n) \ .$$

Отметим, что требование неотрицательности чисел a_i и b_i является лишним при условии, что эти числа входят в область определения функции φ. Мы оставили это условие, поскольку лемма будет в дальнейшем применяться к функции $\sqrt{x}$, определённой на неотрицательных числах.

Proof. Вначале покажем, что если $a \geqslant b \geqslant \varepsilon > 0$, то

$$\varphi(a) + \varphi(b) \geqslant \varphi(a+\varepsilon) + \varphi(b-\varepsilon) \ . \tag{7}$$

Другими словами, если большее из чисел a, b увеличить на ε, а меньшее уменьшить на ε, то сумма $\varphi(a) + \varphi(b)$ не увеличится.

Действительно, обозначим $\delta = \frac{\varepsilon}{a-b+2\varepsilon}$. Тогда, в силу неравенства Иенсена,

$$\delta\varphi(b-\varepsilon) + (1-\delta)\varphi(a+\varepsilon) \leqslant \varphi(\delta(b-\varepsilon) + (1-\delta)(a+\varepsilon)) = \\ \varphi(a+\varepsilon-\delta(a-b+2\varepsilon)) = \varphi(a) \ .$$

Аналогично,

$$(1-\delta)\varphi(b-\varepsilon) + \delta\varphi(a+\varepsilon) \leqslant \varphi(b) \ .$$

Суммируя два последние неравенства, придём к неравенству (7).

Далее доказательство леммы проведём индукцией по n. Базис индукции: $n = 1$. В этом случае утверждение леммы следует из (6) и возрастания функции φ.

Индуктивный переход: $n \geqslant 2$. Увеличим число a_1 и одновременно уменьшим число a_n на одну и ту же величину ε. При этом ε выберем как максимально возможное число, для которого после указанного преобразования будут выполнены все неравенства (6). В результате преобразования сумма $\varphi(a_1) + \ldots + \varphi(a_n)$ не увеличится, так как по условию леммы $a_1 \geqslant a_n$ (см. также неравенство (7)). Следовательно, если мы докажем лемму для нового набора чисел $a_1, \ldots, a_n$, то тем самым мы докажем её для прежнего набора.

Докажем лемму для нового набора $a_1, \ldots, a_n$. В силу максимальности ε, хотя бы одно из неравенств (6), за исключением первого, обращается в равенство. Действительно, при одновременном увеличении a_1 и уменьшении a_n на одно и то же число сумма $a_1 + \ldots + a_n$ не изменяется, следовательно, первое неравенство системы (6) остаётся справедливым. В то же время левая часть каждого последующего неравенства уменьшается. Поэтому из невозможности выбора большего ε следует, что одно из последующих неравенств достигло своего критического состояния, т. е. обратилось в равенство.

Таким образом, для некоторого $k \geqslant 2$ выполнено

$$a_k + \ldots + a_n = b_k + \ldots + b_n \ . \tag{8}$$

Систему неравенств (6) обозначим через $\Phi(a_1, \ldots, a_n;\ b_1, \ldots, b_n)$. Вычтем равенство (8) из первых $k-1$ неравенств этой системы. При этом система (6) распадётся на две независимые системы того же вида, а именно,

$$\Phi(a_1, \ldots, a_{k-1};\ b_1, \ldots, b_{k-1}) \ , \\ \Phi(a_k, \ldots, a_n;\ b_k, \ldots, b_n) \ .$$

Применив к этим двум системам предположение индукции, получим

$$\varphi(a_1) + \ldots + \varphi(a_{k-1}) \geqslant \varphi(b_1) + \ldots + \varphi(b_{k-1}) \ ,$$
$$\varphi(a_k) + \ldots + \varphi(a_n) \geqslant \varphi(b_k) + \ldots + \varphi(b_n) \ .$$

Наконец, сложив два последних неравенства, мы завершим доказательство леммы. □

6 Случай $d = 3$. Доказательство теоремы 2

В данном разделе мы считаем, что $d = 3$. Рассмотрим вершины второго яруса (т. е. яруса, предшествующего выходам) и обозначим их через $v_1, \ldots, v_t$. Пусть $d^+(v_i)$ — число рёбер, входящих в вершину v_i, а $d^-(v_i)$ — число рёбер, выходящих из v_i. Определим число a_i следующим образом:

$$a_i = d^+(v_i) \cdot d^-(v_i) \ , \quad i = 1, \ldots, t \ .$$

Без ограничения общности мы можем считать, что $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_t$.

