compression of files of information retrieval systems. With this application in mind, bounds on the average codelength of an alphabetical code were studied.

The major results of this correspondence are as follows.

1) A necessary and sufficient condition for the existence of a binary alphabetical code was given.
2) An upper bound for L_{opt} (the average codelength of the optimal alphabetical code) was given.

This upper bound shows the redundancy of the optimal alphabetical code in comparison with the Huffman code.

Though this correspondence presents a theoretical bound on L_{opt}, the redundancy of the optimal alphabetical code varies with the distribution of probabilities. To verify the efficiency of the code in practice, the author encoded the descriptors (keywords) of the ERIC thesaurus. There are 8696 descriptors (the average length of a descriptor is about 17 characters), and the alphabet size of the source symbols is 39 (26 capital letters, 10 numeric characters, 2 symbols, and a space character). In this preliminary experiment, L_{Huff} and L_{opt} are 4.254 bits and 4.423 bits, respectively. The redundancy of the optimal alphabetical code is about 5 percent in comparison with the Huffman coding, which shows the usefulness of the alphabetical code.

References

[1] N. Abramson, *Information Theory and Coding*. New York: McGraw-Hill, 1963.
[2] R. G. Gallager, "Variations on a theme by Huffman," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 6, pp. 668–674, 1978.
[3] E. N. Gilbert and E. F. Moore, "Variable length binary encodings," *Bell Syst. Tech. J.*, vol. 38, no. 4, pp. 933–967, 1959.
[4] T. V. Hu and A. C. Tucker, "Optimal computer search trees and variable-length alphabetical codes," *SIAM J. Appl. Math.*, vol. 21, no. 4, pp. 514–532, 1971.
[5] A. Itai, "Optimal alphabetic trees," *SIAM J. Comput.*, vol. 5, no. 1, 9–18, 1976.
[6] D. E. Knuth, *The Art of Computer Programming: Sorting and Searching*, vol. 3, New York: Addison-Wesley, 1973.
[7] D. A. Lelewer and D. S. Hirschberg, "Data compression," *ACM Comput. Surveys*, vol. 19, no. 3, pp. 261–296, 1987.
[8] H. Miyakawa, *Information Theory*. Tokyo: Corona, 1979 (in Japanese).

Index System and Separability of Constant Weight Gray Codes

A. J. van Zanten

Abstract —**A number system is developed for the conversion of natural numbers to the codewords of the Gray code $G(n,k)$ of length n and weight k, and vice versa. As an application sharp lower and upper bounds are derived for the value of $|i-j|$, where i and j are indices of codewords g_i and g_j of $G(n,k)$ such that they differ in precisely $2m$ bits.**

Index Terms —**Gray codes, constant weight codes, index system, ranking problem, number system, separability.**

Manuscript received July 11, 1989; revised January 29, 1991.
The author is with the Department of Mathematics and Informatics, Delft University of Technology, Julianalaan 132, 2628 BL Delft, The Netherlands.
IEEE Log Number 9100444.

I. Introduction

An n-bit Gray code is an ordered sequence of all 2^n n-bit strings (codewords) such that successive codewords differ by the complementation of a single bit. A Gray code is an example of an ordered code. In this correspondence, the term Gray code stands for the so-called binary-reflected Gray code $G(n)$, $n \geq 1$ (cf. e.g., [11]).

Gray codes are used to minimize the number of erroneous bits in bit strings, when transmitted as analog signals (cf. [1]). In fact when bit strings are Gray-coded a one-level error in the analog signal causes an error in one bit. More generally the minimum analog error required to generate m bit errors is equal to $\lceil 2^m/3 \rceil$, as was shown by Yuen in [12]. In [3], Cavior proved that the maximum analog error corresponding to m-bit errors equals $\lfloor 2^n - 2^m/3 \rfloor$. So one has sharp bounds for the *separability* of the code $G(n)$.

Apart from the use made of Gray codes in transmitting information, they also play a role in a number of other mathematical disciplines, such as the theory and construction of minimal-change algorithms to produce various combinatorial objects like permutations, combinations and partitions [2], [4], [11], the analysis of odd–even merging [6], and the theory behind some mathematical puzzles [7].

