ON COUNTING SEQUENCES AND SOME RESEARCH QUESTIONS

I Nengah Suparta

Department of Mathematics Education
Faculty of Mathematics and Natural Sciences
Ganesha University of Education
Email: isuparta@yahoo.com

Abstract
A counting sequence is a list of all binary words of the same length. Counting sequences of any fixed odd Hamming distance between successive words (codewords) are discussed. Gray codes are examples of counting sequences having a single-bit change between neighboring codewords. We describe some results on Gray codes and highlight some research questions. The spectrum of bit changes or transition counts for individual variables for some uniform counting sequences is considered. We show some recent minor findings and pose remaining open questions.

Keywords: Counting sequences, Gray codes, Hamming distance, transition counts, uniform counting sequences.

1. Introduction
A counting sequence is a list of all $2^n$ n-bit distinct binary strings (called codewords). Counting sequences which have specific structures are sometimes desirable to have. For instance, counting sequence having property that the number of bit changes from one codeword to its successor is as large as possible is needed when testing a physical circuit for reliable behavior in worst-case conditions (see e.g. [9, Exercise 67, p. 35]). This kind of counting sequence is called maximum counting sequence. In many applications, having evenly distribution of bits changes for any bit position in a counting sequence is required. This refers to balanced counting sequences. The existence of balanced maximum counting sequences is proven by Suparta and van Zanten in [19].

If in a counting sequence we have that any neighboring codewords differ in a fixed number of bit positions, then we call the sequence uniform. In particular, balanced uniform counting sequences are of considerable interest in combinatorial logic circuits.

Gray codes are very well known example of uniform counting sequences where any successive codewords in the codes differ exactly in one bit position. Although this type of code has been named after its inventor, Frank Gray from Bell Laboratories, and was a patented invention due to Gray in 1953, the code itself actually was demonstrated by the French engineer ‘Emile Baudot’ in 1878 in a telegraph device. Among all kinds of Gray codes, the binary reflected Gray code, also known as the standard Gray code, is the most celebrated examples of Gray codes. This code was used to reduce the coding errors in a pulse code communication system [8]. The usefulness of the binary reflected Gray code and its widespread appearance are undisputed,
for instance in algebraic coding theory (cf. [26]), in the design of combinatorial algorithms (cf. [14]), while its optimality with respect to various applications has proved itself frequently (cf. [2]). For certain applications however, in statistic and computer science, specialized Gray codes with properties not possessed by the standard Gray codes are requested [6]. For instance, when designing experiments, or when designing and testing electrical circuits and information systems, balanced Gray codes are required (cf. [3, 10, 12, 21, 22]).

In this talk we will mainly present some kinds of uniform counting sequences and pose some related open problems.

As the commencement of the discussion we will introduce some basic definitions and notations with respect to counting sequences, and in particular to Gray codes.

2. Basic definitions and notations

In this section we will give definitions in a more formal ways, and introduce some notations. A string \( x = x_n x_{(n-1)} \ldots x_1 \), where \( x_j \in \{0, 1\} \) for all \( j, 1 \leq j \leq n \), is called a binary codeword of length \( n \) or simply \( n \)-codeword.

A counting sequence of length \( n \), also called sequence-\( n \), is an ordered list of all \( 2^n \) distinct \( n \)-codewords. Codewords will be indexed from 0 until \( 2^n - 1 \), and bit positions will be counted from right to left from 1 to \( n \). So, the codeword of index \( i, 0 \leq i \leq 2^n - 1 \), will be written as \( x_i = x_{in} x_{(n-1)} \ldots x_1 \). If in a counting sequence, \( x_0 \) and \( x_{2^n} \) share the property as imposed to any two consecutive codewords, then the sequence is called cyclic. In this case, the codewords \( x_0 \) and \( x_{2^n} \) are identified.

Hamming distance of \( n \)-codewords \( x \) and \( y \), \( d_H(x, y) \), is the number of bit positions where they differ. Whereas the (cyclic) list distance of \( n \)-codewords \( x \) and \( x_j \) in a code \( G \), denoted by \( D(x, x_j) \), is equal to \( D(x, x_j) := \min\{|j - i|, 2^n - |j - i|\} \).

