Math-Net.Ru

Общероссийский математический портал

В. В. Подольский, А. А. Шерстов, Небольшое уменьшение степени многочлена с заданной знаковой функцией может экспоненциально увеличить его вес и длину, *Матем. заметки*, 2010, том 87, выпуск 6, 885–899

DOI: <https://doi.org/10.4213/mzm8734>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 44.207.124.84

26 апреля 2023 г., 17:33:52

Небольшое уменьшение степени многочлена с заданной знаковой функцией может экспоненциально увеличить его вес и длину

В. В. Подольский, А. А. Шерстов

Булева функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ называется *знаковой функцией* целочисленного многочлена $p(x)$, если $f(x) = \text{sgn}(p(x))$ для всех $x \in \{-1, +1\}^n$. При этом многочлен $p(x)$ называется *пороговым элементом* для булевой функции f . *Весом* порогового элемента называется сумма абсолютных значений коэффициентов p . Доказывается, что небольшое изменение степени порогового элемента для заданной функции может сильно сказаться на величине необходимого веса. Более точно, для всякого $d = 1, 2, \dots, n-1$ явно строится функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$, требующая веса $\exp\{\Theta(n)\}$ при представлении пороговыми элементами степени d и представимая пороговым элементом степени $d+1$ с весом лишь $O(n^2)$. Нижняя оценка $\exp\{\Theta(n)\}$ для степени d остается верной и для размера булевой схемы глубины 2 с функцией голосования на верхнем уровне и *произвольными* элементами входной степени d на нижнем уровне. Этот разрыв в величине весов экспоненциально больше, чем известные ранее. Аналогичный результат доказывается для *длины* порогового элемента, т.е. числа одночленов в нем.

Библиография: 35 названий.

1. Введение. Реализация булевых функций действительными многочленами играет важную роль в теории сложности, начиная от сложности вычислений и заканчивая квантовыми вычислениями и теорией обучения. Некоторые из этих разнообразных применений описаны в обзорах [1]–[4]. В данной работе мы рассматриваем одну из таких моделей, *пороговые элементы*. А именно, зафиксируем булеву функцию $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$. Эта функция называется *знаковой функцией* целочисленного многочлена $p(x_1, \dots, x_n)$, если $f(x_1, \dots, x_n) = \text{sgn}(p(x_1, \dots, x_n))$ для всех $x_1, \dots, x_n \in \{-1, +1\}$, где под функцией sgn мы понимаем знаковую функцию:

$$\text{sgn } t = \begin{cases} 1, & \text{если } t > 0, \\ 0, & \text{если } t = 0, \\ -1, & \text{если } t < 0. \end{cases}$$

Работа первого автора выполнена при поддержке Российского фонда фундаментальных исследований (грант № 06-01-00122) и программы “Ведущие научные школы” (грант № НШ-845.2008.1). Работа второго автора выполнена при поддержке Национального научного фонда США (гранты CCF-0728536 и “CAREER AWARD” на имя профессора А. Р. Клайванса).

При этом многочлен $p(x_1, \dots, x_n)$ называется пороговым элементом для булевой функции f . Основными мерами сложности порогового элемента являются: *степень*, обычная степень многочлена; *вес*, сумма абсолютных значений его коэффициентов; и *длина*, количество одночленов в нем. *Пороговой степенью* f называется наименьшая степень порогового элемента для f . *Пороговым весом* f называется наименьший вес порогового элемента для f . *Пороговой длиной* f называется наименьшая длина порогового элемента для f .

Пороговым элементам и их приложениям посвящено большое количество работ в сложности булевых схем, коммуникационной сложности и теории обучения. Изучение пороговых элементов началось в 1968 г. с книги Минского и Пейперта [5]. С тех пор понятие пороговой степени неоднократно использовалось разнообразными способами для доказательства нижних оценок на размер схем и, более широко, для изучения сложных классов [6]–[11]. С помощью нижних оценок на пороговую степень были доказаны некоторые нижние оценки на коммуникационную сложность, начиная с теоремы о пороговой степени и спектральной норме [11], [12] и заканчивая недавними результатами о знаковом ранге [13], [14]. Наконец, в теории обучения понятие пороговой степени использовалось в нескольких ключевых алгоритмах [15], [16] и нижних оценках [17].

Кроме пороговой степени, в книге Минского и Пейперта положено начало активному изучению порогового веса и его приложений. Впоследствии в теории обучения Джексон [18] построил широко известный алгоритм, *гармоническое решето*, для изучения ДНФ-формул и других функций с малым пороговым весом. Клайванс и др. [19], используя другой метод, основанный на линейном программировании, построили алгоритм для изучения конъюнкций пороговых элементов с низкой степенью и малым весом, в модели с допустимой вероятностью ошибки (PAC learning). Еще один пример использования пороговых элементов в теории обучения – работа Клайванса и Серведио [20] об изучении функций, зависящих лишь от малого числа входных переменных. Кроме того, понятие порогового веса неоднократно использовалось в теории сложности, в частности, в сложности вычислений с оракулом [21], [22] и коммуникационной сложности [23]. Наконец, пороговый вес изучался как самостоятельная мера сложности. Было разработано множество аналитических и комбинаторных методов для доказательства нижних оценок на пороговый вес [24]–[29], [17]. Весьма активно изучалось также понятие пороговой длины [24], [27], [30], [28], [17].

Кроме уже упомянутых исследований, изучались соотношения между степенью и весами пороговых элементов для заданных функций. Более точно, изучался наименьший вес, обозначаемый через $W(f, d)$, порогового элемента степени d для функции f , где f фиксирована, а d изменяется. Типы булевых функций, изучавшихся с этой точки зрения, включают пороговые элементы [25], [27], [31], списки разрешения [21], ДНФ-формулы [22] и другие функции [32], [33]. В последних результатах [32], [33] доказаны первые дважды-экспоненциальные нижние оценки на $W(f, d)$ для большого числа различных d .

2. Наши результаты. Эта работа продолжает исследование взаимосвязи между весами и степенями пороговых элементов. Основной вопрос, которым мы задаемся, звучит следующим образом: насколько чувствительна может быть величина $W(f, d)$ к небольшим изменениям d ? Во всех предыдущих работах эта величина изменяется весьма плавно как функция от d . Более точно, $W(f, d)$ и $W(f, d + 1)$

полиномиально связаны для всех допустимых d (т.е. для всех d не меньших, чем пороговая степень f). Наш основной результат – построение для всякого данного d булевой функции f с экспоненциальным разрывом между $W(f, d)$ и $W(f, d + 1)$.

