Balanced Gray Codes

Girish S. Bhat
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206
gsbhat1@eos.ncsu.edu

Carla D. Savage *
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206
cds@cayley.csc.ncsu.edu

Submitted: June 3, 1996; Accepted: August 28, 1996.

Abstract

It is shown that balanced n-bit Gray codes can be constructed for all positive integers n. A *balanced* Gray code is one in which the bit changes are distributed as equally as possible among the bit positions. The strategy used is to prove the existence of a certain subsequence which will allow successful use of the construction proposed by Robinson and Cohn in 1981. Although Wagner and West proved in 1991 that balanced Gray code schemes exist when n is a power of 2, the question for general n has remained open since 1980 when it first attracted attention.

1 Introduction

An n-bit binary Gray code is an exhaustive listing of n-bit strings in which successive strings differ in exactly one bit position. Alternatively, an n-bit binary Gray code can be viewed as a Hamilton path in the n-cube and a *cyclic* binary Gray code as a Hamilton cycle. One such cyclic Gray code, the Binary Reflected Gray Code (BRGC), was patented by Frank Gray [1] as a solution to a communications problem involving digitization of analogue data. Since

*Supported in part by NSF grant DMS9302505

then, binary Gray codes have been used in a wide variety of other applications including databases, experimental design, and puzzle solving [4, 5, 6, 7, 8].

As discussed, for example, in [7] the BRGC scheme, though sufficient to solve the communications problem, is not adequate for certain other applications because of its lack of "uniformity". The term "uniformity" refers to the manner in which the bits change in the Gray code. Several different measures of uniformity and techniques to construct Gray codes satisfying these measures have been proposed in literature. Two such measures are the distribution of "transition counts" [2, 5, 6, 8] and the "gap" [7] of a code. Gray codes which are uniform with respect to the former measure are referred to as *balanced* Gray codes.

To make this notion precise, associate with an n-bit cyclic Gray code $L_n = w_1, w_2, \ldots w_{2^n}$, the *transition sequence* of bit positions $s = s_1, s_2, \ldots, s_{2^n}$, where for $1 \leq i \leq 2^n - 1$, s_i is the bit position in which w_i and w_{i+1} differ and s_{2^n} is the position in which w_{2^n} and w_1 differ. The *transition count* of bit i, $TC(i)$, in the Gray code L_n, is the number of times i occurs in s.

For example, the BRGC is defined by: $L_1 = 0, 1$; and for $n > 1$, $L_n = L_{n-1} \cdot 0, L_{n-1}^{-1} \cdot 1$, where '$\cdot$' denotes concatenation and L_{n-1}^{-1} lists the elements of L_{n-1} from last to first. So, $L_2 = 00, 10, 11, 01$ and $L_3 = 000, 100, 110, 010, 011, 111, 101, 001$. In L_n, the transition counts are given by: $TC(n) = 2$ and $TC(i) = 2^{n-i}$ for $1 \leq i \leq n - 1$.

A Gray code is called *totally balanced* if for any two bit positions i and j, $TC(i) = TC(j)$. A necessary condition for this to hold for an n-bit Gray code is that n is a power of 2. If n is not a power of 2, following [2], we will call a Gray code *balanced* if for any two bit positions i and j, $|TC(i) - TC(j)| \leq 2$. Thus, the BRCG is totally balanced for $n = 1, 2$, balanced for $n = 1, 2, 3$, but unbalanced for $n \geq 4$.

It has been an open question whether balanced Gray codes exist for all values of n. Several techniques for construction of balanced Gray codes have been proposed. These techniques can broadly be classified into two types. Indirect methods [6, 8] involve the transformation of an existing Gray code to obtain one with the required properties, but they do not guarantee balanced Gray codes. Direct methods [2, 9] involve construction of larger Gray codes from smaller ones. The construction of Wagner and West [9], guarantees a balanced Gray code when the number of bits, n, is a power of two. An ingenious construction proposed in [2] produces a Gray code for all n, but balancing the code requires, for each

n, the existence of a subsequence of the transition sequence of a balanced $(n-2)$-bit Gray code satisfying certain constraints. Robinson and Cohn claim without proof in [2] that such a subsequence can always be found.