Let \( C := x_0, x_1, \ldots, x_{2^n-1} \) be a counting sequence-\( n \) and let

\[
\begin{align*}
    s_i &:= \{j | x_{i+1} \neq x_j\}, \\
    TC(j) &:= \# \{i | j \in s_i, 1 \leq i \leq 2^n - 1\},
\end{align*}
\]

where \( \#A \) stands for the cardinality of the set \( A \).

In a counting sequence-\( n \), the sequence \( S(n) := s_1, s_2, \ldots, s_{2^n-1} \) is called the transition sequence of the sequence-\( n \), and the distribution

\[
TC = (TC(1), TC(2), \ldots, TC(n))
\]

is called transition count spectrum (or transition count distribution) of the sequence-\( n \). Occasionally authors prefer to consider the distribution \( \left(\frac{TC(1)}{2^n}, \frac{TC(2)}{2^n}, \ldots, \frac{TC(n)}{2^n}\right) \) instead of \( TC \), which is referred to as the bit error probability of the sequence-\( n \).

If \( |TC(i) - TC(j)| \leq 2 \), for every \( i, j \), then the sequence is called balanced, and it is called totally balanced if \( |TC(i) - TC(j)| \leq 0 \) for every \( i, j \).
Proceedings of International Conference on Research, Implementation and Education of Mathematics and Sciences 2014, Yogyakarta State University, 18-20 May 2014

If \( s_j = 1 \) in \( S(n) \), then the sequence \(-n\) is called a Gray code of length \( n \). There are many types of constructed Gray codes based on interests; amongst are balanced Gray codes [3, 10, 15, 17], minimum run-length Gray codes [6, 7], maximum crossover Hamming distance Gray codes [12, 17], Gray codes with certain separability property [4, 13, 24, 25]. Some related research questions on Gray codes which are still worth to see will be discussed in the third section of this talk.

If \( s_j = t \) where \( t \) is an odd positive integer, then the sequence \(-n\) is called uniform sequence \(-n\) and is denoted by \( sequence-(n, t) \). One interesting research question on this matter is on balanced sequence \(-n, t\). Suparta [17] proved that for any realizable pairs of \( n \) and \( t \) where \( n \) and \( t \) are relatively prime, balanced sequence \(-n, t\) exists. The remaining case, when \( \gcd(n, t) > 1 \), is still unsolved problems. This will be discussed in section four of this talk.

3. Some research questions on Gray codes

One know that there are many types of Gray codes constructions (classes). On some classes, separability property or capacity of those Gray codes is of considerable interest. The very well known class of Gray codes is the standard Gray codes which are constructed recursively as follows. The standard Gray code of length 1 is the list of 0, 1. The reflection of this list is 1, 0. Add 0 at front of codewords in the original code, and 1 at front of codewords in the list of its reflection. We get the standard Gray code of length 2 just by concatenating the resulting lists: 00, 01, 11, 10. We proceed similarly, then we get the standard Gray code of length 3: 000, 010, 011, 010, 110, 111, 110, 100. Let \( G(n - 1) \) be the standard Gray code of length \( n - 1 \), \( G(n - 1)R \) be the list of \( G(n - 1) \) in the reversed order, and \( aG(n - 1) \) be the list of the \( n \)-bit code whose codewords are codewords of \( G(n - 1) \) with prefix \( a \in \{0, 1\} \). Then, the standard Gray code of length \( n \) is equal to the list \( 0G(n - 1), 1G(n - 1)R \).
3.1. Gray codes with maximum run length

**Definition 3.1.1.** Let $G$ be a Gray code and $S = s_1, s_2, \ldots, s_{2^n-1}$ be the transition sequence of $G$, then the *minimum run length* of $G$ is

$$mrl(G) = \min \{|i-j|: s_i = s_j, i \neq j\}.$$  

Thus, the distance between the appearance of the same transitions in the transition sequence $S$ of the code $G$ is at least $mrl(G)$.

On this matter one would like to find an $n$-digit Gray code $G$ with maximal $mrl(G)$ over all $n$. The minimum run length of reflected binary Gray codes of length $n$ is equal to 2 for every $n$.

Let $mrl(n)$ be the maximum $mrl(G)$ for every Gray code $G$ of length $n$. The values of $mrl(n)$ for every $n < 8$ are known to be 1, 2, 2, 2, 4, 4, 5 for $n$ equals 1, 2, ..., 7 respectively.