ТЕОРЕМА 1. *Для всех n и всех $d = 1, 2, \dots, n - 1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

$$W(f, d) = \exp\{\Theta(n)\} \quad \text{и} \quad W(f, d + 1) = O(n^2).$$

Другими словами, мы покажем, что необходимый для вычисления заданной функции вес пороговых элементов может сильно измениться в ответ на небольшие изменения степени пороговых элементов.

Кроме того, мы доказываем следующий более сильный результат, относящийся к булевым схемам из функций голосования. *Функцией голосования* называется пороговый элемент степени 1 с коэффициентами ± 1 . Полезно иметь в виду, что пороговый элемент с весом W – это функция голосования от W функций логического сложения $\oplus$ (и наоборот).

ТЕОРЕМА 2. *Для всех n и всех $d = 1, 2, \dots, n - 1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

- (1) *всякая функция голосования, примененная к произвольным элементам входной степени не выше d и вычисляющая f , должна применяться к не менее, чем $\exp\{\Theta(n)\}$ элементам, что также является достаточным;*
- (2) *f вычислима функцией голосования от $O(n^2)$ функций логического сложения $\oplus$ входной степени не выше $d + 1$, а также функцией голосования от $O(n^2 d)$ функций голосования входной степени не выше $d + 1$.*

Другими словами, для вычисления функции f , которую мы строим, требуется функция голосования от экспоненциального числа элементов входной степени не выше d , независимо от того, что это за элементы. Этот результат сильнее, чем теорема 1, которая допускает только функции логического сложения $\oplus$.

Кроме того, мы доказываем результат, аналогичный теореме 1, для пороговой длины. Через $L(f, d)$ будем обозначать наименьшую длину порогового элемента для f степени не более d .

ТЕОРЕМА 3. *Для всех n и всех $d = 1, 2, \dots, n - 1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

$$L(f, d) = \exp\{\Theta(d)\} \quad \text{и} \quad L(f, d + 1) = O(d).$$

Оценка на $L(f, d)$ в этом результате близка к оптимальной, так как существует не более $(n + 1)^d$ одночленов степени не выше d и, следовательно, длина всякого порогового элемента степени не выше d не превышает $(n + 1)^d$.

3. Наши методы. Доказательства теорем 1, 2 и 3 сочетают методы дискретного преобразования Фурье, линейной алгебры и теории приближений. Отправной точкой в доказательстве служит экспоненциальный разрыв между $W(f, 1)$ и $W(f, 2)$ для некоторой функции f . Для этой цели мы используем функцию из работы [34].

Затем мы обобщаем этот результат до экспоненциального разрыва между $W(f, d)$ и $W(f, d + 1)$ для всякого $d = 1, 2, \dots, \epsilon n$, подходящим образом сдвигая спектр Фурье изначальной функции. Для этого мы доказываем следующий аналитический

результат: если спектр Фурье функции f сосредоточен на характерах из некоторого аффинного подпространства $\mathbb{F}_2^n$, тогда то же самое можно предполагать о многочленах, приближающих f или представляющих f своим знаком.

Наконец, на третьем этапе доказательства мы решаем задачу для случая $d > \epsilon n$. Здесь задействовано совершенно другое рассуждение. Мы используем результат работы [24] для оценки величины $W(f, d)$ через коэффициенты Фурье функции f . Это же рассуждение позволяет нам доказать теорему 3. Однако, это рассуждение не дает доказательства теоремы 2, для чего мы приводим еще одно рассуждение, использующее связь между приближением функции f многочленом большой степени и представлением f знаком многочлена. Кроме доказательства теоремы 2, это рассуждение также уточняет оценку в теореме 1.

4. Предварительные сведения. В данной работе под *булевой функцией* мы понимаем отображение $X \rightarrow \{-1, +1\}$, где X – конечное множество, а значение -1 соответствует “истине”. Для строки $x \in \{-1, +1\}^n$ и подмножества $S \subseteq \{1, \dots, n\}$, вводим обозначение $x|_S = (x_{i_1}, \dots, x_{i_d}) \in \{-1, +1\}^d$, где $i_1 < \dots < i_d$ – элементы S . Через $\mathbb{F}_2$ обозначаем поле из двух элементов.

Преобразование Фурье над $\mathbb{F}_2^n$. Рассмотрим векторное пространство действительных функций на множестве $\{-1, +1\}^n$ со скалярным произведением

$$\langle f, g \rangle = 2^{-n} \sum_{x \in \{-1, +1\}^n} f(x)g(x)$$

и нормой

$$\|f\|_\infty = \max_{x \in \{-1, +1\}^n} |f(x)|.$$

Для подмножества $S \subseteq \{1, \dots, n\}$ определим функцию $\chi_S: \{-1, +1\}^n \rightarrow \{-1, +1\}$ формулой

$$\chi_S(x) = \prod_{i \in S} x_i.$$

Эти функции называются *характерами*. Характеры образуют ортонормированный базис в рассматриваемом векторном пространстве. Следовательно, всякая функция $f: \{-1, +1\}^n \rightarrow \mathbb{R}$ единственным образом представляется в виде

$$f = \sum_{S \subseteq \{1, \dots, n\}} \hat{f}(S) \chi_S, \quad (4.1)$$

где $\hat{f}(S) = \langle f, \chi_S \rangle$ – коэффициент Фурье функции f , соответствующий характеру χ_S . Для функций вида $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ ортонормированность характеров влечет следующее тождество, называемое *равенством Парсеваля*:

$$\sum_{S \subseteq \{1, \dots, n\}} \hat{f}(S)^2 = 1. \quad (4.2)$$

Пороговые элементы. Для функции $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ *пороговой степенью* f , обозначаемой через $\deg_\pm(f)$, называется наименьшее d такое, что $f(x) \equiv \text{sgn } p(x)$ для некоторого многочлена $p \in \text{span}\{\chi_S : |S| \leq d\}$. Если пользоваться терминологией введения, то $\deg_\pm(f)$ есть наименьшая степень порогового элемента

для f . Для $d \geq \deg_{\pm}(f)$ через $L(f, d)$ обозначаем минимальное число ненулевых λ_S среди всех представлений вида

$$f(x) \equiv \operatorname{sgn} \left(\sum_{|S| \leq d} \lambda_S \chi_S(x) \right), \quad (4.3)$$

где λ_S – целые. Мы называем $L(f, d)$ *минимальной длиной порогового элемента степени d для f* . Аналогично, через $W(f, d)$ мы обозначаем минимум $\sum_{|S| \leq d} |\lambda_S|$ по всем целым λ_S таким, что верно (4.3). Мы называем $W(f, d)$ *минимальным весом порогового элемента степени d для f* .

