ELSEVIER

Discrete Mathematics 239 (2001) 191–198

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Note

Enumeration of binary orthogonal arrays of strength 1

Jian-Zhou Zhang[a,*,1], Zhi-Sheng You[a,2], Zheng-Liang Li[b]

[a] *Institute of Image & Graphics, College of Computer, Sichuan University, Chengdu, Sichuan 610064, People's Republic of China*

[b] *Department of Applied Mathematics, University of Electronic Science and Technology of China, Chengdu, Sichuan 610051, People's Republic of China*

Abstract

A $k2^m \times n$ (0,1) matrix is called a *binary orthogonal array* of strength m if in any m columns of the matrix every one of the possible 2^m ordered (0,1) m-tuples occurs in exactly k rows and no two rows are identical. In this paper, the enumeration of binary orthogonal arrays is studied, and a closed expression for the enumeration of binary orthogonal arrays of strength 1 is given using the inclusion–exclusion principle and the edge-induced subgraph.

MSC: 05A15; 05B15; 94A60

Keywords: Binary orthogonal arrays; Enumeration; Inclusion–exclusion principle; Edge-induced subgraph; Connected component

1. Introduction

A *binary orthogonal array* of strength m is a $k2^m \times n$ (0,1) matrix, such that in any m columns of the matrix every one of the possible 2^m ordered (0,1) m-tuples occurs in exactly k rows and no two rows are identical. Orthogonal arrays were introduced by Rao in [11]. Binary orthogonal arrays are closely related to binary resilient functions introduced independently by Chor et al. [5] in complexity theory and Bennett et al. [1] in quantum cryptography, and correlation-immune Boolean functions by Siegenthaler [13] in cryptography. Therefore, the structure, construction and enumeration of binary orthogonal arrays have been intensively investigated in [2,4,7,9,12]. In this paper, the enumeration of binary orthogonal arrays of strength 1 is considered.

* Corresponding author.

E-mail address: jzhou99@263.net (J.-Z. Zhang).

[1] Supported partially by China Postdoctoral Science Foundation.

[2] Supported by National Natural Science Foundation of China, No.69732010.

PII: S0012-365X(01)00045-0

The enumeration of binary orthogonal arrays is closely related to the enumeration of correlation-immune functions whose truth-tables are binary orthogonal arrays (see [4] for details). Unfortunately, only by constructing and counting some special classes of correlation-immune functions in [9,12,14,16] were a few of lower bounds on the number of correlation-immune functions of strength 1 presented. However, we have not known how to construct all correlation-immune functions, and it is infeasible to show that all of them have been constructed. Therefore, it is too difficult to give the exact number by improving lower bounds. With a few exceptions, [16] gives the exact numbers of $4\times n$ and $6\times n$ binary orthogonal arrays of strength 1, but their derivation are tedious, it is then impossible to derive a closed expression for the enumeration of binary orthogonal arrays from their derivation. Moreover, in [10], the enumeration of balanced 2-colorings of the n-cube, which is equivalent to the enumeration of correlation-immune functions of strength 1, was studied. The counting formulae of balanced 2-colorings of the n-cube based respectively on n-variable form of Pólya theorem and superposition approach were given, but the cycle index polynomials need computing.

In this paper, using the inclusion–exclusion principle and the edge-induced subgraph, a closed expression for the enumeration of binary orthogonal arrays of strength 1 is given. Its derivation is easily understood, and the closed expression simplifies computing. The paper is organized as follows. The derivation of a closed expression for the enumeration of binary orthogonal arrays of strength 1 is given in Section 2. Section 3 discusses the formula obtained in Section 2. In Section 4, an illustrative example is presented. We conclude the paper with several problems inspiring further research on the enumeration of binary orthogonal arrays of strength m in Section 5.

