On Perfectly Balanced Boolean Functions

O. A. Logachev
Information Security Institute,
Lomonosov University, Moscow
e-mail: logol@iisi.msu.ru

Abstract

Perfectly balanced functions were introduced by Sumarokov in [1]. A well known class of such functions are those linear either in the first or in the last variable. We present a novel technique to construct perfectly balanced functions not in the above class.

Keywords: Boolean function, perfectly balanced function, function with defect zero.

1 Introduction

Let $\mathbb{N}$ be the set of natural numbers. For $n \in \mathbb{N}$ let $V_n = \mathbb{F}_2$ be the n-dimensional vector space over the field $\mathbb{F}_2 = GF(2)$. We use $\oplus$ for the addition modulo 2. A Boolean function over V_n is a mapping $V_n \to \mathbb{F}_2$. For any $n \in \mathbb{N}$ we denote by $\mathcal{F}_n$ the set of all Boolean functions in variables $\{x_1, \ldots, x_n\}$. We also identify $\mathcal{F}_n$ with $\mathbb{F}_2[x_1, \ldots, x_n]/(x_i^2 \oplus x_i, i = 1, \ldots, n)$, the quotient ring of the ring of polynomials with coefficients in $\mathbb{F}_2$ w.r.t. the ideal generated by the polynomials $x_i^2 \oplus x_i$, $i \in \{1, \ldots, n\}$. Then for any $f \in \mathcal{F}_n$ we have the algebraic normal form

$$f(x) = \bigoplus_{a_1, \ldots, a_n \in \mathbb{F}_2} g(a_1, \ldots, a_n) x_1^{a_1} \ldots x_n^{a_n} = \bigoplus_{a \in V_n} g(a) x^a, \tag{1.1}$$

where $g \in \mathcal{F}_n$ and $f \to g$ is called Möbius Transform of $\mathcal{F}_n$. By $\deg(f)$ we denote the algebraic degree of a function $f \in \mathcal{F}_n$.

Let $f \in \mathcal{F}_n$ and $i \in \{1, \dots, n\}$. We use the following notation

$$\deg(f, x_i) = \deg(f(x \oplus e_i) \oplus f(x)) + 1,$$

where e_i, $i = 1, \dots, n$, are the vectors of the canonical basis of V_n. If $\deg(f, x_i) = 1$, then we say that f depends linearly on x_i. The weight $\mathrm{wt}(f)$ of f is the number of $x \in V$ such that $f(x) = 1$. A function f is balanced if $\mathrm{wt}(f) = \mathrm{wt}(f \oplus 1) = 2^{n-1}$.

Let $A = \bigcup_{s=1}^{\infty} \mathbb{F}_2^s$. By definition, put $B = \bigcup_{t=n}^{\infty} \mathbb{F}_2^t$. A Boolean function $f \in \mathcal{F}_n$ induces a mapping $B \to A$ of the form

$$b = (b_1, \dots, b_l) \to (f(b_1, \dots, b_n), \dots, f(b_{l-n+1}, \dots, b_l)) \tag{1.2}$$

for any $b \in B$.

Perfectly balanced functions (i.e., Boolean functions f such that the mapping (1.2) is onto) were introduced by Sumarokov in [1]. A well known class of such functions (cf. [2]) consists of all functions that are linear either in the first or in the last variable. The aim of this paper is to develop a novel technique to construct perfectly balanced functions not in the above class.