Нам потребуется следующий результат из работы [24].

ТЕОРЕМА 4 (Брук). Пусть $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ – булева функция с пороговой степенью не выше d . Тогда

$$L(f, d) \geq \frac{1}{\max_{|S| \leq d} |\widehat{f}(S)|}.$$

В действительности в работе [24] был доказан более общий результат, но нам будет достаточно этого частного случая.

Еще одну тесно связанную с $W(f, d)$ величину обозначаем через $W^*(f, d)$ и определяем как наименьшее k , для которого существуют $f_i: \{-1, +1\}^d \rightarrow \{-1, +1\}$, $i = 1, 2, \dots, k$, и d -элементные подмножества $S_1, S_2, \dots, S_k \subseteq \{1, \dots, n\}$ такие, что

$$f(x) \equiv \operatorname{sgn} \left(\sum_{i=1}^k f_i(x|_{S_i}) \right). \quad (4.4)$$

Нетрудно понять, что именно для величины $W^*(f, d)$ доказывается нижняя оценка в теореме 2.

Заметим, что понятия $L(f, d)$ и $W(f, d)$ не имеют смысла для $d < \deg_{\pm}(f)$, поскольку $\deg_{\pm}(f)$ есть наименьшая степень порогового элемента для f . То же самое верно и для $W^*(f, d)$, поскольку всякая булева функция от d переменных может быть представлена в виде многочлена степени не выше d .

ЛЕММА 1. Пусть дана функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$. Тогда для всякого $d \geq \deg_{\pm}(f)$ выполнено

$$W^*(f, d) \leq W(f, d) \leq (2\sqrt{2})^d W^*(f, d).$$

ДОКАЗАТЕЛЬСТВО. Первое неравенство очевидно. Для доказательства второго положим $k = W^*(f, d)$ и зафиксируем функции $f_i: \{-1, +1\}^d \rightarrow \{-1, +1\}$, $i = 1, 2, \dots, k$, и d -элементные подмножества $S_1, S_2, \dots, S_k \subseteq \{1, \dots, n\}$, удовлетворяющие (4.4). Тогда следующее представление является пороговым элементом степени d для f :

$$f(x) \equiv \operatorname{sgn} \left(\sum_{i=1}^k \sum_{A \subseteq \{1, \dots, d\}} 2^d \widehat{f}_i(A) \chi_A(x|_{S_i}) \right).$$

Вклад f_i в вес этого порогового элемента не превышает

$$2^d \sum_{A \subseteq \{1, \dots, d\}} |\widehat{f}_i(A)| \leq 2^d \cdot 2^{d/2} \left(\sum_{A \subseteq \{1, \dots, d\}} \widehat{f}_i(A)^2 \right)^{1/2} = (2\sqrt{2})^d,$$

где первый и второй переходы следуют из неравенства Коши–Буняковского и равенства Парсеваля (4.2) соответственно. Суммирование вкладов всех f_i завершает доказательство.

Наконец, определим

$$E(f, d) = \min_p \|f - p\|_\infty,$$

где минимум берется по всем $p \in \text{span}\{\chi_S : |S| \leq d\}$. Другими словами, $E(f, d)$ – это минимальное отклонение при равномерном (чебышевском) приближении f многочленами степени d . Нетрудно понять, что если f принимает значения ± 1 , то $0 \leq E(f, d) \leq 1$.

ЛЕММА 2. Пусть дана функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$. Тогда для всякого $d \geq \deg_\pm(f)$ выполнено

$$W(f, d) \geq W^*(f, d) \geq \frac{1 + E(f, d)}{1 - E(f, d)}.$$

Аналогичные лемме 2 утверждения появлялись ранее в нескольких работах [35], [28], [23] с похожими на наше доказательствами.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ 2. Положим $k = W^*(f, d)$ и зафиксируем функции $f_i: \{-1, +1\}^d \rightarrow \{-1, +1\}$, $i = 1, 2, \dots, k$, и d -элементные подмножества $S_1, S_2, \dots, S_k \subseteq \{1, \dots, n\}$, удовлетворяющие (4.4). Заметим, что выражение

$$\sum_{i=1}^k f_i(x|_{S_i})$$

представимо в виде многочлена степени не выше d , причем

$$1 \leq f(x) \sum_{i=1}^k f_i(x|_{S_i}) \leq k$$

для всех $x \in \{-1, +1\}^n$. Следовательно,

$$E(f, d) \leq \max_{x \in \{-1, +1\}^n} \left| f(x) - \frac{2}{k+1} \sum_{i=1}^k f_i(x|_{S_i}) \right| \leq \frac{k-1}{k+1},$$

что и требовалось доказать.

5. О функциях с аффинно разреженным спектром. Для произвольной булевой функции f трудно охарактеризовать многочлены заданной степени, представляющие f своим знаком или приближающие f . Как мы увидим в этом пункте, нетривиальная характеристика все же возможна при дополнительных ограничениях на спектр Фурье f . Более конкретно, мы рассмотрим функции f , носителями спектра которых являются характеры из аффинного подпространства $\mathbb{F}_2^n$, и покажем, что это свойство переносится на все многочлены, приближающие f или представляющие f своим знаком. Чтобы пояснить идеи, стоящие за этим результатом, начнем с простого примера.