In this paper, we re-examine the construction of Robinson and Cohn and prove that for each n, we can satisfy the constraints required to produce a balanced n-bit Gray code. When n is a power of 2, the resulting Gray code will be totally balanced.

We review the construction of Robinson and Cohn in Section 2. In Section 3, we show how to find the subsequence which can be used with the Robinson-Cohn construction to produced balanced Gray codes for all n. Suggestions for further investigation follow in Section 4.

2 The Construction of Robinson and Cohn

We describe the direct technique suggested by Robinson and Cohn for the construction of balanced Gray codes. The technique is an extension of Gilberts *ultracomposite* method [3] for constructing the BRGC for n-cubes by combining Hamilton cycles from two $(n-1)$-cubes.

In Robinson and Cohn's approach, a Hamilton path for an n-cube is constructed by combining Hamilton paths from four copies of the $(n-2)$-cube. A stepwise description of the construction is as follows.

1. Consider the transition sequence

$$s = (s_1, s_2, \ldots, s_{2^{n-2}})$$

   of an arbitrary $(n-2)$-bit Gray code. Select a subsequence $t = (t_1, \ldots, t_l)$ of s, with l even, such that t_1 and t_2 are consecutive elements of s, as are t_{l-1} and t_l.

2. Let the four copies of the $(n-2)$-cube from which the n-cube is composed be labeled 00, 01, 11, 10, according to the last two bits of their vertices. In each of the four subcubes, consider the Gray codes defined, respectively, by the transition sequences $s^{(00)} = s^{(01)} = s^{(11)} = s^{(10)} = s$. Delete transitions in the following fashion.

   - From $s^{(00)}$, delete the elements corresponding to the odd-indexed transitions $t_1, t_3, \ldots, t_{l-1}$ of the subsequence t selected in (1).

