EXPLICIT DEFINITION OF THE BINARY REFLECTED GRAY CODES

Marston Conder

Department of Mathematics, The University of Auckland

Private Bag 92019, Auckland, NEW ZEALAND

e-mail: conder@math.auckland.ac.nz

Abstract

It is shown that for $1 \leq j \leq n$ and $1 \leq k \leq 2^n$, the jth letter of the kth word of the binary reflected Gray code of length n is equal to the parity of the binomial coefficient ${}^{2^n-2^{n-j}-1}C_{[2^n-2^{n-j-1}-k/2]}$ modulo 2. Also it is shown how this observation and the usual iterative definition of the binary reflected Gray codes are revealed in a modified version of Sierpinski's gasket (Pascal's triangle modulo 2).

1 Introduction

A Gray code of length n is a sequence of n-bit strings (which we shall call words) of letters from some alphabet, with the property that each word differs from the next in just one position. The most celebrated of these codes are the binary reflected Gray codes G_n, which may be defined inductively as follows: G_1 consists of the words 0 and 1 (in that order), and G_{n+1} is obtainable by first listing G_n with each word prefixed by 0 and then listing G_n in reverse order with each word prefixed by 1. For example, $G_2 = (00, 01, 11, 10)$ and then $G_3 = (000, 001, 011, 010, 110, 111, 101, 100)$.

The Gray code G_n provides a Hamilton cycle in the n-dimensional hypercube $Q_n = \{0,1\}^n$. Also Gray codes may be applied in signal processing, statistical analysis to the calculation of correlation coefficients for a variable

subset, and even have some connection with the "Towers of Hanoi" problem. Further details can be found in several references, such as [3], [4] or [6] or those given at the end of [2].

In this paper a direct definition of the binary reflected Gray codes is given in closed form, in terms of residues of certain binomial coefficients modulo 2. Specifically, we prove the following.

Theorem: *For $1 \le j \le n$ and $1 \le k \le 2^n$, the jth letter of the kth word of the binary Gray code of length n is the parity (modulo 2) of the binomial coefficient ${}^{2^n-2^{n-j}-1}C_{[2^n-2^{n-j-1}-k/2]}$.*

This is somewhat surprising but perhaps not entirely unexpected, however an extensive search has not uncovered such an explicit definition (in closed form) elsewhere in the literature.

The theorem arose from an observation made by the author during a seminar by Bob Doran at the University of Auckland on Frank Gray and the Gray code, namely that the words of the binary reflected Gray codes of small length occur in certain sections of a modified form of *Sierpinski's gasket*, or Pascal's triangle modulo 2 (see [1], or §2.2 of [5]). This observation is explained in the next Section, and following some other preliminaries on binomial coefficients the theorem is proved in Section 3.

2 Binomial coefficients modulo 2

First let ${}^nC_k$ be the standard binomial coefficient, defined as the number of k-element subsets of an n-element set, and equal to the coefficient of x^k in the binomial expansion of $(1+x)^n$, for $0 \le k \le n$. Recall that these coefficients satisfy the additive identity ${}^nC_{k-1} + {}^nC_k = {}^{n+1}C_k$ for $1 \le k \le n$, which is a fundamental property of Pascal's triangle.

The triangle's symmetry comes from the identity ${}^nC_k = {}^nC_{n-k}$, and from this it follows for example that ${}^nC_{n/2}$ is always even when n is even. On the other hand, when $n+1$ is a power of 2, every coefficient ${}^nC_k$ is odd; to see this, note that ${}^nC_k$ may be written as a product of rationals of the form $(n+1-j)/j$ for $1 \le j \le k$, and in each case the highest power of 2 dividing the numerator is equal to the highest power of 2 dividing the denominator.

Definition 2.1 : For integers r and s satisfying $0 \le r \le 2s+1$, define $d_{rs} = {}^sC_{[r/2]}$ (where $[r/2]$ is the greatest integer not exceeding $r/2$).

Clearly $d_{0s} = d_{1s} = d_{2s,s} = d_{2s+1,s} = 1$ for all $s \ge 0$, and whenever r is even also $d_{rs} = d_{r+1,s}$, $d_{r-1,s} + d_{r,s} \equiv d_{r,s+1} \bmod 2$, and $d_{rr} \equiv 0 \bmod 2$. The proof is straightforward. A picture of the corresponding triangle of parities of these coefficients modulo 2 is given in Fig.1 for $s \le 15$, with x's for 1's and blanks for 0's. We call this a modified Sierpinski gasket.

Now if we group the rows of this gasket into subsets of sizes 1, 2, 4, 8 and so on (as increasing powers of 2), extend the definition of the coefficients d_{rs} to the case $r > 2s+1$ in some cases, and in each group consider only those rows with index s differing from the index of the row at the top of the next group by a power of 2, then the words of a reflected binary code are revealed as columns in each group. This is illustrated in Fig.2, where we read the words of the code from right to left.