ПРИМЕР. Булева функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ называется *нечетной*, если $f(-x) \equiv -f(x)$. Пусть заданы нечетная булева функция f и многочлен p такой, что $f(x) \equiv \operatorname{sgn} p(x)$. Мы утверждаем, что

$$f(x) \equiv \operatorname{sgn} \left\{ \sum_{|S| \text{ нечетно}} \widehat{p}(S) \chi_S(x) \right\}, \quad (5.1)$$

т.е. из разложения Фурье p можно удалить четные члены. Действительно, из нечетности f следует, что $f(x) \equiv -\operatorname{sgn} p(-x)$, откуда

$$\begin{aligned} f(x) &\equiv \operatorname{sgn} \left\{ \frac{1}{2} p(x) - \frac{1}{2} p(-x) \right\} \\ &\equiv \operatorname{sgn} \left\{ \frac{1}{2} \sum_{|S| \text{ четно}} \widehat{p}(S) \chi_S(x) + \frac{1}{2} \sum_{|S| \text{ нечетно}} \widehat{p}(S) \chi_S(x) \right. \\ &\quad \left. - \frac{1}{2} \sum_{|S| \text{ четно}} \widehat{p}(S) \chi_S(x) + \frac{1}{2} \sum_{|S| \text{ нечетно}} \widehat{p}(S) \chi_S(x) \right\}, \end{aligned}$$

что влечет за собой (5.1).

Эта простая идея была использована в [31]. Немного более сложная конструкция была применена в [32]. Нам потребуется более общий результат, не рассматривавшийся ранее.

Для упрощения обозначений в этом пункте будет удобно предполагать, что входные булевы переменные принимают значения 0 и 1, а не 1 и -1 . А именно, мы рассматриваем векторное пространство действительных функций на $\mathbb{F}_2^n$ со скалярным произведением

$$\langle f, g \rangle = 2^{-n} \sum_{x \in \mathbb{F}_2^n} f(x) g(x)$$

и нормой

$$\|f\|_\infty = \max_{x \in \mathbb{F}_2^n} |f(x)|.$$

Для векторов $x, y \in \mathbb{F}_2^n$ введем обозначение

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

Тогда соответствующее преобразование Фурье функции $f: \mathbb{F}_2^n \rightarrow \mathbb{R}$ задается выражением

$$f = \sum_{w \in \mathbb{F}_2^n} \widehat{f}(w) \chi_w,$$

где характеры $\chi_w: \mathbb{F}_2^n \rightarrow \{-1, +1\}$ определяются формулой $\chi_w(x) = (-1)^{\langle w, x \rangle}$ и соответствующие коэффициенты Фурье суть $\widehat{f}(w) = \langle f, \chi_w \rangle$.

ТЕОРЕМА 5 (об аффинно разреженном спектре). Пусть функция $\phi: \mathbb{F}_2^n \rightarrow \mathbb{R}$ задана. Пусть $\mathcal{A}$ – аффинное подпространство в $\mathbb{F}_2^n$ такое, что $\widehat{\phi}(w) \neq 0$, только если $w \in \mathcal{A}$. Тогда для всякого многочлена $p: \mathbb{F}_2^n \rightarrow \mathbb{R}$ верно неравенство

$$\|\phi - p\|_\infty \geq \left\| \phi - \sum_{w \in \mathcal{A}} \widehat{\phi}(w) \chi_w \right\|_\infty. \quad (5.2)$$

Кроме того, если ϕ принимает значения ± 1 и $\phi(x) \equiv \operatorname{sgn} p(x)$, то

$$\phi(x) \equiv \operatorname{sgn} \left(\sum_{w \in \mathcal{A}} \widehat{p}(w) \chi_w \right).$$

ЗАМЕЧАНИЕ. Аффинное подпространство, соответствующее примеру, – это пространство всех двоичных векторов с нечетным числом единиц.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 5. Для функции ϕ , принимающей значения ± 1 , нетрудно видеть, что $\phi(x) \equiv \operatorname{sgn} p(x)$ тогда и только тогда, когда $\|\phi - \epsilon p\|_\infty < 1$ для всех достаточно малых $\epsilon > 0$. Следовательно, вторая часть теоремы сразу следует из (5.2). Далее в доказательстве сосредоточимся на (5.2).

Можно предполагать, что $\mathcal{A}$ строго содержится в $\mathbb{F}_2^n$, так как в противном случае теорема тривиальна. Запишем $\mathcal{A} = v + \mathcal{L}$ для подходящей точки $v \in \mathbb{F}_2^n$ и линейного подпространства $\mathcal{L} \subset \mathbb{F}_2^n$. Пусть $v_1, v_2, \dots, v_k$ – базис ортогонального дополнения $\mathcal{L}$. Для $\lambda \in \mathbb{F}_2^k$, предположение о спектре Фурье функции ϕ влечет за собой

$$\phi \left(x + \sum \lambda_i v_i \right) \equiv \sum_{w \in \mathcal{A}} \widehat{\phi}(w) (-1)^{\langle w, x \rangle} (-1)^{\sum \lambda_i \langle w, v_i \rangle} \equiv (-1)^{\sum \lambda_i \langle v, v_i \rangle} \phi(x).$$

В частности,

$$\|\phi - p\|_\infty = \max_x \left| \phi(x) - (-1)^{\sum \lambda_i \langle v, v_i \rangle} p \left(x + \sum \lambda_i v_i \right) \right|,$$

откуда по выпуклости модуля

$$\|\phi - p\|_\infty \geq \max_x \left| \phi(x) - \mathbb{E}_\lambda \left[(-1)^{\sum \lambda_i \langle v, v_i \rangle} p \left(x + \sum \lambda_i v_i \right) \right] \right|. \quad (5.3)$$

Далее,

$$\mathbb{E}_\lambda \left[(-1)^{\sum \lambda_i \langle v, v_i \rangle} p \left(x + \sum \lambda_i v_i \right) \right] = \sum_{w \in \mathbb{F}_2^n} \widehat{p}(w) (-1)^{\langle w, x \rangle} \mathbb{E}_\lambda \left[(-1)^{\sum \lambda_i \langle w + v, v_i \rangle} \right].$$

В линейной алгебре доказывается, что математическое ожидание в последнем выражении равно 1, если $w \in v + \mathcal{L} = \mathcal{A}$, и равно нулю в противном случае. Следовательно,

$$\mathbb{E}_\lambda \left[(-1)^{\sum \lambda_i \langle v, v_i \rangle} p \left(x + \sum \lambda_i v_i \right) \right] = \sum_{w \in \mathcal{A}} \widehat{p}(w) \chi_w(x),$$

что завершает доказательство в силу (5.3).