2. A closed expression for the enumeration of binary orthogonal arrays of strength 1

In order to use the inclusion–exclusion principle (see [8] for details), we first give some notations. $\{0,1\}^{2k\times n}$ denotes a set of all $2k\times n$ $(0,1)$-valued matrices. Let

$$A(2k,n)=\{X\in\{0,1\}^{2k\times n};X \text{ is a binary orthogonal array of strength } 1\},$$
$$B(2k,n)=\{X\in\{0,1\}^{2k\times n};\ \text{Every column of } X \text{ has exactly } k \text{ zeroes}\},$$
$$C(2k,n,i,j)=\{X\in B(2k,n);\ \text{The } i\text{th row equals the } j\text{th row in } X\},$$

where $1\leqslant i<j\leqslant 2k$. Obviously, $|A(2k,n)|=0$ if $2k>2^n$, and $|B(2k,n)|=\binom{2k}{k}^n$, here $|\cdot|$ denotes the cardinality of a set. Then

$$A(2k,n)=B(2k,n)-\bigcup_{\{i,j\}\in\mathscr{P}_2}C(2k,n,i,j),$$

where the union is over a set $\mathscr{P}_2$ consisting of all 2-element subsets of $\{1,2,\ldots,2k\}$, $\mathscr{P}_2$ may be viewed as an edge set of the (undirected) complete graph of order $2k$.

Using the inclusion–exclusion principle,

$$|A(2k,n)| = |B(2k,n)| + \sum_{\mathscr{F}\subseteq\mathscr{P}_2,\ \mathscr{F}\neq\emptyset} (-1)^{|\mathscr{F}|} \left| \bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j) \right|, \tag{1}$$

where the sum ranges over all nonempty subsets $\mathscr{F}$ of $\mathscr{P}_2$, $\mathscr{F}$ may be viewed as a nonempty edge subset of the complete graph of order $2k$.

To illustrate how to compute $|\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)|$ in the right-hand side of Eq. (1), the graph–theoretic terminology and notations are needed in this paper. Standard definitions and terminology for graph–theoretic concepts may be found in the book [3]. The notation of the edge-induced subgraph is followed herein. $\langle\mathscr{F}\rangle$ denotes the subgraph induced by the edge subset $\mathscr{F}$ of the complete graph of order $2k$. The vertex set of $\langle\mathscr{F}\rangle$ is $\bigcup_{\{i,j\}\in\mathscr{F}}\{i,j\}$. The edge set of $\langle\mathscr{F}\rangle$ is $\{(i,j);\{i,j\}\in\mathscr{F}\}$. The number of connected components of $\langle\mathscr{F}\rangle$ is denoted by r. The cardinality of the vertex sets of r connected components in $\langle\mathscr{F}\rangle$ are denoted by $q_1,q_2,\ldots,q_r$, respectively. Clearly, $\min_{1\leqslant i\leqslant r} q_i \geqslant 2$ and the cardinality of the vertex set of $\langle\mathscr{F}\rangle$ is

$$\left| \bigcup_{\{i,j\}\in\mathscr{F}} \{i,j\} \right| = q_1 + q_2 + \cdots + q_r = (I, q(\mathscr{F})),$$

where I is an r-dimension vector $(1,1,\ldots,1)$, $q(\mathscr{F}) = (q_1,q_2,\ldots,q_r)$, $(I,q(\mathscr{F}))$ denotes inner product of two vectors I and $q(\mathscr{F})$. The cardinality of $\{1,2,\ldots,2k\} - \bigcup_{\{i,j\}\in\mathscr{F}}\{i,j\}$ is $2k-(I,q(\mathscr{F}))$.

According to the definition of $C(2k,n,i,j)$ and $\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)$, it is easy to know the following **fact**: two rows of a matrix in $\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)$ are identical if their row labels are all in the same component of $\langle\mathscr{F}\rangle$. The fact is also right for the rows in a column.

On the basis of the above fact, for each $\mathscr{F}$, we can count the number of column selections of matrices in $\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)$. For some column, first selecting some components of $\langle\mathscr{F}\rangle$ and assigning one entries located in this column and the rows whose row labels are in the selected components, the cardinality of the set consisting of the vertices in these selected components can be designated by $(\alpha,q(\mathscr{F}))$, where $\alpha\in\{0,1\}^r$, $\{0,1\}^r$ denotes the set of all r-dimension $(0,1)$-valued vectors, $(\alpha,q(\mathscr{F}))$ is also inner product of two vectors α and $q(\mathscr{F})$. If $(\alpha,q(\mathscr{F}))\leqslant k$ (because there are at most and exactly k ones in each column), other $k-(\alpha,q(\mathscr{F}))$ ones in this column are put in some rows whose row labels are in $\{1,2,\ldots,2k\}-\bigcup_{\{i,j\}\in\mathscr{F}}\{i,j\}$ if $2k-(I,q(\mathscr{F}))\geqslant k-(\alpha,q(\mathscr{F}))$. Therefore, for some α, if $(\alpha,q(\mathscr{F}))\leqslant k$ and $2k-(I,q(\mathscr{F}))\geqslant k-(\alpha,q(\mathscr{F}))$, then there are

$$\binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))}$$

column selections.