Положим $m = [n/2]$. Для каждого числа $p = 1, 2, \ldots, m$ преобразуем нашу схему S в схему S_p глубины 2, реализующую тот же оператор H. Преобразование осуществим следующим образом: вершины $v_1, \ldots, v_{p-1}$ переместим на первый ярус схемы, соединив их рёбрами со всеми входами, а вершины $v_p, \ldots, v_t$ удалим из схемы.

При удалении вершины v_i, мы введём новые рёбра для сохранения функциональности схемы. А именно, если вершина w первого яруса предшествует вершине v_i (т. е. из вершины w в вершину v_i ведёт ребро) и если вершина v_i предшествует выходу z_j, то мы проведём новое ребро непосредственно из вершины w к выходу z_j. Так мы поступим для каждой пары вершин (w, z_j). Таким образом, при удалении вершины v_i мы введём a_i новых рёбер.

Вершины $v_1, \ldots, v_{p-1}$ будем считать особыми для схемы S_p (особые вершины мы ввели в разделе 4). Число особых вершин не больше, чем $n/2$, поэтому мы можем применить следствие 2 из леммы 1. Напомним, что через L_i обозначается число рёбер i-го яруса, а через L_2^* — число рёбер 2-го яруса, выходящих из неособых вершин. Для того, чтобы отличать, к какой именно схеме относится число L_i или L_2^*, будем в скобках после этого числа указывать схему. Применяя следствие 2 из леммы 1, получим

$$L_1(S_p) \cdot L_2^*(S_p) \geqslant \frac{n^3}{16} \ . \tag{9}$$

Подсчитаем, чему равны числа $L_1(S_p)$ и $L_2^*(S_p)$. Каждое ребро первого яруса схемы S_p либо было ребром первого яруса в схеме S, либо было добавлено в процессе преобразования схемы. Мы добавили рёбра, которые соединяют вершины $v_1, \ldots, v_{p-1}$ со всеми входами схемы. Поскольку число входов равно $2n$, то общее число добавленных рёбер есть $2n(p-1)$. Итак,

$$L_1(S_p) = L_1(S) + 2n(p-1) \ .$$

Далее, рёбра второго яруса, выходящие из неособых вершин — это в точности те рёбра, которые мы добавили в схему S_p при удалении вершин $v_p, \ldots, v_t$. При удалении вершины v_i добавляется a_i новых рёбер, поэтому

$$L_2^*(S_p) = \sum_{i=p}^{t} a_i \ .$$

Обозначим $\theta = \frac{L_1(S)}{2n}$. Подставим в неравенство (9) найденные значения для $L_1(S_p)$ и $L_2^*(S_p)$ и преобразуем его. В результате получим

$$\sum_{i=p}^{t} a_i \geqslant \frac{n^3}{16(L_1(S) + 2(p-1)n)} = \frac{n^2}{32(\theta + p - 1)} \ . \tag{10}$$

Теперь применим лемму 2. Для этого введём числа $b_1, \ldots, b_t$ следующим образом:

$$b_p = \tfrac{n^2}{32}\left(\tfrac{1}{\theta+p-1} - \tfrac{1}{\theta+p}\right) \ , \quad p = 1, \ldots, m-1 \ ,$$
$$b_m = \tfrac{n^2}{32}\left(\tfrac{1}{\theta+m-1}\right) \ , \quad b_{m+1} = \ldots = b_t = 0 \ .$$