2 Basic definitions

Let $f \in \mathcal{F}_n$ and $m \in \mathbb{N}$. Consider the system of equations

$$f(x_s, x_{s+1}, \dots, x_{s+n-1}) = y_s, \ \ s = 1, 2, \dots, m, \tag{2.1}$$

where $x = (x_1, \dots, x_{m+n-1}) \in V_{m+n-1}$, $y = (y_1, \dots, y_m) \in V_m$. In vectorial form this system can be written as follows

$$y = f_m^*(x),$$

where

$$\begin{aligned} &f_m^*(x_1, x_2, \dots, x_{m+n-1}) \\ &\quad = (f(x_1, \dots, x_n), f(x_2, \dots, x_{n+1}), \dots, f(x_m, \dots, x_{m+n-1})). \end{aligned} \tag{2.2}$$

For any $f \in \mathcal{F}_n$ and any $m \in \mathbb{N}$ consider a set

$$J(f, m) = \{y \in V_m \mid \forall x \in V_{m+n-1} \ f(x) \neq y\}. \tag{2.3}$$

Denote by $\mathrm{Def}_m(f)$ the cardinality of $J(f, m)$.

Definition 2.1 ([1])**.** A function $f \in \mathcal{F}_n$ is said to have defect zero iff $\mathrm{Def}_m(f) = 0$ for any $m \in \mathbb{N}$.

It is easy to see ([1]) that $f \in \mathcal{F}_n$ has defect zero if $\deg(f, x_1) = 1$ or $\deg(f, x_n) = 1$. Let

$$\mathcal{L}_n = \{f \in \mathcal{F}_n \mid \deg(f, x_1) = 1\}$$

and

$$\mathcal{R}_n = \{f \in \mathcal{F}_n \mid \deg(f, x_n) = 1\}.$$

Definition 2.2 ([1])**.** A function $f \in \mathcal{F}_n$ is called perfectly balanced iff

$$\sharp (f_m^*)^{-1}(y) = 2^{n-1}$$

for any $m \in \mathbb{N}$ and for every $y \in V_m$ ($\sharp M$ denotes the cardinality of the set M).

Let $\mathcal{E}_n$ denote the set of all perfectly balanced functions in $\mathcal{F}_n$. From Definition 2.2 it is easy to see that a perfectly balanced function $f \in \mathcal{F}_n$ is balanced, i.e., $\mathrm{wt}(f) = 2^{n-1}$. It follows immediately that $\sharp\mathcal{E}_n/2^{2^n} \to 0$ as $n \to \infty$.

3 Preliminaries

Theorem 3.1 ([1])**.** *A Boolean function has defect zero iff it is perfectly balanced.*

Denote by $\mathcal{D}_n$ the set of Boolean functions in $\mathcal{E}_n$ such that

$$\deg(f, x_1) \geq 1, \;\; \deg(f, x_n) \geq 1.$$

Sumarokov [1] developed a technique to construct functions in $\mathcal{D}_n \setminus (\mathcal{L}_n \cup \mathcal{R}_n)$ was developed.

Example 3.2 ([1])**.** A Boolean function

$$f(x_1, x_2, x_3, x_4) = x_1x_2 \oplus x_2 \oplus x_3 \oplus x_1x_2x_4 \oplus x_2x_4 \oplus 1$$

is a perfectly balanced function in $\mathcal{D}_4 \setminus (\mathcal{L}_4 \cup \mathcal{R}_4)$.

Furthermore Sumarokov [1] proved some upper bounds on m for functions of nonzero defect and defined the following mappings $\gamma_0, \gamma_1, \gamma_2$ from $\mathcal{F}_n$ onto $\mathcal{F}_n$ such that $\gamma_i(\mathcal{E}_n) = \mathcal{E}_n$, $i = 1, 2, 3$:

(1) $\gamma_0 \colon f(x_1, \dots, x_n) \to f(x_1, \dots, x_n) \oplus 1$;

(2) $\gamma_1 \colon f(x_1, \dots, x_n) \to f(x_1 \oplus 1, \dots, x_n \oplus 1)$;

(3) $\gamma_2 \colon f(x_1, \dots, x_n) \to f(x_n, \dots, x_1)$.

For certain applications it is interesting to investigate conditions under which the distribution of the right-hand side of (2.1) is uniform provided that the distribution of the random vector $X_m = (x_1, \dots, x_{m+n-1})$ is uniform.