СЛЕДСТВИЕ 1. Пусть дана функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$. Для целого числа $k \geq 1$ определим $F: \{-1, +1\}^{n+k} \rightarrow \{-1, +1\}$ формулой

$$F(x) \equiv x_1 x_2 \cdots x_k f(x_{k+1}, \dots, x_{k+n}).$$

Тогда

$$E(F, d) = E(f, d - k), \quad W(F, d) = W(f, d - k), \quad d = k, k + 1, \dots.$$

ДОКАЗАТЕЛЬСТВО. Ясно, что $E(F, d) \leq E(f, d - k)$. В обратную сторону, зафиксируем многочлен p степени не выше d такой, что $\|F - p\|_\infty = E(F, d)$. Рассмотрим функцию $\phi: \mathbb{F}_2^{n+k} \rightarrow \{-1, +1\}$, заданную формулой

$$\phi(x_1, \dots, x_{n+k}) = F((-1)^{x_1}, \dots, (-1)^{x_{n+k}})$$

и аффинное подпространство $\mathcal{A} \subset \mathbb{F}_2^{n+k}$, определенное как

$$\mathcal{A} = (\overbrace{1, \dots, 1}^k, 0, \dots, 0) + \text{span}\{(\overbrace{0, \dots, 0}^i, 1, 0, \dots, 0) : i \geq k\}.$$

Применяя теорему 5 к ϕ и $\mathcal{A}$, мы заключаем, что

$$\begin{aligned} \|F - p\|_\infty &\geq \left\| F - \sum_{\substack{S \subseteq \{k+1, \dots, k+n\} \\ |S| \leq d-k}} \widehat{p}(S \cup \{1, \dots, k\}) \chi_{S \cup \{1, \dots, k\}} \right\|_\infty \\ &= \left\| F \cdot \chi_{\{1, \dots, k\}} - \sum_{\substack{S \subseteq \{k+1, \dots, k+n\} \\ |S| \leq d-k}} \widehat{p}(S \cup \{1, \dots, k\}) \chi_S \right\|_\infty. \end{aligned}$$

Поскольку $F \cdot \chi_{\{1, \dots, k\}} = f(x_{k+1}, \dots, x_{k+n})$, мы получаем $\|F - p\|_\infty \geq E(f, d - k)$, что и требовалось.

Доказательство $W(F, d) = W(f, d - k)$ полностью аналогично, за исключением того, что следует применить ту часть теоремы 5, которая относится к представлению функции знаком, а не равномерному приближению. Следствие доказано.

СЛЕДСТВИЕ 2. Пусть дана функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$. Для целого числа $k \geq 1$ определим $F: \{-1, +1\}^{kn} \rightarrow \{-1, +1\}$ формулой

$$F(x) \equiv f\left(\prod_{i=1}^k x_i, \dots, \prod_{i=1}^k x_{(n-1)k+i}\right).$$

Тогда

$$\begin{aligned} E(F, d) &= E\left(f, \left\lfloor \frac{d}{k} \right\rfloor\right), \quad d = 0, 1, 2, \dots, \\ W(F, d) &= W\left(f, \left\lfloor \frac{d}{k} \right\rfloor\right), \quad d = 0, 1, 2, \dots. \end{aligned}$$

ДОКАЗАТЕЛЬСТВО. Вновь, ясно, что $E(F, d) \leq E(f, \lfloor d/k \rfloor)$. В обратную сторону, зафиксируем многочлен p степени не выше d такой, что $\|F - p\|_\infty = E(F, d)$.

Рассмотрим функцию $\phi: \mathbb{F}_2^{kn} \rightarrow \{-1, +1\}$, заданную формулой

$$\phi(x_1, \dots, x_{kn}) = F((-1)^{x_1}, \dots, (-1)^{x_{kn}})$$

и аффинное подпространство $\mathcal{A} \subset \mathbb{F}_2^{kn}$, определенное как

$$\mathcal{A} = \text{span}\{(\overbrace{0, \dots, 0}^{ik}, \overbrace{1, \dots, 1}^k, 0, \dots, 0) : i = 0, 1, \dots, n-1\}.$$

Обозначим через K_i , где $i = 1, \dots, n$, множество $\{(i-1)k+1, \dots, ik\}$. Тогда, применяя теорему 5 к ϕ и $\mathcal{A}$, получаем

$$\|F - p\|_\infty \geq \left\| F - \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| \leq \lfloor d/k \rfloor}} \hat{p}\left(\bigcup_{i \in S} K_i\right) \chi_{\bigcup_{i \in S} K_i} \right\|_\infty = \left\| f - \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| \leq \lfloor d/k \rfloor}} \hat{p}\left(\bigcup_{i \in S} K_i\right) \chi_S \right\|_\infty.$$

Отсюда сразу следует, что $E(F, d) \geq E(f, \lfloor d/k \rfloor)$.

Как и в следствии 1, доказательство $W(F, d) = W(f, \lfloor d/k \rfloor)$ получается аналогично, за исключением того, что следует применить ту часть теоремы 5, которая относится к представлению функции знаком, а не равномерному приближению. Следствие доказано.

6. Случай малых степеней. В этом пункте мы получим требуемый экспоненциальный разрыв между $W(f, d)$ и $W(f, d+1)$ для $d = 1, 2, \dots, \epsilon n$. Случай больших степеней d потребует совершенно другой конструкции, которую мы откладываем до следующего пункта.

Отправным пунктом нашего рассуждения является результат, адаптированный из классической работы Майхилла и Кауца [34].

ТЕОРЕМА 6 (Майхилл и Кауц, 1961 г.). Пусть $f_m: \{-1, +1\}^{2m} \rightarrow \{-1, +1\}$ — функция, определенная формулой

$$f_m(x, y) \equiv \operatorname{sgn} \left\{ 2^{m+1} - 1 - \sum_{i=1}^m 2^i x_i^* - \sum_{i=1}^m (2^{m+1} - 2^i) y_i^* \right\},$$

где полагаем $b^* = (1 - b)/2$. Тогда

$$W(f_m, 1) = \Theta(m2^m).$$

Теперь мы дополним теорему 6 верхней оценкой на вес квадратичных пороговых элементов.

