EuroComb 2005 DMTCS proc. **AE**, 2005, 11–16

Monotone Boolean Functions with s Zeros Farthest from Threshold Functions

Kazuyuki Amano[1] and Jun Tarui[2]

[1] *Graduate School of Information Sciences, Tohoku University, Sendai 980-8579 Japan*
ama@ecei.tohoku.ac.jp
[2] *Department of Information and Communication Engineering, University of Electro-Communications Chofu, Tokyo 182-8585 Japan* tarui@ice.uec.ac.jp

Let T_t denote the t-threshold function on the n-cube: $T_t(x) = 1$ if $|\{i : x_i = 1\}| \geq t$, and 0 otherwise. Define the distance between Boolean functions g and h, $d(g, h)$, to be the number of points on which g and h disagree. We consider the following extremal problem: Over a monotone Boolean function g on the n-cube with s zeros, what is the maximum of $d(g, T_t)$? We show that the following monotone function p_s maximizes the distance: For $x \in \{0, 1\}^n$, $p_s(x) = 0$ if and only if $N(x) < s$, where $N(x)$ is the integer whose n-bit binary representation is x. Our result generalizes the previous work for the case $t = \lceil n/2 \rceil$ and $s = 2^{n-1}$ by Blum, Burch, and Langford [BBL98-FOCS98], who considered the problem to analyze the behavior of a learning algorithm for monotone Boolean functions, and the previous work for the same t and s by Amano and Maruoka [AM02-ALT02].

1 Introduction and Overview

For a Boolean function h and a class $\mathcal{C}$ of Boolean functions, we consider the following extremal problem: what is the maximum distance between $g \in \mathcal{C}$ and h? Equivalently, under the uniform distribution on $\{0, 1\}^n$, how small can the correlation between $g \in \mathcal{C}$ and h be? The distance between Boolean functions g and h, $d(g, h)$, is defined to be the number of points on which g and h disagree. A Boolean function $g : \{0, 1\}^n \to \{0, 1\}$ is *monotone* if, for $x, y \in \{0, 1\}^n$, $x \leq y \Rightarrow g(x) \leq g(y)$, where for $x, y \in \{0, 1\}^n$, $x \leq y$ if and only if $x_i \leq y_i$ for all $i = 1, \ldots, n$. A Boolean function is *fair* if it outputs 1 on exactly half of its inputs. The starting point of our work is the fact that *among all fair monotone Boolean functions, a single variable function $g(x) = x_i$ is farthest from the majority function*; this was conjectured by Blum, Burch, and Langford [BBL98], and was proved by Amano and Maruoka [AM02].

The main concern of the work of Blum, Burch, and Langford is *learning* of monotone Boolean functions. They gave the following simple algorithm for weakly learning a monotone Boolean function under the uniform distribution: Given samples $(x_1, g(x_1)), (x_2, g(x_2)), \ldots, (x_m, g(x_m))$, output, as a hypothesis, a function that is most correlated with those samples among *three* functions $\{0, 1, \mathrm{Majority}\}$, where 0 and 1 are constant functions. With high probability the output of the algorithm has correlation at least $\Omega(1/\sqrt{n})$ with g. Blum, Burch, and Langford showed, using the Kruskal-Katona theorem about the minimum size of a shadow, that any fair monotone Boolean function g has correlation at least $\Omega(1/\sqrt{n})$ with Majority. They conjectured that in fact a single variable function $g(x) = x_i$ is a fair monotone function that is farthest from Majority. Amano and Maruoka proved this conjecture also using the Kruskal-Katona theorem.

In this paper we give a generalization in which we consider any threshold function, not just Majority, and any monotone function with a prescribed number of zeros, not just a fair one. Our proof is self-contained; we do not use the Kruskal-Katona theorem.