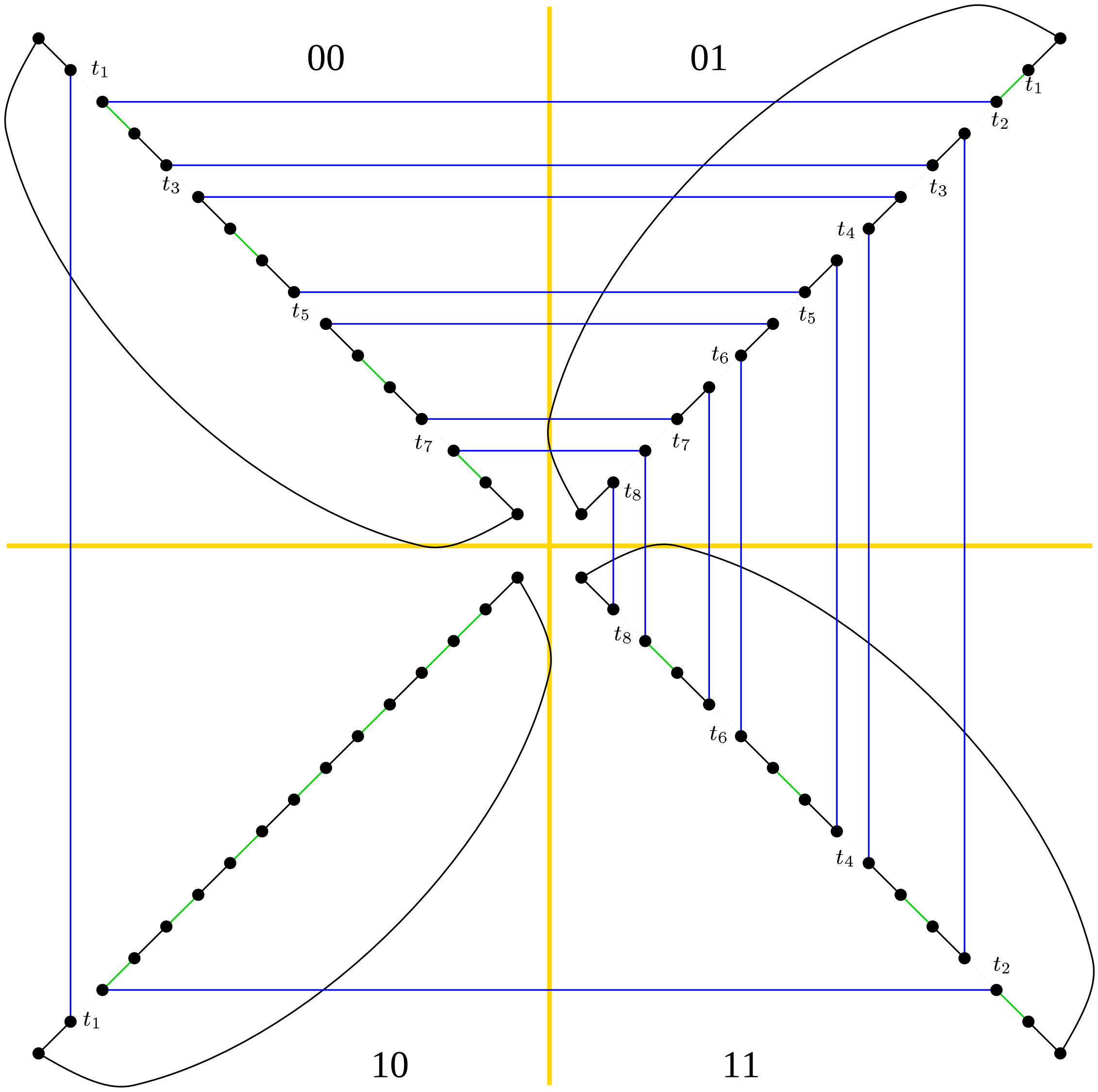

Figure 1: The Robinson-Cohn construction ($l = 8$). Deleted t_i transitions are dotted. In color, added transitions are blue, included t_i transitions are green, excluded t_i are red (and dotted), and original transitions are black.

- From $s^{(01)}$, delete the elements corresponding to $t_2, t_3, \ldots, t_l$.
- From $s^{(11)}$, delete the elements corresponding to the even-indexed transitions $t_2, t_4, \ldots, t_l$.
- From $s^{(10)}$, delete just the element corresponding to transition t_1.

3. Now connect the four subcubes as shown in Figure 2.

It can be checked that the construction described above indeed gives an n-bit Gray code. The distribution of the transition counts in this code depends on the choice of the selected subsequence t. If $TC_n(i)$ denotes the transition count of bit position i in the n-bit Gray code, it is clear that

$$TC_n(n-1) = TC_n(n) = l, \tag{1}$$

where l is the length of the subsequence t. Also, every transition in the subsequence t is deleted once from two different sequences $s^{(ij)}$. For instance, t_1 is deleted twice, once from $s^{(00)}$ and once from $s^{(10)}$. Therefore, if a bit position i occurs b times in t, $2b$ transitions for that particular bit position will be deleted. Consequently, the final transition count for bit position i will be

$$TC_n(i) = 4TC_{n-2}(i) - 2b, \tag{2}$$

where $TC_{n-2}(i)$ is the transition count for bit position i in the $(n-2)$-bit Gray code defined by transition sequence s. So if the subsequence t is chosen strategically, it may be possible that steps 1-3 will result in a balanced n-bit Gray code. The claim of Robinson and Cohn is that if the original $(n-2)$-bit binary code is balanced, then it is always possible to choose such a subsequence.

3 Choosing the Subsequence

In this section, we will show how to use the construction of Robinson and Cohn to produce, for all positive integers n, an n-bit Gray code in which every bit position has transition count either a_n or $a_n + 2$.

