Оценка количества классов τ-Inf-эквивалентности для пороговых функций

И. В. Грибушин (МГУ им. М. В. Ломоносова, Москва)

В работе исследуются относительные влияния переменных булевой функции. Множество булевых функций разбивается на классы τ-Inf-эквивалентности в зависимости от максимального относительного влияния переменных. Приводятся нижняя и верхняя оценки количества классов τ-Inf-эквивалентности для пороговых функций. Они равны $2^{n/2}$ и $n2^{2n}$.

Ключевые слова: пороговые функции, влияние переменных булевой функции, относительное влияние переменных булевой функции, классы τ-Inf-эквивалентности.

Рассмотрим булеву функцию $f : \{-1, 1\}^n \to \{-1, 1\}$ от n переменных.

Обозначим I_f — множество, на котором $f = 1$. $[n] := \{1, \ldots, n\}$, $[k, n] := \{k, \ldots, n\}$.

Определение 1. [?] Характеристической функцией множества $S \subseteq [n]$ называется:

$$\chi_S := \prod_{i \in S} x_i, \quad \chi_\varnothing := 1.$$

Определение 2. [?] Пусть x равномерно распределено на пространстве $\{-1, 1\}^n$. Коэффициентом Фурье функции f относительно $S \subseteq [n]$ называется:

$$\hat{f}(S) := \hat{f}_S := \mathbf{E} f(x) \chi_S(x).$$

Из равенства Парсеваля выводится следующее утверждение.

Утверждение 1. [?]

$$\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1.$$

Если $x = (x_1, \dots, x_i, \dots, x_n)$, то обозначим

$$x^{\oplus i} := (x_1, \dots, -x_i, \dots, x_n).$$

Также если $x = (x_1, \dots, x_i, \dots, x_j, \dots, x_n)$, $i < j$, то обозначим

$$x^{i \leftrightarrow j} := (x_1, \dots, x_j, \dots, x_i, \dots, x_n).$$

Определение 3. [?] Влиянием i-ой переменной на функцию f называется:

$$Inf_i := \mathbf{Pr}_{x \in \{-1,1\}^n}[f(x) \neq f(x^{\oplus i})].$$

Определение 4. Функции f и $\tilde{f}$ называются эквивалентными, если они имеют одинаковые множества влияний переменных:

$$\{Inf_i(f) | i \in [n]\} = \{Inf_i(\tilde{f}) | i \in [n]\}.$$

Замечание 1. Следующие три операции сохраняют эквивалентность в смысле определения 4:

1) $f \sim \tilde{f} = -f$;
2) $f \sim \tilde{f} = f(x^{\oplus i})$;
3) $f \sim \tilde{f} = f(x^{i \leftrightarrow j})$.

Замечание 2. Принимая во внимание операцию 3, можно считать, что переменные любой функции f упорядочены по убыванию своих влияний:

$$Inf_1 \geqslant Inf_2 \geqslant \dots \geqslant Inf_n.$$

Замечание 3. [?]

$$Inf_i(f) = \sum_{i \in S \subseteq [n]} \hat{f}(S)^2.$$

Утверждение 2. [?] *Для любой монотонной булевой функции f имеет место:*

$$Inf_i = \hat{f}_{\{i\}}.$$

Определение 5. [?] Полным влиянием функции f называется:

$$Inf(f) := \sum_{i=1}^{n} Inf_i(f).$$

Определение 6. Относительным влиянием i-ой переменной на функцию f называется:

$$\tau_i = \frac{Inf_i(f)}{Inf(f)}.$$

Определение 7. [?] f называется τ-регулярной для некоторого $\tau > 0$, если $\forall i \in [n]$:

$$Inf_i(f) \leqslant \tau Inf(f).$$

Определение 8. Пусть $p : \{-1,1\}^n \to \mathbf{R}$ — линейный многочлен, $f = \operatorname{sign}(p)$, тогда функция $f : \{-1,1\}^n \to \{-1,1\}$ называется пороговой.

Из определения 7 имеем:

$$\max_{i\in[n]} Inf_i(f) \leqslant \tau Inf(f) \Rightarrow \max_{i\in[n]} \frac{Inf_i(f)}{Inf(f)} \leqslant \tau.$$

Так как для любой τ-регулярной функции имеет смысл рассматривать только наименьшее значение τ, получаем:

$$\boxed{\tau(f) = \max_{i\in[n]} \frac{Inf_i(f)}{Inf(f)}} \qquad (1)$$

Получили, что $\tau(f)$ является максимальным относительным влиянием среди всех относительных влияний переменных функции f.

Определение 9. Функции f и g называются τ-Inf-эквивалентными, если $\tau(f) = \tau(g)$.