Let T_t denote the t-threshold function: for $x \in \{0,1\}^n$, $T_t(x) = 1$ if $|\{i : x_i = 1\}| \geq t$, and 0 otherwise. Throughout the paper t is an integer; it will be convenient to allow t to be negative; for $t \leq 0$, T_t is the constant 1 function on the n-cube. The majority function is function T_t with $t = \lceil n/2 \rceil$. For a Boolean function g, let $\sharp_0(g)$ and $\sharp_1(g)$ respectively denote the number of points on which $g = 0$ and on which $g = 1$. Similarly, for Boolean functions g and h, and $a, b \in \{0,1\}$, let $\sharp_{ab}(g,h)$ denote the number of points x such that $g(x) = a$ and $h(x) = b$.

The problem we consider is the following: Among all monotone $g : \{0,1\}^n \to \{0,1\}$ with $\sharp_0(g) = s$ ($0 \leq s < 2^n$), what is the maximum of $d(g, T_t)$? Maximizing $d(g,h)$ for a fixed function h and a function g with $\sharp_0(g) = s$ is equivalent to maximizing $\sharp_{01}(g,h)$ since for any g and h, $d(g,h) = \sharp_0(h) - \sharp_0(g) + 2\sharp_{01}(g,h)$. For $0 \leq s < 2^n$, define the monotone function $p_s : \{0,1\}^n \to \{0,1\}$ as follows. For $x \in \{0,1\}^n$, $p_s(x) = 0$ if and only if $N(x) < s$, where $N(x)$ is the integer whose n-bit binary representation is x.

Theorem Let $g : \{0,1\}^n \to \{0,1\}$ be monotone with $\sharp_0(g) = s$. Then, for any integer t,

$$\sharp_{01}(g, T_t) \leq \sharp_{01}(p_s, T_t), \quad \text{and hence} \quad d(g, T_t) \leq d(p_s, T_t).$$

2 Proof of Theorem

For an integer $i \geq 0$, let $b(i)$ be the number of 1's in the binary representation of the integer i. Let $0 \leq l \leq m$. For an integer t, define $f_t(l,m)$ to be the number of integers i such that $l \leq i < m$ and $b(i) \geq t$. Note that for $t \leq 0$, $f_t(l,m) = f_0(l,m) = m - l$. Also note that the following hold: $f_t(l,m) \leq f_{t-1}(l,m)$; for $0 \leq k < 2^r$, $f_t(0,k) = f_{t-1}(2^r, 2^r + k)$; $\sharp_{01}(p_s, T_t) = f_t(0,s)$.

It turns out that the main work we do to prove the theorem is an analysis of $f_t(l,m)$. This aspect of our proof is somewhat similar to a proof of the edge-isoperimetric inequality on the Boolean cube explained in the book by Bollobas [Bo86, § 16 Isoperimetric Problems]. We now state a key lemma, Lemma 1, and two auxiliary lemmas, Lemmas 2 and 3. We wll give a proof of the theorem using Lemma 1, and then give proofs of the three lemmas.

Lemma 1 For $0 \leq l \leq m$ and any integer t,

$$f_t(0,m) + f_{t-1}(0,l) \leq f_t(0, m+l).$$

Lemma 2 For $k, l \geq 0$ and any integer t,

$$f_t(0,k) \leq f_t(l, l+k).$$

Lemma 3 For $k, l, q \geq 0$ such that $l + k \leq 2^q$ and for any integer t,

$$f_t(l, l+k) \leq f_t(2^q - k, 2^q).$$

Proof of Theorem using Lemma 1. The proof is by induction on n. The base case $n = 1$ is trivial. For induction assume that $n > 1$ and that the assertion holds for $n-1$. Let g_0 and g_1 be the Boolean functions on the $(n-1)$-cube obtained from g by fixing the first bit to be 0 and 1 respectively; i.e., for $e = 0, 1$, and for $x \in \{0,1\}^{n-1}$, $g_e(x) = g(ex)$. Let $m = \sharp_0(g_0)$ and $l = \sharp_0(g_1)$. Clearly, $m + l = s$, and since g is monotone, $m \geq l$. Thus we have