We can make the connection as follows:

Definition 2.2 : For each positive integer n, and for integers j and k satisfying $1 \le j \le n$ and $1 \le k \le 2^n$, define $b_{k,j}^{(n)}$ to be the residue of d_{rs} modulo 2, where $s = 2^n - 2^{n-j} - 1$ and $r = s + 1 + 2^n - k$. Equivalently, define $b_{k,j}^{(n)} \equiv {}^{2^n-2^{n-j}-1}C_{[2^n-2^{n-j-1}-k/2]}$ modulo 2.

To prove our theorem we show $b_{k,j}^{(n)}$ is the jth letter of the kth word of the reflected binary Gray code of length n. Our proof requires the following additional observations:

Lemma 2.3 : If $0 \le s,t < 2^m$ then ${}^{2^m+s}C_t \equiv {}^sC_t$ modulo 2.

Proof. Consider each of ${}^{2^m+s}C_t$ and ${}^sC_t$ as a product of rationals of the form $(2^m+s+1-j)/j$ and $(s+1-j)/j$ respectively for $1 \le j \le t$. In each case the highest power of 2 dividing the numerator $(2^m+s+1-j)$ is equal to the highest power of 2 dividing $(s+1-j)$, since $s < 2^m$. [Note: this argument works even in cases where $t > s$.]

Corollary 2.4 : If $0 \le s,t < 2^m$ then ${}^{2^m+s}C_{2^m+t} \equiv {}^sC_t$ modulo 2.

Proof. If $t > s$ then ${}^{2^m+s}C_{2^m+t} = 0 = {}^sC_t$, while if $s \le t$ then it follows from the above lemma that ${}^{2^m+s}C_{2^m+t} = {}^{2^m+s}C_{s-t} \equiv {}^sC_{s-t} \equiv {}^sC_t \bmod 2$.

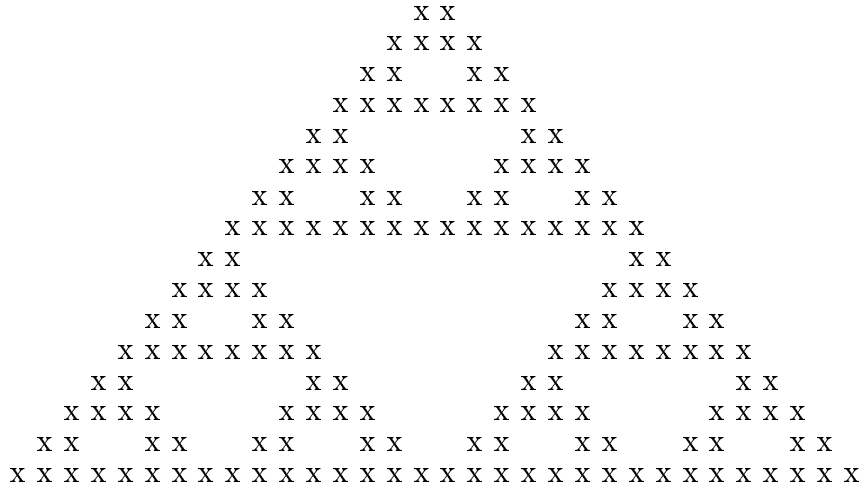

Fig.1: *Modified Sierpinski Gasket*

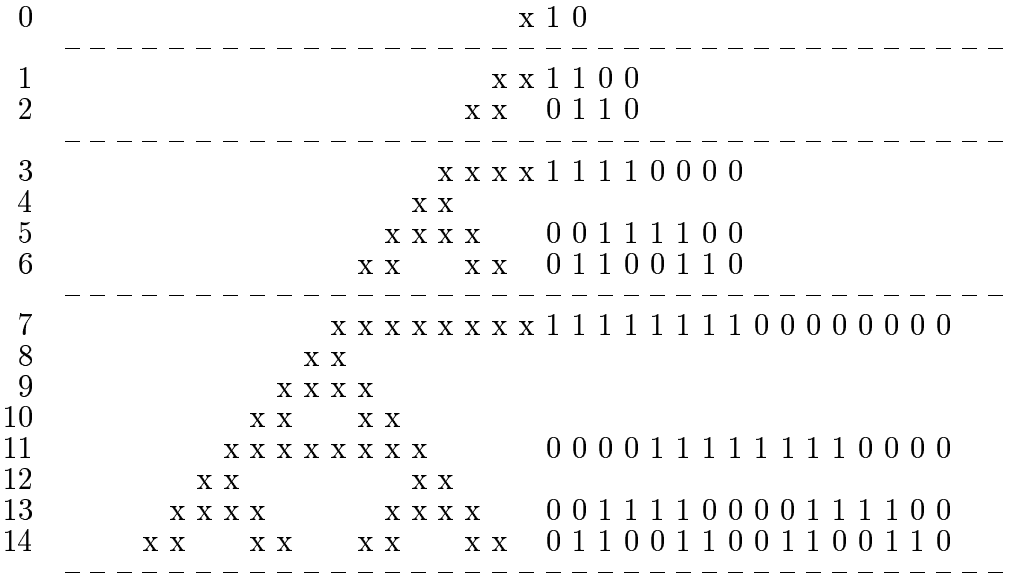

Fig.2: *Occurrence of binary Gray codes in modified Sierpinski gasket*

3 Proof of the Theorem

We now verify that $b_{k,j}^{(n)}$ (as defined in Section 2) is the jth letter of the kth word of the reflected binary Gray code G_n of length n described in the Introduction. Recall that G_1 consists of the words 0 and 1 (in that order), and G_{n+1} is obtainable by first listing G_n with each word prefixed by 0 and then listing G_n in reverse order with each word prefixed by 1.