ТЕОРЕМА 7. Пусть $f_m: \{-1, +1\}^{2m} \rightarrow \{-1, +1\}$ — функция, определенная в теореме 6. Тогда

$$W(f_m, 2) \leq O(m^2).$$

ДОКАЗАТЕЛЬСТВО. Как и раньше, мы используем обозначение $b^* = (1 - b)/2$. Проверка показывает, что

$$f_m(x_1, \dots, x_m, y_1, \dots, y_m) \equiv \left(\bigvee_{1 \leq i < j \leq m} y_i \wedge y_j \right) \vee \left(\bigvee_{i=1}^m y_i \wedge (x_i \vee x_{i+1} \vee \dots \vee x_m) \right).$$

Из этого представления ясно, что

$$f_m(x, y) \equiv \operatorname{sgn} \left\{ \frac{1}{2} - \sum_{1 \leq i < j \leq m} y_i^* y_j^* - \sum_{i=1}^m y_i^* (x_i^* + x_{i+1}^* + \dots + x_m^*) \right\},$$

откуда следует требуемая верхняя оценка $W(f_m, 2)$.

Итак, имеем

ТЕОРЕМА 8. *Для любых n и $d = 1, 2, \dots, n-1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

$$W(f, d) = \Theta((n-d)2^{(n-d)/2}), \quad (6.1)$$

$$W(f, d+1) = O((n-d)^2). \quad (6.2)$$

ДОКАЗАТЕЛЬСТВО. Положим $m = \lfloor (n-d+1)/2 \rfloor$ и определим f как

$$f(x) = x_1 x_2 \cdots x_{d-1} f_m(x_d, x_{d+1}, \dots, x_{d+2m-1}),$$

где f_m — функция из теоремы 6. Тогда нужные оценки на $W(f, d)$ и $W(f, d+1)$ следуют из следствия 1 и теорем 6 и 7. Теорема доказана.

7. Случай больших степеней. В этом пункте мы получим искомый разрыв между $W(f, d)$ и $W(f, d+1)$ для $d > \epsilon n$, а также разрыв между $L(f, d)$ и $L(f, d+1)$. Ключевым здесь является следующее утверждение.

ЛЕММА 3. *Для любой функции $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ справедлива оценка*

$$L(f, n-1) \geq \frac{1}{\sqrt{1 - \widehat{f}(\{1, \dots, n\})^2}}.$$

ДОКАЗАТЕЛЬСТВО. Положим $\alpha = \widehat{f}(\{1, \dots, n\})^2$. Из равенства Парсеваля (4.2) имеем $|\widehat{f}(S)| \leq \sqrt{1-\alpha}$ для всякого подмножества $S \subsetneq \{1, \dots, n\}$, откуда искомая оценка следует по теореме 4. Лемма доказана.

Теперь докажем основной результат о длине, который влечет за собой теорему 3 из введения.

ТЕОРЕМА 9. *Для всех n и всякого $d = 1, 2, \dots, n-1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

$$2^{(d-1)/2} \leq L(f, d) < 2^{d+1}, \quad (7.1)$$

$$L(f, d+1) \leq d+3. \quad (7.2)$$

ДОКАЗАТЕЛЬСТВО. Функция f будет зависеть только от первых $d+1$ переменных. А именно, положим

$$f(x) = \begin{cases} -1, & \text{если } x_1 = \cdots = x_{d+1} = 1, \\ x_1 \cdots x_{d+1} & \text{иначе.} \end{cases} \quad (7.3)$$

Тогда

$$\widehat{f}(\{1, \dots, d+1\}) = 1 - 2^{-d}. \quad (7.4)$$

В силу леммы 3 имеем искомую нижнюю оценку $L(f, d)$. Поскольку выполнено $|\widehat{f}(\{1, \dots, d+1\})| < 1$, из разложения (4.1) получаем

$$f(x) = \operatorname{sgn} \left\{ \sum_{S \subsetneq \{1, \dots, d+1\}} \widehat{f}(S) \chi_S(x) \right\},$$

что доказывает верхнюю оценку $L(f, d)$. Наконец, представление f в виде

$$f(x) \equiv \operatorname{sgn}\{(d+1-x_1-x_2-\cdots-x_{d+1})x_1x_2\cdots x_{d+1}-1\} \quad (7.5)$$

влечет за собой требуемую оценку $L(f, d+1)$. Теорема доказана.

Теперь докажем аналогичную теорему для весов пороговых элементов. Самая сложная часть, нижняя оценка на веса пороговых элементов степени d , тривиально следует из предыдущей теоремы, поскольку $W(f, d) \geq L(f, d)$. Однако мы покажем, как другим методом можно получить точную нижнюю оценку, и вдобавок получим нижнюю оценку на $W^*(f, d)$. В этом нам поможет следующее наблюдение о равномерном приближении на булевом кубе.

ТЕОРЕМА 10. Пусть дана функция $f: \{-1, +1\}^n \rightarrow \mathbb{R}$. Тогда

$$E(f, n-1) = |\widehat{f}(\{1, \dots, n\})|.$$

ДОКАЗАТЕЛЬСТВО. Для доказательства верхней оценки достаточно взять в качестве приближающего многочлена $\sum_{|S| < n} \widehat{f}(S) \chi_S$. Для нижней оценки рассмотрим произвольный многочлен p степени не выше $n-1$. Тогда

$$\begin{aligned} \max_{x \in \{-1, +1\}^n} |f(x) - p(x)| &= \max_{x \in \{-1, +1\}^n} |f(x)x_1 \cdots x_n - p(x)x_1 \cdots x_n| \\ &\geq \left| \sum_{x \in \{-1, +1\}^n} [f(x)x_1 \cdots x_n - p(x)x_1 \cdots x_n] \right| = |\widehat{f}(\{1, \dots, n\})|. \end{aligned}$$

Теорема доказана.

ТЕОРЕМА 11. Для всех n и всякого $d = 1, 2, \dots, n-1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$, для которой верны оценки (7.1) и (7.2), а также соотношения

$$W^*(f, d) = W(f, d) = 2 \cdot 2^d - 1, \quad (7.6)$$

$$W^*(f, d+1) \leq W(f, d+1) \leq 2d+3. \quad (7.7)$$

ДОКАЗАТЕЛЬСТВО. В качестве f возьмем функцию (7.3), для которой оценки (7.1) и (7.2) уже были получены в теореме 9. Из теоремы 10 и равенства (7.4) следует, что $E(f, d) = 1 - 2^{-d}$, а значит, $W^*(f, d) \geq 2 \cdot 2^d - 1$ по лемме 2.