$$\begin{aligned}
\sharp_{01}(g, T_t) &= \sharp_{01}(g_0, T_t) + \sharp_{01}(g_1, T_{t-1}) \\
&\leq \sharp_{01}(p_m, T_t) + \sharp_{01}(p_l, T_{t-1}) \\
&= f_t(0, m) + f_{t-1}(0, l) \\
&\leq f_t(0, m + l) \\
&= \sharp_{01}(p_s, T_t),
\end{aligned}$$

where the first inequality is by the inductive assumption and the second inequality is by Lemma 1. □

Proofs of Lemmas 2 and 3. For $0 \leq i < 2^q$, the q-bit binary representation of i has bit 1 at position j ($1 \leq j \leq q$) if and only if the q-bit binary representation of $2^q - 1 - i$ has bit 0 at position j. Hence Lemma 2 readily yields Lemma 3.

Now we prove Lemma 2. The assertion is trivial when $l = k = 0$. Assume that $l + k \geq 1$ and let $r = \lfloor \log_2(l + k) \rfloor$ so that we have $2^r \leq l + k < 2^{r+1}$. The proof is by induction on r; more precisely, we prove Lemma 2 by inductively assuming that the assertion of Lemma 2 holds *and* the corresponding assertion of Lemma 3 holds.

In the base case when $r = 0$, we have $l + k = 1$ and thus either (i) $l = 0, k = 1$ or (ii) $l = 1, k = 0$; in both cases the claim is immediate. For induction assume that $r > 0$ and that for $r - 1$ the assertion holds, and hence the corresponding assertion of Lemma 3 also holds.

CASE 1: $2^r \leq l$: In this case $2^r \leq l \leq l + k < 2^{r+1}$ and

$$f_t(l, l + k) = f_{t-1}(l - 2^r, l + k - 2^r) \geq f_{t-1}(0, k) \geq f_t(0, k),$$

where the first inequality is by the inductive assumption.

CASE 2: $l < 2^r$ and $k < 2^r$:

$$\begin{aligned}
f_t(l, l + k) &= f_t(l, 2^r) + f_t(2^r, l + k) \\
&= f_t(l, 2^r) + f_{t-1}(0, l + k - 2^r) \\
&\geq f_t(l + k - 2^r, k) + f_{t-1}(0, l + k - 2^r) \\
&\geq f_t(l + k - 2^r, k) + f_t(0, l + k - 2^r) \\
&= f_t(0, k),
\end{aligned}$$

where the first inequality is by Lemma 3.

CASE 3: $l < 2^r$ and $k \geq 2^r$: In this case $l < 2^r \leq 2^r + l \leq l + k$ and

$$
\begin{aligned}
f_t(l, l+k) &= [\, f_t(l, 2^r) + f_t(2^r, 2^r + l)\,] + f_t(2^r + l, l+k) \\
&= [\, f_t(l, 2^r) + f_{t-1}(0, l)\,] + f_t(2^r + l, l+k) \\
&\geq [\, f_t(0, l) + f_t(l, 2^r)\,] + f_t(2^r + l, l+k) \\
&= f_t(0, 2^r) + f_t(2^r + l, l+k) \\
&= f_t(0, 2^r) + f_{t-1}(l, l+k-2^r) \\
&\geq f_t(0, 2^r) + f_{t-1}(0, k-2^r) \\
&= f_t(0, 2^r) + f_t(2^r, k) \\
&= f_t(0, k),
\end{aligned}
$$

where the second inequality is by the inductive assumption. □

Proof of Lemma 1. The asertion is trivial when $l = m = 0$. Asume that $m \geq 1$ and let $r = \lfloor \log_2 m \rfloor$ so that we have $2^r \leq m < 2^{r+1}$. The proof is by induction on r. In the base case when $r = 0$ we have $m = 1$, and $l = 0$ or $l = 1$; in both cases the claim is immediate. For induction assume that $r > 0$ and that the claim holds for $r - 1$.