If we define

$$\begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix} = \begin{cases} 0 & \text{if } k-(\alpha,q(\mathscr{F}))<0, \\ 0 & \text{if } 2k-(I,q(\mathscr{F}))<k-(\alpha,q(\mathscr{F})), \\ 1 & \text{if } 2k-(I,q(\mathscr{F}) \geqslant k-(\alpha,q(\mathscr{F}))=0, \end{cases}$$

then, for each $\mathscr{F}$, there are

$$\sum_{\alpha\in\{0,1\}^r} \begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix}$$

column selections in each column of matrices in $\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)$. Hence

$$\begin{aligned} \left|\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)\right| &= \prod_{l=1}^{n} \sum_{\alpha\in\{0,1\}^r} \begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix} \\ &= \left[\sum_{\alpha\in\{0,1\}^r} \begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix}\right]^n. \end{aligned}$$

With this result, from Eq. (1), it follows that

$$|A(2k,n)| = \begin{pmatrix} 2k \\ k \end{pmatrix}^n + \sum_{\mathscr{F}\subseteq\mathscr{P}_2,\mathscr{F}\neq\emptyset} (-1)^{|\mathscr{F}|} \left[\sum_{\alpha\in\{0,1\}^r} \begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix}\right]^n. \tag{2}$$

Finally, the second part in the right-hand side of Eq. (2) can be further simplified because many terms of them are zeros. In terms of the definition of $B(2k,n)$, there does not exist a matrix in $B(2k,n)$ which has the same $k+1$ rows. Thus, if $\max_{1\leqslant i\leqslant r} q_i \geqslant k+1$, then $\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)=\emptyset$, i.e., $|\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)|=0$. In fact, if $\max_{1\leqslant i\leqslant r} q_i \geqslant k+1$, without loss of generality, suppose $q_1 \geqslant k+1$, when the first coordinate of α is 1, then $k-(\alpha,q(\mathscr{F}))<0$, hence

$$\begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix} = 0.$$

Otherwise, when the first coordinate of α is 0, due to

$$\begin{aligned} 2k-(I,q(\mathscr{F})) &= 2k-\sum_{i=1}^{r} q_i = 2k-q_1-\sum_{i=2}^{r} q_i \leqslant k-1-\sum_{i=2}^{r} q_i \\ &< k-\sum_{i=2}^{r} q_i \leqslant k-(\alpha,q(\mathscr{F})), \end{aligned}$$

$$\begin{pmatrix} 2k-(I,q(\mathscr{F})) \\ k-(\alpha,q(\mathscr{F})) \end{pmatrix} = 0.$$

Therefore, if $\max_{1\leqslant i\leqslant r} q_i \geqslant k+1$,

$$\sum_{\alpha\in\{0,1\}^r} \binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))} = 0,$$

i.e., $|\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)| = 0$.

Hence, we may suppose that $\max_{1\leqslant i\leqslant r} q_i \leqslant k$, i.e., $q_i \leqslant k$ $(i=1,2,\ldots,r)$, moreover $\sum_{i=1}^r q_i \leqslant 2k$, thus

$$|\mathscr{F}| \leqslant \sum_{i=1}^r \binom{q_i}{2} = \frac{\sum_{i=1}^r q_i^2 - \sum_{i=1}^r q_i}{2} \leqslant \frac{k\sum_{i=1}^r q_i - \sum_{i=1}^r q_i}{2}$$

$$= \frac{(k-1)\sum_{i=1}^r q_i}{2} \leqslant k(k-1).$$

This shows that $|\bigcap_{\{i,j\}\in\mathscr{F}} C(2k,n,i,j)|=0$ if $|\mathscr{F}| > k(k-1)$ (because $\max_{1\leqslant i\leqslant r} q_i \geqslant k+1$).

According to the above illustration, Eq. (2) can be further rewritten as

$$|A(2k,n)| = \binom{2k}{k}^n + \sum_{\mathscr{F}\subseteq\mathscr{P}_2, 1\leqslant|\mathscr{F}|\leqslant k(k-1)} (-1)^{|\mathscr{F}|} \left[\sum_{\alpha\in\{0,1\}^r} \binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))}\right]^n. \tag{3}$$

3. Some remarks on Eq. (3)

In this section, we give some remarks on Eq. (3).