In many of these applications the question arises of converting a natural number (written in its decimal representation) to its Gray code representation or vice versa of converting a Gray codeword to the integer it represents. If we denote a codeword of $G(n)$ by g_i and let the index i run through the ordered set of integers $0, 1, \cdots, 2^n - 1$, these questions are equivalent to asking for nonrecursive rules that describe the bijective mapping between i and g_i. We refer to this topic as the problem of the *index system* of $G(n)$. Actually the aforementioned minimum and maximum analog errors are sharp bounds for $|i-j|$, where i and j are indices of codewords g_i and g_j such that these words differ in precisely m bits.

In general this problem exists for any ordered code. A solution in the case of $G(n)$ can easily be found (cf. [5], [11, ch. 5]). It appears that for the description of the mapping $i \leftrightarrow g_i$ the binary number system is the appropriate number system for expressing the values of the index i. In [10] Mansour presents a related set of rules, using a weighting system for the bit positions of a codeword.

In this correspondence we are concerned with the subcode $G(n,k)$ of $G(n)$ consisting of those words of $G(n)$ with precisely k 1-bits, $0 < k < n$. We call this code the *constant weight Gray code* of length n and weight k. Like $G(n)$, this code is also of minimal-change type in the sense that each codeword differs in precisely two bits from its successor (cf. [11]), and is also used in algorithms to produce combinatorial objects [2]. In particular we are interested in the index system of $G(n,k)$ considered on its own, i.e., after (re)numbering the codewords by the ordered set of integers $0, 1, \cdots, \binom{n}{k} - 1$ we shall derive rules in Section IV that describe the mapping between i and g_i. It appears that the appropriate number system for expressing the values of i is a number system (cf. Section III), which shows some resemblance to the binomial number system mentioned in [8], [9], and which is used for the index system of the lexicographic code $L(n,k)$ in [5, Ch. 5]. The code $L(n,k)$ consists of the same codewords as $G(n,k)$ but arranged in lexicographic order. Its relationship with the binomial number system is briefly discussed in Section II.

0018-9448/91/0700-1229$01.00

In Section V, we discuss an application of the index system of $G(n,k)$, analogous to the results of Yuen and Cavior. We derive sharp lower and upper bounds for the value of $|i-j|$, where i and j are the indices of codewords g_i and g_j of $G(n,k)$ such that they differ in precisely $2m$ bits.

II. Preliminaries

The n-bit Gray code $G(n)$ is usually denoted as a $2^n xn$-matrix

$$G(n)=\begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{2^n-1}\end{bmatrix}, \tag{1}$$

where

$$g_i = g_{in-1}g_{in-2}\cdots g_{i0} \tag{2}$$

is the ith codeword, $0 \le i \le 2^n-1$, with bits g_{ij}, $0\le j \le n-1$. For the definition of $G(n)$ and for elementary properties we refer to [11, ch. 5]. Among other things it is proved there that, if $(b_n b_{n-1}\cdots b_1 b_0)_2$ is the binary representation of the index i, one has

$$g_{ij} = b_{j+1} + b_j (\text{mod} 2), \qquad 0 \le j < n, \tag{3}$$

or, written more concisely,

$$g_i = i \oplus \left\lfloor \frac{i}{2} \right\rfloor, \tag{4}$$

where $\oplus$ stands for the exclusive-or-operation.

The inverse mapping is given by

$$b_j = \sum_{1=j}^{n-1} g_{il}(\text{mod} 2), \qquad 0\le j<n. \tag{5}$$

In Section IV we shall exploit a property concerning the relative order of two codewords of $G(n)$, which is an immediate consequence of (5). We formulate this property as a lemma.

Lemma: Let g_i and g_j be two codewords of $G(n)$, and let the bit with index k be the first bit from the left in which these codewords differ or, more specifically,

$$\begin{aligned} &g_{il}=g_{jl}, \qquad l=k+1,k+2,\cdots,n-1,\\ &g_{ik}>g_{jk}. \end{aligned}$$

Then $i>j$ if $\Sigma_{l=k+1}^{n-1} g_{il}$ is even and $i<j$ if $\Sigma_{l=k+1}^{n-1} g_{il}$ is odd.