If we write $\lg n$ for $\log_2 n$, then we have the following fact which is mentioned in [6]

**Theorem 3.1.2.** For integers $n \geq 2$ we have $mrl(n) \geq [n - 2.001 \lg n]$.

**Problem 3.1.3.** Does a function $B$ exist such that for every $n \geq 2$, $mrl(n) \geq B(n) > |n - 2.001 \lg n|$?

Goddyn and Gvozdjak [6] presented a hitherto construction on how to produce 10-bit Gray code $C$ of $mrl(C) = 8$. This 10-bit Gray code is the best known Gray code in accordance with minimum run length. Furthermore, They posed in [6] as unsolved the problem of generalizing this 10-bit construction in a more constructive manner.

3.2. Gray codes and their Separability capacity

Separability problem of Gray codes roughly deals with the relationship between the Hamming distance between any two codewords and their list distance. This is expressed by the separability function of the code. Separability problem is formulated as the following. If two codewords in a Gray code differ in $m$ bit positions, how far are they separated from each other in the list of codewords? It is known that the larger this list distance, the smaller the number of bit errors will be when transmitting codewords by means of analog signals (cf. [27]). In more precise statement is when we index the codewords in the list from 0 until $2^n - 1$, and if two codewords $x_i$ and $x_j$ have Hamming distance $d_H(x_i; x_j) = m$, can we find an integral-valued bounding function $b$ such that the list distance satisfies $d_L(x_i; x_j) \geq b(m)$, for $1 \leq m \leq n$?

Yuen [24] and Cavior [4] observed that the separability function of the binary standard Gray codes is $b(n) = \left\lfloor \frac{2^n}{3} \right\rfloor$. Van Zanten and Suparta [25] generalized this study to cyclic $N$-ary
Standard Gray codes and derived the separability function for the cyclic $N$-ary Gray code of length $n$ as $b(n) = \left\lceil \frac{n^n}{N^{n-1}} \right\rceil$. The related question is about finding out Gray codes which have separability capacity better than the standard Gray codes.

**Problem 3.2.1** Does there exist a Gray code and a bound $b(m)$, such that if $d_H(x_i, x_j) = m$ then $D(x_i, x_j) > b(m) \geq \frac{2^m}{3}$ for all $m$-values with $2 < m \leq n$?

A weaker version of the above requirement yields the following problem.

**Problem 3.2.2.** Does there exist a Gray code and a bound $b(m)$, such that if $d_H(x_i, x_j) = m$, then $D(x_i, x_j) \geq b(m) \geq \frac{2^m}{3}$ for all $m$-values with $2 < m \leq n$, whereas at least for one $m$-value $D(x_i, x_j) > \frac{2^m}{3}$, for all pairs $x_i$ and $x_j$ with $d_H(x_i, x_j) = m$?

Using a Gray map, Park and Bose in [13] constructed a new class of Gray codes and proved that this class of codes in some case has separability property

$$d_H(x_i, x_j) = m \Rightarrow D(x_i, x_j) \geq \begin{cases} \frac{4}{15}2^m, & \text{if } m \text{ is odd,} \\ \frac{7}{15}2^m, & \text{if } m \text{ is even.} \end{cases}$$

Observe that the separability capacity of the codes is better than the standard Gray codes if $m \geq 4$ is even, but it is worse otherwise. Suparta [17] introduced a class of Gray codes of length $n$ with the bound function

$$b(m) = \begin{cases} \frac{2^m}{3}, & \text{for } m, 1 \leq m < n, \\ \frac{2^m}{3} + 2^{m-3}, & \text{for } m = n. \end{cases}$$

This is a positive answer of Problem 3.2.2. We can see that the resulting class of Gray codes has the same bound $b(m)$ for all $1 \leq m < n$, but a better bound for $m = n$. Of course, this class of codes seems not to be optimal, and the problem is still open especially for Problem 3.2.1. We notice that for $n = 4$, the resulting Gray code constructed by Suparta in [17] is optimal in terms of separability capacity.

### 3.3. Gray codes with prescribed count spectrum

For some applications, one needs Gray codes with some additional properties. In experimental design, one is interested in Gray codes which possess balanced distribution of bits changes (see e.g. [10, 11, 12, 21]). Apart from applications sides, existence of Gray codes in accordance with transition count spectrum is an interesting combinatorial question. In this section we will formulize the question in a more formal statement after discussing some known facts.