First, because

$$\binom{2k}{k} = \binom{2k-1}{k} + \binom{2k-1}{k-1} = \binom{2k-1}{k} + \binom{2k-1}{k} = 2\binom{2k-1}{k}$$

and

$$\binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))} = \binom{2k-(I,q(\mathscr{F}))}{k-(I,q(\mathscr{F}))+(\alpha,q(\mathscr{F}))} = \binom{2k-(I,q(\mathscr{F}))}{k-(I-\alpha,q(\mathscr{F}))},$$

moreover $I-\alpha\in\{0,1\}^r$, Eq. (3) shows that $|A(2k,n)|$ divides by 2^n.

Second, because

$$|A(2k,n+1)| = \binom{2k}{k}^{n+1} + \sum_{\mathscr{F}\subseteq\mathscr{P}_2, 1\leqslant|\mathscr{F}|\leqslant k(k-1)} (-1)^{|\mathscr{F}|} \left[\sum_{\alpha\in\{0,1\}^r} \binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))}\right]^{n+1}$$

$$= \binom{2k}{k}^n \binom{2k}{k} + \sum_{\mathscr{F}\subseteq\mathscr{P}_2, 1\leqslant|\mathscr{F}|\leqslant k(k-1)} \left[(-1)^{|\mathscr{F}|} \cdot \left[\sum_{\alpha\in\{0,1\}^r} \binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))}\right]^n \left[\sum_{\alpha\in\{0,1\}^r} \binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))}\right]\right],$$

this shows that $|A(2k,n+1)|$ can be computed by combining terms in $|A(2k,n)|$. Therefore, Eq. (3) provides a fast algorithm from $|A(2k,n)|$ to $|A(2k,n+1)|$.

Third, we know that $|A(2k,n)| = 0$ if $2k > 2^n$, namely when $2k > 2^n$,

$$\sum_{\mathscr{F}\subseteq\mathscr{P}_2, 1\leqslant|\mathscr{F}|\leqslant k(k-1)} (-1)^{|\mathscr{F}|} \left[\sum_{\alpha\in\{0,1\}^r} \binom{2k-(I,q(\mathscr{F}))}{k-(\alpha,q(\mathscr{F}))}\right]^n = -\binom{2k}{k}^n.$$

This shows that the second part in the right-hand side of Eq. (3) is possibly written as a simpler form. This is under investigation.

Finally, according to the relation between binary orthogonal arrays and correlation-immune Boolean functions, the number of correlation-immune Boolean functions of n variables of strength 1 is

$$\sum_{k=0}^{2^{n-1}} \frac{|A(2k,n)|}{(2k)!}.$$

4. Example

In this section, let us calculate $|A(4,n)|$ by using Eq. (3). Here $k=2$, hence $\mathscr{F}$ satisfies $1 \leqslant |\mathscr{F}| \leqslant 2$ and $\max_{1\leqslant i\leqslant r} q_i \leqslant 2$.

When $|\mathscr{F}| = 1, \langle\mathscr{F}\rangle$ is a complete graph K_2, $r=1$ and $q(\mathscr{F}) = (2)$. $\mathscr{F}$ has six selections in K_4, and

$$\left|\bigcap_{\{i,j\}\in\mathscr{F}} C(4,n,i,j)\right| = \left[\sum_{\alpha\in\{0,1\}^1} \binom{4-(I,q(\mathscr{F}))}{2-(\alpha,q(\mathscr{F}))}\right]^n = \left[\sum_{\alpha\in\{0,1\}^1} \binom{4-2}{2-(\alpha,q(\mathscr{F}))}\right]^n = \left[\binom{2}{2} + \binom{2}{0}\right]^n = 2^n.$$

When $|\mathscr{F}|=2$, $r=2$ (notice $q_1=3$ if $r=1$) and connected components of $\langle\mathscr{F}\rangle$ are all K_2, hence $q(\mathscr{F})=(2,2)$. $\mathscr{F}$ has three selections in K_4, and

$$\left|\bigcap_{\{i,j\}\in\mathscr{F}} C(4,n,i,j)\right| = \left[\sum_{\alpha\in\{0,1\}^2}\binom{4-(I,q(\mathscr{F}))}{2-(\alpha,q(\mathscr{F}))}\right]^n$$

$$= \left[\sum_{\alpha\in\{0,1\}^2}\binom{4-4}{2-(\alpha,q(\mathscr{F}))}\right]^n$$

$$= \left[\binom{0}{2}+\binom{0}{0}+\binom{0}{0}+\binom{0}{-2}\right]^n = 2^n.$$

Therefore,

$$|A(4,n)| = \binom{4}{2}^n - 6\times 2^n + 3\times 2^n = 6^n - 3\times 2^n = 2^n(3^n-3).$$

This result is the same as that in [16], but the calculation is straightforward by Eq. (3).