The subcode $G(n,k)$ is defined as the $\binom{n}{k}\times n$-submatrix of $G(n)$ consisting of all codewords with exactly k 1-bits, $0<k\le n$. (For a recursive definition of $G(n,k)$ we refer to [11].) As was already announced in the Introduction, we renumber the rows of $G(n,k)$ by the ordered set of integers $0,1,\cdots,\binom{n}{k}-1$. Two successive codewords of $G(n,k)$, which are indicated by g_i and g_{i+1} with respect to the new index values differ in exactly 2 bits, or stated in terms of the Hamming distance

$$d(g_i,g_{i+1})=2, \qquad 0\le i<\binom{n}{k}-1. \tag{6}$$

Since all codewords have constant weight we have in general for two arbitrary words g_i and g_j that

$$d(g_i,g_j)=2m, \qquad 0\le m\le \min\{k,n-k\}. \tag{7}$$

The integer m is called the Johnson distance between g_i and g_j.

Finally we discuss the index system for $L(n,k)$ that will serve as a guiding principle for deriving the index system of $G(n,k)$. Basic to the index system of $L(n,k)$ is the following property of binomial coefficients (cf. [5, problem 24]).

If k is any integer ≥ 1, then any nonnegative integer n can uniquely be represented as

$$n=\binom{a_k}{k}+\binom{a_{k-1}}{k-1}+\cdots+\binom{a_1}{1}, \tag{8}$$

with

$$a_k>a_{k-1}>\cdots>a_1\ge 0.$$

For a proof we refer to [9]. Implicit in this proof is the construction of the digits $a_k, a_{k-1},\cdots a_1$, respectively. First one chooses a_k as large as possible such that $\binom{a_k}{k}\le n$. Then one chooses a_{k-1} as large as possible such that $\binom{a_{k-1}}{k-1}\le n-\binom{a_k}{k}$, etc. This property provides us with a number system for nonnegative integers, for any fixed value of k, usually called the binomial number system. With respect to this number system (for some fixed value of k), we write

$$n=(a_k a_{k-1}\cdots a_1). \tag{9}$$

Now let l be a codeword of $L(n,k)$ with ones in positions $b_k, b_{k-1},\cdots,b_1$, and with $n-1\ge b_k>b_{k-1}>\cdots>b_1\ge 0$. We introduce the following classes of codewords:

$$\begin{aligned} &\qquad\quad\;\; b_k \qquad b_{k-1} \qquad b_1\\ L_{b_k}&=\{0\cdots 00x\cdots\cdots\cdots\cdots x\},\\ L_{b_{k-1}}&=\{0\cdots 010\cdots 00x\cdots\cdots\cdots x\},\\ &\quad\vdots\\ L_{b_1}&=\{0\cdots 010\cdots 010\cdots 0x\cdots x\}. \end{aligned}$$

For each codeword of class L_{b_i}, $k\ge i\ge 1$, one has to choose precisely i crossmarked places to fill in i ones, whereas the remaining places have to be filled in with zeros. It is obvious that the number of codewords in $L(n,k)$ that precede l is equal to

$$|L_{b_k}|+|L_{b_{k-1}}|+\cdots+|L_{b_1}|. \tag{10}$$

Hence, if the word $0^{n-k}1^k\in L(n,k)$ has index 0 we have for the lexicographic index $\text{ind}_L(l)$ that

$$\text{ind}_L(l)=\binom{b_k}{k}+\binom{b_{k-1}}{k-1}+\cdots+\binom{b_1}{1}=(b_k b_{k-1}\cdots b_1). \tag{11}$$

The inverse problem of converting an index n to the corresponding codeword of $L(n,k)$ amounts to expressing n in the binomial number system by means of the earlier mentioned construction of the digits $a_k, a_{k-1},\cdots,a_1$.

III. The Alternating Binomial Number System

In this section we introduce another binomial number system, based on the following theorem.

Theorem 1: Let k be any integer ≥ 1. Any nonnegative integer n, if k is even, and any positive integer n, if k is odd, can be uniquely represented as

$$n=\binom{a_k}{k}-\binom{a_{k-1}}{k-1}+\binom{a_{k-2}}{k-2}-\cdots\pm\binom{a_1}{1},$$

with

$$a_k>a_{k-1}>a_{k-2}>\cdots>a_1\ge 1.$$

Proof: We distinguish between the cases of k is even and k is odd. Let k be even and $n\ge 0$. First we show the existence of such a representation.