**Theorem 3.3.1** Let $(TC(1), TC(2), \ldots, TC(n))$ be the transition count spectrum of a cyclic Gray code-$n$ which is ordered in non-decreasing way, i.e. $TC(i) \leq TC(i+1)$, for all $i$, $1 \leq i \leq n-1$. 

M-5
Then, one has $\sum_{i=1}^{n} TC(i) \geq 2^k$, for all $k, 1 \leq k \leq n$.

The proof of this theorem can be seen in [17].

An obvious corollary of Theorem 3.3.1 is the following

**Corollary 3.3.2** $\left\lceil \frac{2^i}{i} \right\rceil \leq c_i \leq \frac{2^{2n-2i+1}}{n+1}$ for all $i, 1 \leq i \leq n$.

**Proof.** The lower bound for $c_i$ is trivial. The maximal value of $c_i$ will be reached when $\sum_{j=1}^{i-1} c_j$ reaches its minimum $2^{i-1}$. We conclude that the maximal value of $c_i$ equals $\frac{2^{2n-2i+1}}{n+1}$.

We can extend Theorem 3.3.1 to the non-cyclic case as follows.

**Theorem 3.3.3** Let $(c_1, c_2, ..., c_n)$ be the transition count spectrum of a non-cyclic Gray code $G(n)$ of length $n$ which is ordered in non-decreasing way. Then, one has $\sum_{i=1}^{k} c_i \geq 2^k - 1$, for all $k, 1 \leq k \leq n$.

Let $p$ be a positive integer and $p_1 + p_2 + ... + p_k = p$, where $p_i$ is a positive integer for all $i \in \{1, 2, ..., k\}$. The distribution or spectrum $A = (p_1, p_2, ..., p_k)$ is called a partition of $p$. Furthermore, if $p_i$ is even for all $i$, the spectrum $A$ is called an even partition of $p$. It is easy to observe that there always exists an even partition of $2^n$. The question is about the converse of the theorem which is formulated as the following problem.

**Problem 3.3.4**

Let $A = (p_1, p_2, ..., p_n)$ be an even partition of $2^n$, $p_i \leq p_{i+1}$, for all $i, 1 \leq i \leq n - 1$, and satisfy the condition $\sum_{i=1}^{n} p_i \geq 2^k$, for all $k, 1 \leq k \leq n$. Does there exist a cyclic Gray code of length $n$ having transition count spectrum $A$?

The construction of Ludman and Sampson in [12] seems to be a tool for the problem, but the construction is not satisfactory for large length $n$ of codes.


By using an extended construction of Bakos [1], Suparta [17] proved the following.

**Theorem 3.3.5**

Let $G(n-2)$ be a Gray code with transition count spectrum $(TC(1), TC(2), ..., TC(n-2))$ and $A = (p_1, p_2, ..., p_n)$ be an even partition of $2^n$. A Gray code of length $n$ with transition count spectrum $A$ exists if and only if

(i) $p_k = p_{k+1}$ for some $k \in \{1, 2, ..., n-1\}$;

(ii) $2TC(i) \leq p_i \leq 4TC(i)$ for every $i \in \{1, 2, ..., k-1\}$,
   \[ 2TC(i) \leq p_{i+2} \leq 4TC(i) \text{ for every } i, k \leq i \leq n-2, \]

(iii) there is at least one $i_0 \in \{1, 2, ..., n-2\}$ such that either $p_{i_0} \leq 4(TC(i_0) - 1)$, $i_0 \in \{1, 2,$}
\[ \ldots, k - 1 \) or \( p_{i_0 + 2} \leq 4(TC(i_0) - 1), k \leq i_0 \leq n - 2. \]

Theorem 3.3.5 gives only a minor answer for Problem 3.3.5. Thus, the question of finding the answer of the whole problem is still an open question.

### 3.4 Graphs induced by Gray codes

Let \( S(n) := s_1, s_2, \ldots, s_{2^n - 1} \) be the transition sequence of a Gray code \( G \). The graphs \( \Gamma_G \) induced by the Gray code \( G \) has vertex set \( \{1, 2, \ldots, n\} \) and edge set \( \{\{s_i, s_{i+1}\} : i \in \{1, 2, \ldots, 2^n - 1\}\} \). We emphasize that graph \( \Gamma_G \) is undirected simple graphs. The vertices of the graph \( \Gamma_G \) correspond to bit positions; and vertices \( i \) and \( j \) are adjacent when bit positions \( i \) and \( j \) flip consecutively in the code \( G \). We can see immediately that \( S(n) \) determines graph \( \Gamma_G \) with \( n \) vertices. It is easy to observe that standard Gray code of length \( n \) induces a star or complete graph \( K_{1,n} \). Wilmer and Ernst in [23] proved some graphs induced by Gray codes. Furthermore, Suparta [17, 18] proved that there exists Gray code of length \( n \) inducing complete graph of \( n \) vertices. This result constitutes an answer of Wilmer and Ernst conjecture in [23].