$|A(6,n)|$ can be also computed by using Eq. (3), the answer is

$$|A(6,n)| = 2^n(10^n - 15\times 4^n + 45\times 2^n - 40).$$

5. Related problems

We have given a closed expression for enumeration of binary orthogonal arrays of strength 1 by using the inclusion–exclusion principle and edge-induced subgraph. However, the enumeration of binary orthogonal arrays of strength $m \geqslant 2$ remains open.

Several lower bounds on the number of binary orthogonal arrays of strength m were given by constructing and counting some special classes of binary orthogonal arrays of strength m in [15]. Moreover, in [6], an asymptotic formula for the number of correlation-immune Boolean functions was presented. Therefore, many problems need solving to count binary orthogonal arrays of strength m.

Acknowledgements

Dr. Zhang gratefully acknowledges National Laboratory of Pattern Recognition, Institute of Automation, The Chinese Academy of Sciences, Beijing, where he was a postdoctoral fellowship from September 1996 to November 1998. The authors would like to thank anonymous referees for helpful comments and suggestions for improvement of the manuscript.

References

[1] C.H. Bennett, G. Brassard, J.M. Robert, Privacy amplification by public discussion, SIAM J. Comput. 17 (1988) 210–229.
[2] J. Bierbrauer, K. Gopalakrishnan, D.R. Stinson, Orthogonal arrays, resilient functions, error correcting codes and linear programming bounds, SIAM J. Discrete Math. 9 (1996) 424–452.
[3] J.A. Bondy, U.S.R. Murty, Graph Theory with Applications, American Elsevier, New York, 1976.
[4] P. Camion, C. Carlet, P. Charpin, N. Sendrier, On correlation-immune functions, in: J. Feigenbaum (Ed.), CRYPTO '91, Lecture Notes in Computer Science, Vol. 576, Springer, Berlin, 1992, pp. 86–100.
[5] B. Chor, O. Goldreich, J. Hastad, J. Freidmann, S. Rudich, R. Smolensky, The bit extraction problem or t-resilient functions, Proceedings of the 26th IEEE Symposium on Foundations of Computer Science, 1985, pp. 396–407.
[6] O.V. Denisov, Asymptotic formula for the number of correlation-immune Boolean functions of order k, Discrete Math. 2 (1991) 25–46 (in Russian).
[7] K. Gopalakrishnan, D.R. Stinson, Three characterizations of non-binary correlation-immune and resilient functions, Designs, Codes Cryptography 5 (1995) 241–251.
[8] C. Ko, W.D. Wei, Combinational Theory, Vol. 1, Chinese Science Press, Beijing, 1984 (in Chinese).
[9] C. Mitchell, Enumerating Boolean functions of cryptographic significance, J. Cryptology 2 (1990) 155–170.
[10] E.M. Palmer, R.C. Read, R.W. Robinson, Balancing the n-cube: a census of colorings, J. Algebraic Combin. 1 (1992) 257–273.
[11] C.R. Rao, Factorial experiments derivable from combinatorial arrangements of arrays, J. Roy. Statist. Soc. 9 (1947) 128–139.
[12] W.J. Shan, Structure and construction of functions with correlation immunity, Acta Math. Appl. Sinica 14 (1991) 331–336 (in Chinese).
[13] T. Siegenthaler, Correlation-immunity of nonlinear combining functions for cryptographic application, IEEE Trans. Inform. Theory 30 (1984) 776–780.
[14] H.J. Tian, Y.X. Yang, J.Y. Wang, Enumerating correlation-immune functions of order one, J. Electron. 19 (1997) 631–636 (in Chinese).
[15] Q.Y. Wen, G.Z. Xiao, The enumeration of correlation-immune Boolean functions of m-order, J. Electron. 19 (1997) 852–854 (in Chinese).
[16] Y.X. Yang, Enumerating Boolean functions with correlation immunity, J. Electron. 15 (1993) 140–146 (in Chinese).