Поскольку $|\widehat{f}(\{1, \dots, d+1\})| < 1$, имеем

$$f(x) \equiv \operatorname{sgn} \left(\sum_{S \subsetneq \{1, \dots, d+1\}} 2^d \widehat{f}(S) \chi_S(x) \right). \quad (7.8)$$

По определению каждая величина $2^d \widehat{f}(S)$ – целое число. Следовательно, f вычислима пороговым элементом степени d с весом не более

$$2^d \sum_{S \subsetneq \{1, \dots, d+1\}} |\widehat{f}(S)| \leq 2^d \sqrt{2^{d+1} - 1} \sqrt{1 - \widehat{f}(\{1, \dots, d+1\})^2} < 2 \cdot 2^d,$$

где первый переход следует из неравенства Коши–Буняковского и теоремы Парсеваля (4.2). Так как последнее неравенство строгое, а $W(f, d)$ – целое число, получаем требуемую верхнюю оценку для $W(f, d)$, совпадающую с нижней оценкой для $W^*(f, d)$.

Наконец, представление f в виде (7.5) влечет требуемую оценку на $W(f, d+1)$. Теорема доказана.

8. Основной результат и обобщения. Из теорем 8 и 11 немедленно получается теорема 1. Остается доказать более сильную теорему 2.

ТЕОРЕМА 2 (переформулирована). *Для всех n и всех $d = 1, 2, \dots, n-1$ существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

- (1) *всякая функция голосования, примененная к произвольным элементам входной степени не выше d и вычисляющая f , должна применяться к не менее чем $\exp\{\Theta(n)\}$ элементам, что также является достаточным;*
- (2) *f вычислима функцией голосования от $O(n^2)$ функций логического сложения $\oplus$ входной степени не выше $d+1$, а также функцией голосования от $O(n^2d)$ функций голосования входной степени не выше $d+1$.*

ДОКАЗАТЕЛЬСТВО. Сначала рассмотрим случай $1 \leq d \leq n/5$. Возьмем в качестве f функцию из теоремы 8. Из равенства (6.1) и леммы 2 следует, что

$$W(f, d) = \Theta(n2^{(n-d)/2}) \quad \text{и} \quad W^*(f, d) \geq \Omega(n2^{n/10}),$$

что устанавливает первое свойство f из формулировки теоремы. Напомним также, что по построению f вычислима функцией голосования от $O(n^2)$ функций логического сложения от не более чем $(d+1)$ переменных каждая. Поскольку всякая функция логического сложения от не более чем $d+1$ переменных выражается в виде вещественной суммы $O(d)$ функций голосования (см. Хайнал и др. [35; предложение 2.1] и Брук [24; § 2]), второе свойство также верно.

Рассмотрим теперь оставшийся случай $n/5 < d < n$. Возьмем в качестве f функцию, построенную в теореме 11. Тогда первое свойство немедленно следует из (7.6), а второе свойство следует из (7.7) и упомянутого выше выражения функции логического сложения через сумму функций голосования.

В свете результатов об аффинно разреженном спектре из п. 1 теоремы 1 и 2 допускают различные обобщения. Мы ограничимся одним показательным примером таких обобщений.

ТЕОРЕМА 12. *Для всех n и всех $d = 1, 2, \dots, n-1$ и $k = 1, 2, \dots, \max\{d, n-d\}$, существует (явно заданная) функция $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ такая, что*

$$W(f, d) = W(f, d+1) = \dots = W(f, d+k-1) = 2^{\Theta(n/k)},$$

$$W(f, d+k) = O\left(\frac{n^2}{k^2}\right).$$

ДОКАЗАТЕЛЬСТВО. Положим $l_1 = \lfloor d/k \rfloor$, $l_2 = \lfloor (n-d)/k \rfloor$ и $l = l_1 + l_2$. Заметим, что $l = \Theta(n/k)$.

Поскольку $k \leq \max\{d, n-d\}$, имеем $1 \leq l_1 \leq l-1$. Как следствие, теорема 1 дает функцию $g: \{-1, +1\}^l \rightarrow \{-1, +1\}$ такую, что

$$W(g, l_1) = 2^{\Theta(l)}, \tag{8.1}$$

$$W(g, l_1 + 1) = O(l^2). \tag{8.2}$$

Определим

$$f(x) = x_{kl+1}x_{kl+2} \cdots x_{kl+(d-kl_1)} \cdot g\left(\prod_{i=1}^k x_i, \dots, \prod_{i=1}^k x_{(l-1)k+i}\right).$$

Ввиду следствий 1 и 2 требуемые свойства функции f следуют из соответствующих свойств (8.1) и (8.2) функции g .

9. Открытые вопросы. Рассмотрим булеву функцию $f: \{-1, +1\}^n \rightarrow \{-1, +1\}$ и соответствующие величины

$$W(f, d), \quad d = 0, 1, 2, \dots, n.$$

Какую информацию содержат эти числа об f ? Ясно, что функцию f нельзя однозначно восстановить из (9), поскольку эти величины неизменны при переходе к $-f$, а также при перестановке переменных. Однако, содержательное спектральное описание f в терминах (9) должно быть возможно.

Например, представляется вероятным, что экспоненциальный разрыв между $W(f, d)$ и $W(f, d+1)$ предполагает спектральную близость f к функции логического сложения $\oplus$ на $d+1$ переменной. В частности, экспоненциальный разрыв между $W(f, d)$ и $W(f, d+1)$, вероятно, невозможен, если f – линейный пороговый элемент и d велико, скажем, $d = n^\epsilon$. Верна ли эта гипотеза? Заметим, что эта интуиция полностью соответствует доказательству теоремы 1.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