Recently Suparta [20] constructed Gray codes inducing the complete bipartite graph \( K_{m,n} \) for any pair of positive integers \( m \) and \( n \). As an instance, Gray code which has transition sequence \( 1, 2, 1, 3, 4, 3, 1, 3, 4, 2, 4, 3, 1, 3, 4, 3 \) induces the complete bipartite graphs \( K_{2,2} \). This extends the results of Wilmer and Ernst in [23].

A graph \( \Gamma_G \) is said to have bi-directional edge if the transition sequence of the Gray code \( G \) has some subsequence of the form \( i, j, i \). The graph which is induced by Gray code which has transition sequence \( 1, 2, 1, 3, 4, 3, 1, 3, 4, 2, 4, 3, 1, 3, 4, 3 \) is with bi-directional edge since the transition sequence contains subsequence \( 3, 1, 3 \) or \( 4, 2, 4 \). Wilmer and Ernst in [23] proved the existence of Gray code of any length \( n \geq 6 \), the graph of which contains no bi-directional edges.

The following problems are still open to study:

**Problem 3.4.1** Does there exist an \( n \)-bit Gray code, for every \( n \geq 6 \), which induces the complete graph \( K_n \), and which has no bi-directional edges?

**Problem 3.4.1** Does there exist an \( n \)-bit Gray code, for every \( m, n \geq 5 \), which induces the complete graph \( K_{m,n} \), and which has no bi-directional edges?

### 4. On balanced sequence-(\( n, t \))

Gray codes are examples of counting sequences with \( t = 1 \). It is known that balanced Gray codes do exist. Thus, balanced sequence-(\( n, 1 \)) exists for any value of \( n \). There are some constructions (See [9, p.88], [15]) for obtaining sequence-(\( n, t \)) for any realizable values of \( n \) and \( t \). But these constructions do not give guarantee that the resulting sequence-(\( n, t \)) is

---

M-7
balanced. Although, Robinson and Cohn [15] firmly believed that balanced sequence-\((n, t)\)
exists for any realizable pairs of \(n\) and \(t\), and formulated the following conjecture.

**Conjecture 4.1** (Robinson-Cohn [15]). *For every realizable pair of \(n\) and \(t\), \(2 \leq t < n\), a balanced \((n, t)\)-sequence exists.*

Suparta in [17] introduced a slight modification of the construction of Robinson and Cohn in [15] due to anti-transition sequence of a code, and proved the following theorem.

**Theorem 4.2** For every \(m \geq 1\), there exists a balanced uniform \((2m, 2m - 1)\)-sequence, and if \(m\) is a power of two, there exists a totally balanced uniform \((2m, 2m - 1)\)-sequence.

Moreover, Suparta [17] also introduced a construction for constructing sequence-\((n, t)\) based on transition sequence of Gray codes, and proved that balanced sequence-\((n, t)\) exists for any realizable pair of \(n\) and \(t\) satisfying \(\gcd(n, t) = 1\). We formalize the result in the following theorem.

**Theorem 4.3.** For any realizable pair of \(n\) and \(t\), where \(\gcd(n, t) = 1\), there exists a balanced \((n, t)\)-sequence.

Theorem 4.2 and 4.3 give only a minor affirmative solution of Conjecture 4.1. The remaining cases are still open to study.

5. Conclusion

We mentioned some interesting properties of uniform counting sequences which are worth to observe. Some related facts have been derived. But those facts gave only partial or even minor answers to the whole cases. Then we formulated the remaining cases as research problems. For surveys on related material, see [6, 7] for Gray codes with optimal run lengths, [4, 13, 17, 24, 25] for separability of Gray codes, [18, 23] for graphs induced by Gray codes, [9, 12, 17, 21] for Gray codes with prescribed count spectrum, and [15, 17] for balanced uniform counting sequences. An extended survey can be obtained from [16].

6. Bibliography