Theorem 3.3 ([3])**.** *Let $\{X_m = (x_1, \dots, x_{m+n-1})\}_{m=1}^{\infty}$ be a sequence of random variables, where X_m is distributed uniformly over V_{m+n-1}. Then the random variable $Y_m = f_m^*(X_m)$ is distributed uniformly for any $m \in \mathbb{N}$ iff the function f is perfectly balanced.*

4 Main result

For any $k \in \mathbb{N}$ and any $l \in \mathbb{N}$ consider a mapping $\Xi_{k,l} \colon \mathcal{F}_k \times \mathcal{F}_l \to \mathcal{F}_{k+l-1}$ of the form

$$\Xi_{k,l}(f, g) = f[g] = h \in \mathcal{F}_{k+l-1}, f \in \mathcal{F}_k, g \in \mathcal{F}_l,$$

where

$$\begin{aligned} h(x_1, \dots, x_{k+l-1}) &= f[g](x_1, \dots, x_{k+l-1}) \\ &= f(g(x_1, \dots, x_k), g(x_2, \dots, x_{k+1}), \dots, g(x_k, \dots, x_{k+l-1})). \end{aligned}$$

Our main result is the next theorem.

Theorem 4.1. *Let $f \in \mathcal{F}_k$, $g \in \mathcal{F}_l$. A Boolean function $h = f[g] \in \mathcal{F}_{k+l-1}$ is perfectly balanced iff both functions f and g are perfectly balanced.*

Proof. Let f and g be perfectly balanced functions and m be any natural number. Then for any vector $z = (z_1, \dots, z_m) \in V_m$ we have $\sharp(f_m^*)^{-1}(z) = 2^{k-1}$. Furthermore for every vector $y = (y_1, \dots, y_{m+k-1}) \in (f_m^*)^{-1}(z)$ we have $\sharp(g_{m+k-1}^*)^{-1}(y) = 2^{l-1}$. It now follows that

$$\begin{aligned} \sharp(h_m^*)^{-1}(z) &= \sharp(f[g]_m^*)^{-1}(z) \\ &= \sum_{y \in \sharp(f_m^*)^{-1}(z)} \sharp(g_{m+k-1}^*)^{-1}(y) = 2^{k+l-2} = 2^{(k+l-1)-1}, \end{aligned}$$

for any $m \in \mathbb{N}$ and any $z \in V_m$, i.e. a function $h \in \mathcal{F}_{k+l-1}$ is perfectly balanced.

Let the function $h = f[g] \in \mathcal{F}_{k+l-1}$ be perfectly balanced. Assume the contrary, namely, that either f or g is not perfectly balanced. First assume that f is not perfectly balanced. By Theorem 3.1, f is not a function of defect zero. Then there exist a natural number m and a vector $z = (z_1, \ldots, z_m) \in V_m$ such that $z \in J(f, m)$. Therefore we have $z \in J(f[g], m)$, i.e., $f[g]$ is not a function of defect zero. By Theorem 3.1, $f[g]$ is not perfectly balanced. This contradiction proves that f is perfectly balanced.

Now, assume that g is not perfectly balanced. Then there exist a natural number r and a vector $y^* = (y_1^*, \ldots, y_r^*) \in V_r$ such that $\sharp(g_r^*)^{-1}(y^*) = 2^{l-1}+\alpha$, where $0 < \alpha \le 2^{l-1}$. Using the vector y^*, we construct a set $M_{r,t}$, $t = 1, 2, \ldots$ of vectors in $V_{r(t+1)+(l-1)t}$ of the form

$$\begin{aligned} y = (&y_1^*, \ldots, y_r^*; y_{r+1}, \ldots, y_{l+r-1}; y_1^*, \ldots, y_r^*; \ldots; \\ &y_1^*, \ldots, y_r^*; y_{tr+(t-1)(l-1)}, \ldots, y_{tr+t(l-1)}; y_1^*, \ldots, y_r^*), \end{aligned}$$