Заметим, что при таком определении чисел b_i система неравенств (6) следует из неравенств (10), взятых для разных p. Действительно, число b_i определено как разность двух соседних чисел, фигурирующих в правой части неравенства (10). Поэтому в сумме $b_p + \ldots + b_t$, после представления каждого слагаемого в виде разности двух чисел, промежуточные слагаемые сократятся и останется первое число, стоящее справа в (10). Заметим также, что последние неравенства системы (6), не имеющие аналогов в системе (10), верны в силу того, что $b_{m+1} = \ldots = b_t = 0$, а числа a_i неотрицательны.

Применяя лемму 2 к функции $\varphi(x) = \sqrt{x}$, получим

$$\sum_{i=1}^{t} \sqrt{a_i} \geqslant \sum_{i=1}^{m} \left(\frac{n^2}{32}\left(\frac{1}{\theta+i-1} - \frac{1}{\theta+i}\right)\right)^{1/2} = \sum_{i=1}^{m} \frac{n}{(32(\theta+i-1)(\theta+i))^{1/2}} \geqslant$$
$$\frac{n}{4\sqrt{2}} \sum_{i=1}^{m} \frac{1}{\theta+i} = \frac{n}{4\sqrt{2}} \left(\ln(\theta+m) - \ln\theta + O(1)\right) \ . \tag{11}$$

С другой стороны, в силу определения чисел a_i,

$$\sqrt{a_i} = \left(d^+(v_i) \cdot d^-(v_i)\right)^{1/2} \leqslant \frac{d^+(v_i) + d^-(v_i)}{2} \ . \tag{12}$$

Наконец, учтём, что рёбра второго яруса схемы S — это в точности рёбра, входящие в вершины $v_1, \ldots, v_t$, а рёбра третьего яруса — это рёбра, выходящие из этих вершин. Следовательно,

$$L_2(S) = \sum_{i=1}^{t} d^+(v_i) \ , \quad L_3(S) = \sum_{i=1}^{t} d^-(v_i) \ .$$

Поэтому, суммируя неравенства (12), получим

$$\sum_{i=1}^{t} \sqrt{a_i} \leqslant \frac{1}{2} \sum_{i=1}^{t} \left(d^+(v_i) + d^-(v_i)\right) = \frac{1}{2} \left(L_2(S) + L_3(S)\right) \ . \tag{13}$$

В заключение рассмотрим два случая. Если $\theta \geqslant \ln n$, то $L(S) \geqslant L_1(S) = 2\theta n \geqslant 2n \ln n$. Если же $\theta < \ln n$, то из (11) и (13) следует

$$L(S) \geqslant L_2(S) + L_3(S) = \Omega\left(\frac{n}{2\sqrt{2}} \ln\left(1 + \frac{m}{\theta}\right)\right) = \Omega(n \ln n) \ .$$

Итак, в любом случае $L(S) = \Omega(n \log n)$, а значит, мы доказали теорему 2.

7 Благодарности

Работа переведена на английский язык [автор перевода]. Автор выражает благодарность Программному комитету конференции CSR-2008 за заинтересованность в данной работе. 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. Pudlak P., Savicky P.: On Shifting Networks. Theoretical Comput. Sci. **116** (1993) 415–419
4. Pudlak P.: Communication in Bounded Depth Circuits. Combinatorica **14:2** (1994) 203–216
5. Radhakrishnan J., Ta-Shma A.: Bounds for Dispersers, Extractors and Depth-Two Superconcentrators. SIAM J. of Discrete Mathematics **13:1** (2000) 2–24
6. Raz R., Shpilka A.: Lower Bounds for Matrix Product, in Bounded Depth Circuits with Arbitrary Gates. SIAM J. Comput. **32:2** (2003) 488–513
7. Cherukhin D. Yu.: A lower bound for complexity in the class of schemes of depth 2 without constraints on a basis. Moscow Univ. Math. Bull. **60:4** (2005) 42–44