For $n \geq 1$, let a_n be the unique even integer satisfying

$$a_n \leq \frac{2^n}{n} < a_n + 2. \tag{3}$$

n	2^n	a_n	k_n	c_n	d_n
1	2	2		1	0
2	4	2		2	0
3	8	2	3	2	1
4	16	4	2	4	0
5	32	6	1	4	1
6	64	10	3	4	2
7	128	18	3	6	1
8	256	32	4	8	0
9	512	56	8	5	4
10	1024	102	13	8	2

Table 1: Values of constants.

Note that if $m = \lfloor 2^n/n \rfloor$ is even, then $a_n = m$, otherwise, $a_n = m - 1$.

We begin by defining certain constants associated with the construction. Let c_n, d_n denote the number of bit positions which would have transition counts a_n, $a_n + 2$, respectively, in the required Gray code. Note that the unique integers satisfying both

$$c_n + d_n = n \tag{4}$$

and

$$c_n a_n + d_n(a_n + 2) = 2^n \tag{5}$$

are

$$c_n = n - d_n; \qquad d_n = \frac{2^n - na_n}{2}. \tag{6}$$

Since $a_n = 2\lfloor 2^{n-1}/n \rfloor$, d_n is just the residue of 2^{n-1} modulo n, so

$$c_n > 0. \tag{7}$$

In the proof of the construction, we make use of one further constant, for $n \geq 3$:

$$k_n = \frac{4a_{n-2} - a_n}{2}. \tag{8}$$

Values of these constants are shown in Table 1 for $n = 1, \ldots, 10$.

Lemma 1 *For $n \geq 7$, $a_{n-2} \geq k_n + 2$.*

4-bit Balanced Gray Code	5-bit Balanced Gray Code	
0000	00000	10111
1000	10000	10101
1100	11000	10001
1101	11100	11001
1111	11110	11101
1110	11111	01101
1010	01111	01100
0010	01110	01000
0110	00110	01010
0100	00010	11010
0101	00011	11011
0111	01011	10011
0011	01001	10010
1011	00001	10110
1001	00101	10100
0001	00111	00100

Figure 2: Balanced Gray codes for $n = 4, 5$.

Proof. Using (8),

$$a_{n-2} - k_n - 2 = a_{n-2} - (4a_{n-2} - a_n)/2 - 2 = a_n/2 - a_{n-2} - 2.$$

From (3), $a_n > 2^n/n - 2$ and $a_{n-2} \le 2^{n-2}/(n-2)$, so

$$a_{n-2} - k_n - 2 > \frac{2^{n-1}}{n} - \frac{2^{n-2}}{n-2} - 3 = \frac{2^{n-2}(n-4)}{n(n-2)} - 3.$$

which is greater than -1 for $n \ge 7$. □

Lemma 2 *If for each $n \ge 6$ we can find integers v_n, y_n satisfying both conditions*

(A) $\quad 0 \le v_n \le c_{n-2}; \quad 0 \le y_n \le d_{n-2}$

and

(B) $\quad (c_{n-2} - v_n)k_n + v_n(k_n - 1) + (d_{n-2} - y_n)(k_n + 3) + y_n(k_n + 4) = l \in \{a_n, a_n + 2\},$

then for all $n \ge 1$, there is an n-bit Gray code in which every bit position changes a_n or $a_n + 2$ times.

Proof. For $n = 1, 2, 3$, the BRGC scheme gives a Gray code satisfying this property. In Figure 2 we exhibit Gray codes for $n = 4, 5$ in which every bit position has transition count a_n or $a_n + 2$. Let $n \geq 6$ and assume inductively that an $(n-2)$-bit Gray code exists with transition sequence $s = s_1, s_2, \ldots, s_{2^{n-2}}$, in which every bit position occurs either a_{n-2} or $a_{n-2} + 2$ times. Further assume the existence of integers v_n, y_n satisfying conditions (A) and (B) of the lemma. By (5), c_{n-2} of the bit positions $1, \ldots, n-2$ have transition count a_{n-2} and d_{n-2} have transition count $a_{n-2} + 2$. Partition the bit positions with transition count a_{n-2} into two sets, U_n, V_n, of sizes $c_{n-2} - v_n$ and v_n, respectively. Partition the bit positions with transition count $a_{n-2} + 2$ into two sets, X_n, Y_n, of sizes $d_{n-2} - y_n$ and y_n, respectively.