First it is easy to see from Fig.2 or Definition 2.2 that the theorem is true for $n = 1$: in fact $b_{1,1}^{(1)} \equiv {}^0C_1 \equiv 0$ and $b_{2,1}^{(1)} \equiv {}^0C_0 \equiv 1$ modulo 2.

A similarly easy calculation shows that $b_{k,1}^{(n)} \equiv \begin{cases} 0 & \text{when } 1 \le k \le 2^{n-1} \\ 1 & \text{when } 2^{n-1} < k \le 2^n \end{cases}$ for ${}^{2^n-2^{n-1}-1}C_{[2^n-2^{n-2}-k/2]} = {}^{2^{n-1}-1}C_{[2^{n-1}+2^{n-2}-k/2]}$ is zero when $1 \le k \le 2^{n-1}$, and odd when $2^{n-1} < k \le 2^n$ (since ${}^{2^a-1}C_b$ is odd whenever $0 \le b \le 2^a - 1$).

On the other hand, for $j \ge 2$ and for all $k \le 2^{n-1}$ we note that

$$\begin{aligned}
b_{2^n-k+1,j}^{(n)} &\equiv {}^{2^n-2^{n-j}-1}C_{[2^n-2^{n-j-1}-2^{n-1}+k/2-1/2]} && \text{by definition} \\
&\equiv {}^{2^n-2^{n-j}-1}C_{[2^{n-1}-2^{n-j-1}+k/2-1/2]} \\
&\equiv {}^{2^n-2^{n-j}-1}C_{2^n-2^{n-j}-1-[2^{n-1}-2^{n-j-1}+k/2-1/2]} && \text{as } {}^sC_t = {}^sC_{s-t} \\
&\equiv {}^{2^n-2^{n-j}-1}C_{2^{n-1}-2^{n-j-1}-[k/2+1/2]} \\
&\equiv {}^{2^{n-1}-2^{n-j}-1}C_{2^{n-1}-2^{n-j-1}-[k/2+1/2]} && \text{by Lemma 2.3} \\
&\equiv {}^{2^n-2^{n-j}-1}C_{2^n-2^{n-j-1}-[k/2+1/2]} && \text{by Corollary 2.4} \\
&\equiv {}^{2^n-2^{n-j}-1}C_{[2^n-2^{n-j-1}-k/2]} && \text{as } -[k/2+1/2] = [-k/2] \\
&\equiv b_{k,j}^{(n)} \text{ modulo 2,}
\end{aligned}$$

which is the required "reflective" property of the code.

Finally also for $j \ge 2$ and all $k \le 2^{n-1}$ we have

$$\begin{aligned}
b_{k,j-1}^{(n-1)} &\equiv {}^{2^{n-1}-2^{(n-1)-(j-1)}-1}C_{[2^{n-1}-2^{(n-1)-(j-1)-1}-k/2]} && \text{by definition} \\
&\equiv {}^{2^{n-1}-2^{n-j}-1}C_{[2^{n-1}-2^{n-j-1}-k/2]} \\
&\equiv {}^{2^n-2^{n-j}-1}C_{[2^n-2^{n-j-1}-k/2]} && \text{by Corollary 2.4} \\
&\equiv b_{k,j}^{(n)} \text{ modulo 2,}
\end{aligned}$$

which is now enough to complete the proof by induction on n.

Acknowledgements

The author is grateful to Charlie Colbourn, Bob Doran, Peter Eades and Frank Ruskey for discussions and information on the topic of Gray codes, and to the N.Z. Marsden Fund for research support.

References

[1] Marston Conder & Cameron Walker, Vertex-transitive non-Cayley graphs with arbitrarily large vertex-stabilizer, *J. Algebraic Combinatorics*, to appear 1997.

[2] J.H. Conway, N.J.A. Sloane & A.R. Wilks, Gray codes for reflection groups, *Graphs and Combinatorics* **5** (1989), 315–325.

[3] Alan Doerr & Kenneth Levasseur, *Applied Discrete Structures for Computer Science* (Macmillan, 1989).

[4] F.G. Heath, Origins of the binary code, *Scientific American* **227** (1972), 76–83.

[5] H.-O. Peitgen, H. Jürgens & D. Saupe, *Chaos and Fractals: New Frontiers of Science* (Springer-Verlag, 1992).

[6] E.M. Reingold, J. Nievergelt & N. Deo, *Combinatorial Algorithms: Theory and Practice* (Prentice Hall, 1977).