where components of y not asterisked are arbitrary. It follows easily that $\sharp M_{r,t} = (2^{l-1})^t$. Using the definition of the set $M_{r,t}$, one can shown that for any $y \in M_{r,t}$ with $\left(g^*_{r(t+1)+(l-1)t}\right)^{-1}(y) \neq \emptyset$, the next inclusion holds:

$$\left(g^*_{r(t+1)+(l-1)t}\right)^{-1}(y) \subseteq \underbrace{(g_r^*)^{-1}(y^*) \times \ldots \times (g_r^*)^{-1}(y^*)}_{t+1}.$$

Furthermore it is clear that

$$g^*_{r(t+1)+(l-1)t}\left(\underbrace{(g_r^*)^{-1}(y^*) \times \ldots \times (g_r^*)^{-1}(y^*)}_{t+1}\right) \subseteq M_{r,t}.$$

Let μ_t denote an expected number of vectors in the set $\underbrace{(g_r^*)^{-1}(y^*) \times \ldots \times (g_r^*)^{-1}(y^*)}_{t+1}$ per vector of the set $M_{r,t}$:

$$\mu_t = \frac{2^{l-1}+\alpha)^{t+1}}{(2^{l-1})^t} = 2^{l-1}\left(1+\frac{\alpha}{2^{l-1}}\right)^{t+1}.$$

Since $(1+\alpha/2^{l-1}) > 1$, it follows that there exists a natural number t_0, such that $\mu_{t_0} > 2^{(k+l-1)-1}$.

Consequently there exists a vector $y \in M_{r,t_0}$ with property

$$\sharp \left(g^*_{r(t_0+1)+(l-1)t_0} \right)^{-1} (y) > 2^{(k+l-1)-1}. \tag{4.1}$$

Let $z = f^*_{r(t_0+1)+(l-1)t_0-k+1}(y)$. Using (4.1) we get

$$\sharp \left(f[g]^*_{(t_0+1)(r+l-1)-k+1} \right)^{-1} (z) > 2^{(k+l-1)-1},$$

i.e., a function $f[g]$ is not perfectly balanced. This contradiction proves the theorem. $\square$

Using Theorem 4.1, we can construct perfectly balanced functions in $\mathcal{D}_n \setminus (\mathcal{L}_n \cup \mathcal{R}_n)$.

Example 4.2. Let $f(x_1, x_2, x_3) = x_1 + x_2 x_3 \in \mathcal{L}_3$ and $g(x_1, x_2, x_3) = x_1 x_2 \oplus x_3 \in \mathcal{R}_3$. Then

$$\begin{aligned} h(x_1, x_2, x_3, x_4, x_5) &= f(g(x_1, x_2, x_3), g((x_2, x_3, x_4), g(x_3, x_4, x_5)) \\ &= x_1 x_2 \oplus x_3 \oplus x_2 x_3 x_4 \oplus x_2 x_3 x_5 \oplus x_3 x_4 \oplus x_4 x_5 \in \mathcal{D}_5 \setminus (\mathcal{L}_5 \cup \mathcal{R}_5). \end{aligned}$$

References

[1] S. N. Sumarokov. Defects of Boolean functions and invertability of a certain class of coding circuits. Obozrenie prikladnoi i promyshlennoi matematiki, v. 1, no. 1, 1994, p. 33–55 (in Russian).

[2] F. P. Preparata. Convolutional transformations of binary sequences: Boolean functions ande their resynchronizing properties. IEEE Trans. Electron. Comput., 1966, v. 15, no. 6, p. 898–909.

[3] O. A. Logachev, A. A. Sal'nikov, V. V. Yashchenko. Boolean functions in coding theory and cryptology. MCCME, Moscow, 2004 (in Russian).