We claim that one can construct a subsequence $t = t_1, \ldots, t_l$ of s so that for each bit position $i \in \{1, \ldots, n-2\}$, the number of occurrences of i in t is

$$\begin{array}{ll} k_n & \text{if } \ i \in U_n; \\ k_n - 1 & \text{if } \ i \in V_n; \\ k_n + 3 & \text{if } \ i \in X_n; \\ k_n + 4 & \text{if } \ i \in Y_n; \end{array} \tag{9}$$

and furthermore so that t includes the first two and last two elements of s:

$$t_1 = s_1; \quad t_2 = s_2; \quad t_{l-1} = s_{2^{n-2}-1}; \quad t_l = s_{2^{n-2}}. \tag{10}$$

We can guarantee that (9) is satisfied if, when $i \in U_n \cup V_n$,

$$TC_{n-2}(i) = a_{n-2} \geq k_n$$

and otherwise, if $i \in X_n \cup Y_n$,

$$TC_{n-2}(i) = a_{n-2} + 2 \geq k_n + 4.$$

Note from Table 1 that for $n = 6$, $d_{n-2} = 0$, so that only $a_{n-2} \geq k_n$ need be satisfied, which it is. For $n \geq 7$, we have $a_{n-2} \geq k_n + 2$ by Lemma 1. To see that (10) can be satisfied as well, note that no bit position appears more than twice among $\{s_1, s_2, s_{2^{n-2}-1}, s_{2^{n-2}}\}$. It can be checked from (3), (8), and Table 1 that for $n \geq 6$, we have $k_n \geq 3$. Thus, each bit position is required by (9) to appear at least $k_n - 1 \geq 2$ times in the subsequence t. So, there is no difficulty in arranging for t to satisfy (10) as well as (9), as claimed.

We henceforth assume that t has been chosen to satisfy (9) and (10). By condition (B), t has length $l \in \{a_n, a_n+2\}$. Now, from the sequences s and t, use the construction of Robinson and Cohn to construct an n-bit Gray code. By (1) and (2), in the resulting Gray code, the transition counts of the bit positions are given by

$$TC_n(i) = \begin{cases} 4a_{n-2} - 2k_n = a_n & \text{if } \; i \in U_n \\ 4a_{n-2} - 2(k_n - 1) = a_n + 2 & \text{if } \; i \in V_n \\ 4(a_{n-2}+2) - 2(k_n+3) = a_n + 2 & \text{if } \; i \in X_n \\ 4(a_{n-2}+2) - 2(k_n+4) = a_n & \text{if } \; i \in Y_n \\ l \in \{a_n, a_n+2\} & \text{if } \; i \in \{n-1, n\}, \end{cases} \tag{11}$$

which are all in $\{a_n, a_n+2\}$. It is straightforward to confirm that $\sum_{i=1}^{n} TC_n(i) = 2^n$. □

It remains to show that the hypotheses of Lemma 2 can be satisfied.

Theorem 1 *For all $n \geq 1$, there is an n-bit Gray code in which every bit position changes a_n or a_n+2 times.*

Proof. This is clear for $n = 1, 2, 3$ and, from Figure 2, for $n = 4, 5$. For $n \geq 6$, we show that there exist integers v_n, y_n satisfying both (A) and (B) of Lemma 2.

Condition (B) simplifies via (6) and (8) to

$$a_n + d_n - d_{n-2} - v_n + y_n \in \{a_n, a_n+2\}$$

which is equivalent to

$$d_{n-2} + v_n - y_n \in \{d_n, d_n - 2\}. \tag{12}$$

To satisfy condition (A), we are free to select $v_n \in \{0, 1, \ldots, c_{n-2}\}$ and $y_n \in \{0, 1, \ldots, d_{n-2}\}$. Thus, $d_{n-2} + v_n - y_n$ can assume any integer value in the closed interval $[d_{n-2} - d_{n-2}, d_{n-2} + c_{n-2}] = [0, n-2]$. It remains to show that either d_n or $d_n - 2$ lies in this interval, but this follows immediately from $0 \leq d_n \leq n-1$.