CASE 1: $2^r \leq l$:

$$
\begin{aligned}
f_t(0, m) + f_{t-1}(0, l) &= f_t(0, 2^r) + f_t(2^r, m) + f_{t-1}(0, 2^r) + f_{t-1}(2^r, l) \\
&= f_t(0, 2^r) + f_{t-1}(0, m-2^r) + f_{t-1}(0, 2^r) + f_{t-2}(0, l-2^r) \\
&= f_t(0, 2^r) + f_{t-1}(0, 2^r) + f_{t-1}(0, m-2^r) + f_{t-2}(0, l-2^r) \\
&\leq f_t(0, 2^r) + f_{t-1}(0, 2^r) + f_{t-1}(0, m+l-2^{r+1}) \\
&= f_t(0, m+l),
\end{aligned}
$$

where the inequality is by the inductive assumption.

CASE 2: $l < 2^r$, $(m - 2^r) + l \leq 2^r$:

$$
\begin{aligned}
f_t(0, m) + f_{t-1}(0, l) &= f_t(0, 2^r) + f_t(2^r, m) + f_{t-1}(0, l) \\
&= f_t(0, 2^r) + f_{t-1}(0, m-2^r) + f_{t-1}(0, l) \\
&\leq f_t(0, 2^r) + f_{t-1}(0, m-2^r) + f_{t-1}(m-2^r, m-2^r+l) \\
&= f_t(0, 2^r) + f_{t-1}(0, m-2^r+l) \\
&= f_t(0, 2^r) + f_t(2^r, m+l) \\
&= f_t(0, m+l),
\end{aligned}
$$

where the inequality is by Lemma 2.

CASE 3: $l < 2^r$, $m + l > 2^{r+1}$: We have the following derivation where Lemma 3 is used in the form (1) for the inequality below.

$$f_{t-1}((m+l) - 2^{r+1}, l) \leq f_{t-1}(m - 2^r, 2^r) \quad (1)$$

$$\begin{aligned}
& f_t(0,m) + f_{t-1}(0,l) \\
= \;& f_t(0,2^r) + f_t(2^r,m) + f_{t-1}(0,(m+l)-2^{r+1}) + f_{t-1}((m+l)-2^{r+1},l) \\
= \;& f_t(0,2^r) + f_{t-1}(0,m-2^r) + f_{t-1}(0,(m+l)-2^{r+1}) + f_{t-1}((m+l)-2^{r+1},l) \\
\leq \;& f_t(0,2^r) + f_{t-1}(0,m-2^r) + f_{t-1}(m-2^r,2^r) + f_{t-1}(0,(m+l)-2^{r+1}) \\
= \;& f_t(0,2^r) + f_{t-1}(0,2^r) + f_{t-1}(0,(m+l)-2^{r+1}) \\
= \;& f_t(0,2^r) + f_t(2^r,2^{r+1}) + f_t(2^{r+1},m+l) \\
= \;& f_t(0,m+l). \; \square
\end{aligned}$$

Acknowledgements

The authors thank the anonymous referees for helpful comments.

References

[AM02] K. Amano and A. Maruoka. On Learning Monotone Boolean Functions under the Uniform Distribution, *Proc. of ALT02: The 13th International Conference on Algorithmic Learning Theory*: Lecture Notes in Computer Science, vol. 2533, pp. 57–68, Springer-Verlag, 2002; a journal version to appear in *Theoretical Computer Science*.

[BBL98] A. Blum, C. Burch, and L. Langford. On Learning Monotone Boolean Functions, *Proc. of FOCS98: The 39th Annual IEEE Symposium on Foundations of Computer Science*, pp. 408–415, IEEE Press, 1998.

[Bo86] B. Bollobas. *Combinatorics*, Cambridge University Press, 1986.