Замечание 4. Из формулы (**??**) следует, что $\frac{1}{n} \leqslant \tau(f) \leqslant 1$, где n — количество существенных переменных у функции f.

Утверждение 3. [?] *Относительное влияние переменных любой пороговой функции, существенно зависящей от $n \geqslant 2$ переменных, лежит на отрезке* $\left[\frac{1}{n}, \frac{2^{n-1}-1}{2^{n-1}+n-2}\right]$.

Обозначим $M_\tau(n), T_\tau(n)$ и $B_\tau(n)$ количество классов τ-Inf-эквивалентности для монотонных, пороговых и булевых функций от n переменных соответственно.

Замечание 5. Так как множество монотонных функций является подмножеством множества пороговых функций, и для любой пороговой функции f существует эквивалентная ей монотонная пороговая функция (см. [?]), значит, множество значений τ на монотонных функциях равно множеству значений τ на пороговых функциях, следовательно, $T_\tau(n) = M_\tau(n)$.

Замечание 6. Так как множество пороговых функций является подмножеством множества булевых функций, то $T_\tau(n) \leqslant B_\tau(n)$.

Утверждение 4. [?] *Множество пороговых функций от* 3 *переменных разбивается на* 5 *классов* τ-Inf*-эквивалентности:* $\left\{0, \frac{1}{3}, \frac{1}{2}, \frac{3}{5}, 1\right\}$, *а множество пороговых функций от* 4 *переменных — на* 10 *классов* τ-Inf*-эквивалентности:* $\left\{0, \frac{1}{4}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{5}{12}, \frac{1}{2}, \frac{3}{5}, \frac{7}{10}, 1\right\}$.

Утверждение 5. $B_\tau(n+1) \leqslant n2^{2n}$.

Доказательство. Из определения 3 следует, что Inf_i может принимать значения вида $a/2^n$, где $a \in [0, 2^n]$. Воспользуемся тем, что Inf_i можно упорядочить по убыванию (см. замечание 2), тогда $Inf_1 \in \left\{\frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n}{2^n}\right\}$ и $Inf \in \left\{\frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{n2^n}{2^n}\right\}$. Значит, дробь $\tau = \frac{Inf_1}{Inf}$ может принимать не более $n2^{2n}$ различных значений. Утверждение доказано.

Теорема 1. $T_\tau(n+1) \geqslant 2^{[n/2]}$.

Доказательство. Пусть $g_j(x_1, \ldots, x_n)$ и $f_j(x_1, \ldots, x_{n+1})$ — функции от n и $n+1$ переменной соответственно, где $j \in \left[2^k\right], k \in [n]$.

$$g_j(x_1, \ldots, x_n) = \begin{cases} 1, & \text{если } \sum_{i=1}^{n}(x_i+1)2^{i-2} < j; \\ -1, & \text{иначе.} \end{cases}$$

$$f_j(x_1, \ldots, x_{n+1}) = \begin{cases} 1, & \text{если } x_{n+1} = 1; \\ g_j(x_1, \ldots, x_n), & \text{если } x_{n+1} = -1. \end{cases}$$

Из способа задания функций g_j и f_j следует, что все они являются пороговыми. Найдём влияния переменных для функции f_j:

$$Inf_i(f_j) = \frac{Inf_i(g_j)}{2} \leqslant \frac{1}{2}, \text{где } i \in [n];$$

$$Inf_{n+1}(f_j) = \frac{2^{n+1} - 2|I_{g_j}|}{2^{n+1}} = \frac{2^n - |I_{g_j}|}{2^n} \geqslant \frac{2^n - 2^k}{2^n} = 1 - 2^{k-n}.$$

Если $k \leqslant n-1$, то $Inf_{n+1}(f_j) \geqslant 1 - 2^{-1} = \frac{1}{2} \geqslant Inf_i(f_j)$, где $i \in [n]$, следовательно, получаем, что $\tau(f_j) = \max_{i\in[n+1]} \frac{Inf_i(f_j)}{Inf(f_j)} = \frac{Inf_{n+1}(f_j)}{Inf(f_j)}$. Посмотрим, как меняется $\tau(f_j)$ при увеличении j на 1:

$$
\begin{aligned}
|I_{g_{j+1}}| &= |I_{g_j}| + 1;\\
Inf_i(f_{j+1}) &= Inf_i(f_j) \pm \frac{1}{2^n}, \text{где } i \in [k];\\
Inf_i(f_{j+1}) &= Inf_i(f_j) + \frac{1}{2^n}, \text{где } i \in [k+1, n];\\
Inf_{n+1}(f_{j+1}) &= \frac{2^n - |I_{g_{j+1}}|}{2^n} = \frac{2^n - |I_{g_j}| - 1}{2^n} = Inf_{n+1}(f_j) - \frac{1}{2^n}.
\end{aligned}
$$