Specifically, v_n and y_n can be chosen as follows:

- If $d_{n-2} \geq d_n$ then set $v_n = 0$ and $y_n = d_{n-2} - d_n$;
- otherwise, if $c_{n-2} + d_{n-2} - d_n \geq 0$ then set $y_n = 0$ and $v_n = d_n - d_{n-2}$;
- otherwise, set $y_n = 0$ and $v_n = d_n - d_{n-2} - 2$.

In each case, it can be checked that (12) is satisfied and that $v_n \in \{0, 1, \ldots, c_{n-2}\}$ and $y_n \in \{0, 1, \ldots, d_{n-2}\}$. We use the fact that $c_n > 0$ from (7), so

$$c_{n-2} + d_{n-2} - d_n = n - 2 - d_n = c_n - 2 \geq -1$$

Thus, if neither of the first two cases hold,

$$d_n - d_{n-2} = c_{n-2} + 1 \geq 2.$$

□

We note that the last case in the proof of Theorem 1 can occur only if $n - 2 - d_n = -1$, that is, only if the residue of 2^{n-1} modulo n is -1. We suspect that this never happens for $n \geq 1$.

Note from (3) and (6) that when n is a power of 2, $d_n = 0$, so the Gray code of our construction is totally balanced. This gives an alternative to the construction of [9] when n is a power of 2.

Corollary 1 *For all $n \geq 1$, if n is a power of two, the construction of Lemma 2 and Theorem 1 gives an n-bit Gray code in which every bit position changes $2^n/n$ times.*

4 Concluding Remarks

Is it possible, for all n, to construct a Gray code in which, for any bit positions i and j, $|TC(i) - TC(j)| \leq 1$? Another problem, suggested by one of the referees, is to determine under what conditions a given partition of $2^n - 1$ into n positive integers can represent the transition counts of a Gray code (or, a partition of 2^n into n even integers for a cyclic Gray code.)

Acknowledgement We are most grateful to the referees for their many suggestions and insights to improve the presentation of this paper. In particular, one of the referees considerably simplified the proof of the main theorem and the other referee contributed Figure 1, a more illuminating alternative to the one in the original version of this paper. The second author would like to thank Luis Goddyn for introducing her to this problem and for generously sharing his ideas.

References

[1] F. Gray, *Pulse Code Communication*, U.S Patent No. 2632058 (March 15,1953).

[2] J. Robinson and M. Cohn, *Counting Sequences*, IEEE Trans. Comput. C-30 (1981) 17-23.

[3] E.N. Gilbert, *Gray Codes and Paths on the n-Cube*, Bell System Technical Journal, Vol. 37 (1958) 815-826.

[4] M. Gardner, *The Curious Properties of the Gray Code and How it Can be Used to Solve Puzzles*, Sci. American, 227 (1972) 106-109.

[5] J.E. Ludman, *Gray Code Generation for MPSK Signals*, IEEE Trans. Commun., COM-29 (1981) 1519-1522.

[6] J.E. Ludman and J.L. Sampson, *A Technique for Generating Gray Codes*, Journal of Statistical Planning and Inference, 5 (1981) 171-180.

[7] L. Goddyn, G.M. Lawrence and E. Nemeth, *Gray Codes with Optimized Run Lengths*, Utilitas Mathematica, 34 (1988) 179-192.

[8] V. E. Vickers and J. Silverman, *A Technique for Generating Specialized Gray Codes*, IEEE Trans. Comput. C-29 (1980) 329-331.

[9] D.G.Wagner and J. West, *Construction of Uniform Gray Codes*, Congressus Numerantium, 80 (1991) 217-223.