- [1] R. Beigel, “The polynomial method in circuit complexity”, *Proceedings of the Eighth Annual Structure in Complexity Theory Conference* (San Diego, CA, 1993), IEEE Comput. Soc. Press, Los Alamitos, CA, 1993, 82–95.
- [2] M. E. Saks, “Slicing the hypercube”, *Surveys in Combinatorics*, 1993 (Keele), London Math. Soc. Lecture Note Ser., **187**, Cambridge Univ. Press, Cambridge, 1993, 211–255.
- [3] H. Buhrman, R. de Wolf, “Complexity measures and decision tree complexity: a survey”, *Theoret. Comput. Sci.*, **288**:1 (2002), 21–43.
- [4] A. A. Sherstov, “Communication lower bounds using dual polynomials”, *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, **95** (2008), 59–93.
- [5] М. Минский, С. Пейперт, *Перцептроны*, Мир, М., 1971.
- [6] R. Paturi, M. E. Saks, “Approximating threshold circuits by rational functions”, *Inform. and Comput.*, **112**:2 (1994), 257–272.
- [7] K.-Y. Siu, V. P. Roychowdhury, T. Kailath, “Rational approximation techniques for analysis of neural networks”, *IEEE Trans. Inform. Theory*, **40**:2 (1994), 455–466.
- [8] J. Aspnes, R. Beigel, M. Furst, S. Rudich, “The expressive power of voting polynomials”, *Combinatorica*, **14**:2 (1994), 135–148.
- [9] R. Beigel, N. Reingold, D. Spielman, “PP is closed under intersection”, *J. Comput. System Sci.*, **50**:2 (1995), 191–202.
- [10] M. Krause, P. Pudlák, “On the computational power of depth-2 circuits with threshold and modulo gates”, *Theoret. Comput. Sci.*, **174**:1–2 (1997), 137–156.
- [11] A. A. Sherstov, “Separating AC^0 from depth-2 majority circuits”, *STOC’07 – Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, ACM, New York, NY, 2007, 294–301.
- [12] A. A. Sherstov, “The pattern matrix method for lower bounds on quantum communication”, *STOC’08 – Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, ACM, New York, NY, 2008, 85–94.
- [13] A. A. Sherstov, “The unbounded-error communication complexity of symmetric functions”, *Proceedings of the 49th Symposium on Foundations of Computer Science (FOCS)*, IEEE Comput. Soc. Press, 2008, 384–393.
- [14] A. A. Razborov, A. A. Sherstov, “The sign-rank of AC^0 ”, *Proceedings of the 49th Symposium on Foundations of Computer Science (FOCS)*, IEEE Comput. Soc. Press, 2008, 57–66.

- [15] A. R. Klivans, R. A. Servedio, “Learning DNF in time $2^{\tilde{O}(n^{1/3})}$ ”, *J. Comput. System Sci.*, **68**:2 (2004), 303–318.
- [16] R. O’Donnell, R. A. Servedio, “New degree bounds for polynomial threshold functions”, *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, ACM, New York, NY, 2003, 325–334.
- [17] A. R. Klivans, A. A. Sherstov, “Unconditional lower bounds for learning intersections of halfspaces”, *Mach. Learn.*, **69**:2–3 (2007), 97–114.
- [18] J. C. Jackson, *The Harmonic Sieve: A Novel Application of Fourier Analysis to Machine Learning Theory and Practice*, PhD thesis, Carnegie Mellon Univ., Pittsburgh, PA, 1995.
- [19] A. R. Klivans, R. O’Donnell, R. A. Servedio, “Learning intersections and thresholds of halfspaces”, *J. Comput. System Sci.*, **68**:4 (2004), 808–840.
- [20] A. R. Klivans, R. A. Servedio, “Toward attribute efficient learning of decision lists and parities”, *J. Mach. Learn. Res.*, **7** (2006), 587–602.
- [21] R. Beigel, “Perceptrons, PP, and the polynomial hierarchy”, *Comput. Complexity*, **4**:4 (1994), 339–349.
- [22] N. K. Vereshchagin, “Lower bounds for perceptrons solving some separation problems and oracle separation of AM from PP”, *Third Israel Symposium on the Theory of Computing and Systems* (Tel Aviv, 1995), IEEE Comput. Soc. Press, Los Alamitos, CA, 1995, 46–51.
- [23] H. Buhrman, N. K. Vereshchagin, R. de Wolf, “On computation and communication with small bias”, *Proceedings of the 22nd Annual IEEE Conference on Computational Complexity (CCC)*, IEEE Comput. Soc. Press, 2007, 24–32.
- [24] J. Bruck, “Harmonic analysis of polynomial threshold functions”, *SIAM J. Discrete Math.*, **3**:2 (1990), 168–177.
- [25] K.-Y. Siu, J. Bruck, “On the power of threshold circuits with small weights”, *SIAM J. Discrete Math.*, **4**:3 (1991), 423–435.
- [26] J. Bruck, R. Smolensky, “Polynomial threshold functions, AC^0 functions, and spectral norms”, *SIAM J. Comput.*, **21**:1 (1992.), 33–42.
- [27] M. Goldmann, J. Hästad, A. A. Razborov, “Majority gates vs. general weighted threshold gates”, *Comput. Complexity*, **2**:4 (1992), 277–300.
- [28] M. Krause, P. Pudlák, “Computing Boolean functions by polynomials and threshold circuits”, *Comput. Complexity*, **7**:4 (1998), 346–370.
- [29] M. Krause, “On the computational power of Boolean decision lists”, *Comput. Complexity*, **14**:4 (2005), 362–375.
- [30] M. Goldmann, “On the power of a threshold gate at the top”, *Inform. Process. Lett.*, **63**:6 (1997), 287–293.
- [31] J. Hästad, “On the size of weights for threshold gates”, *SIAM J. Discrete Math.*, **7**:3 (1994), 484–492.
- [32] В. В. Подольский, “Перцептроны с большим весом”, *Пробл. передачи информ.*, **45**:1 (2009), 51–59.
- [33] V. V. Podolskii, “A uniform lower bound on weights of perceptrons”, *Computer Science – Theory and Applications*, Lecture Notes in Comput. Sci., **5010**, Springer-Verlag, Berlin, 2008, 261–272.
- [34] J. Myhill, W. H. Kautz, “On the size of weights required for linear-input switching functions”, *IEEE Trans. Electron. Comput.*, **10**:2 (1961), 288–290.
- [35] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, G. Turán, “Threshold circuits of bounded depth”, *J. Comput. System Sci.*, **46**:2 (1993), 129–154.

В. В. Подольский

Московский государственный университет
им. М. В. Ломоносова
E-mail: podolskii@lpcs.math.msu.su

Поступило

20.06.2009

А. А. Шерстов

University of Texas at Austin
E-mail: sherstov@cs.utexas.edu