Сравним $\tau(f_j)$ и $\tau(f_{j+1})$:

$$
\begin{aligned}
\tau(f_j) - \tau(f_{j+1}) = \frac{Inf_{n+1}(f_j)}{Inf(f_j)} - \frac{Inf_{n+1}(f_{j+1})}{Inf(f_{j+1})} = \\
= \frac{Inf_{n+1}(f_j)Inf(f_{j+1}) - Inf_{n+1}(f_{j+1})Inf(f_j)}{Inf(f_j)Inf(f_{j+1})}.
\end{aligned}
$$

Так как $Inf(f_j)Inf(f_{j+1}) > 0$, оставим только числитель:

$$
\begin{aligned}
&Inf_{n+1}(f_j)Inf(f_{j+1}) - Inf_{n+1}(f_{j+1})Inf(f_j) =\\
&\quad= Inf_{n+1}(f_j)\left(\sum_{i=1}^{n} Inf_i(f_{j+1}) + Inf_{n+1}(f_{j+1})\right) -\\
&\quad- Inf_{n+1}(f_{j+1})\left(\sum_{i=1}^{n} Inf_i(f_j) + Inf_{n+1}(f_j)\right) =\\
&= Inf_{n+1}(f_j)\sum_{i=1}^{n} Inf_i(f_{j+1}) - Inf_{n+1}(f_{j+1})\sum_{i=1}^{n} Inf_i(f_j) =\\
&= \left(Inf_{n+1}(f_{j+1}) + \frac{1}{2^n}\right)\sum_{i=1}^{n} Inf_i(f_{j+1}) - Inf_{n+1}(f_{j+1})\sum_{i=1}^{n} Inf_i(f_j) =\\
&= Inf_{n+1}(f_{j+1})\sum_{i=1}^{n}(Inf_i(f_{j+1}) - Inf_i(f_j)) + \frac{1}{2^n}\sum_{i=1}^{n} Inf_i(f_{j+1}).
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{i=1}^{n}(Inf_i(f_{j+1}) - Inf_i(f_j)) =\\
&\quad= \sum_{i=1}^{k}(Inf_i(f_{j+1}) - Inf_i(f_j)) + \sum_{i=k+1}^{n}(Inf_i(f_{j+1}) - Inf_i(f_j)) =
\end{aligned}
$$

$$= \sum_{i=1}^{k}(Inf_i(f_{j+1}) - Inf_i(f_j)) + \sum_{i=k+1}^{n}\frac{1}{2^n} \geqslant$$

$$\geqslant \sum_{i=1}^{k}\frac{-1}{2^n} + \frac{n-k}{2^n} = \frac{-k}{2^n} + \frac{n-k}{2^n} = \frac{n-2k}{2^n}.$$

Таким образом, мы показали, что $\tau(f_j) - \tau(f_{j+1}) > 0$ при $k \leqslant [n/2]$. То есть для различных f_j, где $j \in [2^k]$, будут получаться различные $\tau(f_j)$. Значит, мы построили последовательность из $2^{[n/2]}$ пороговых функций с различными значениями τ. Из этого следует, что $T_\tau(n+1) \geqslant 2^{[n/2]}$. Теорема доказана.

Соединив вместе замечания 5, 6, утверждение (**??**) и теорему (**??**), получаем следующие оценки количества классов τ-Inf-эквивалентности:

$$2^{[n/2]} \leqslant M_\tau(n+1) = T_\tau(n+1) \leqslant B_\tau(n+1) \leqslant n2^{2n}.$$

Автор выражает благодарность своему научному руководителю А. А. Ирматову за внимание к работе и полезные обсуждения.

Список литературы

[1] Ben-Or M., Linial N. Collective coin flipping // Randomness and computation. — Academic Press, 1990. — 5. — P. 91–115.

[2] Diakonikolas I., Harsha P., Klivans A., Meka R., Raghavendra P., Servedio R., Li-Yang Tan. Bounding the average sensitivity and noise sensitivity of polynomial threshold functions // Proceedings of the 42nd ACM symposium on Theory of computing. — ACM, 2010. — P. 533–542.

[3] Kahn J., Kalai G., Linial N. The Influence of Variables on Boolean Functions // Proceedings of the 29th Annual Symposium on Foundations of Computer Science (FOCS). — 1988. — P. 68–80.

[4] Muroga S. Threshold logic and its applications. — New York: Wiley-Interscience, 1971.

[5] O’Donnell R. Analysis of Boolean functions. — Cambridge: Cambridge University Press, 2014.

[6] Грибушин И. В. О возможных значениях максимума относительного влияния переменных для пороговых функций // Интеллектуальные системы. Теория и приложения. — Т. 20, вып. 1. — 2016. — С. 195–212.