Choose a_k as small as possible such that $\binom{a_k}{k} > n$. Then

$$0 < n_1 := \binom{a_k}{k} - n \le \binom{a_k}{k} - \binom{a_k - 1}{k} = \binom{a_k - 1}{k-1}.$$

Choose a_{k-1} as small as possible such that $\binom{a_{k-1}}{k-1} \ge n_1$. Then $a_{k-1} < a_k$ and

$$0 \le n_2 := \binom{a_{k-1}}{k-1} - n_1 < \binom{a_{k-1}}{k-1} - \binom{a_{k-1} - 1}{k-1} = \binom{a_{k-1} - 1}{k-2}.$$

Choose a_{k-2} as small as possible such that $\binom{a_{k-2}}{k-2} > n_2$. Then $a_{k-2} < a_{k-1}$. Continue with

$$0 < n_3 := \binom{a_{k-2}}{k-2} - n_2 = -n + \binom{a_k}{k} - \binom{a_{k-1}}{k-1} + \binom{a_{k-2}}{k-2}$$

in the same fashion until one has

$$0 < n_{k-1} := \binom{a_2}{2} - n_{k-2} = -n + \binom{a_k}{k} - \binom{a_{k-1}}{k-1} + \cdots + \binom{a_2}{2}.$$

Choose $a_1 = n_{k-1}$. Then we have

$$n = \binom{a_k}{k} - \binom{a_{k-1}}{k-1} + \binom{a_{k-2}}{k-2} - \cdots - \binom{a_1}{1},$$

with

$$a_k > a_{k-1} > a_{k-2} > \cdots > a_1 \ge 1.$$

Let k be odd and $n > 0$.

Choose a_k as small as possible such that $\binom{a_k}{k} \ge n$. Then

$$0 \le n_1 := \binom{a_k}{k} - n < \binom{a_k}{k} - \binom{a_k - 1}{k} = \binom{a_k - 1}{k-1}.$$

Continue with choosing the a_i, $k-1 \ge i \ge 1$, as small as possible such that $\binom{a_i}{i} \ge n_{k-i}$ if i is odd, and $\binom{a_i}{i} > n_{k-i}$ if i is even, as in the k is even case. Since $n > 0$, we finally have $n_{k-1} > 0$ and so we can choose $a_1 = n_{k-1}$ with $a_1 \ge 1$.

We end up with

$$n = \binom{a_k}{k} - \binom{a_{k-1}}{k-1} + \binom{a_{k-2}}{k-2} - \cdots + \binom{a_1}{1},$$

and

$$a_k > a_{k-1} > a_{k-2} > \cdots > a_1 \ge 1.$$

Hence, in all cases we have proved the existence of a representation as stated in the theorem.

To prove the uniqueness of this representation we assume that

$$n = \binom{b_k}{k} - \binom{b_{k-1}}{k-1} + \binom{b_{k-2}}{k-2} - \cdots \pm \binom{b_1}{1},$$
$$b_k > b_{k-1} > \cdots > b_1 \ge 1$$

is any representation of n satisfying the requirements of the theorem. Then we shall show that b_k is the smallest integer such that $\binom{b_k}{k} \ge n$, if k is odd, and $\binom{b_k}{k} > n$, if k is even.

Assume that this is not the case. From the assumption it follows that

$$\binom{b_k}{k} - \binom{b_{k-1}}{k-1} + \binom{b_{k-2}}{k-2} - \cdots \pm \binom{b_1}{1} \le \binom{b_k - 1}{k}$$

or

$$\binom{b_{k-1}}{k-1} - \binom{b_{k-2}}{k-2} + \cdots \mp \binom{b_1}{1} \ge \binom{b_k}{k} - \binom{b_k - 1}{k} = \binom{b_k - 1}{k-1}.$$

However,

$$\begin{aligned}
&\binom{b_{k-1}}{k-1} - \binom{b_{k-2}}{k-2} + \cdots \mp \binom{b_1}{1} \\
&\quad < \binom{b_{k-1}}{k-1} + \binom{b_{k-2}}{k-2} + \cdots + \binom{b_1}{1} \\
&\quad \le \binom{b_{k-1}}{k-1} + \binom{b_{k-1} - 1}{k-2} + \cdots + \binom{b_{k-1} - k + 2}{1} \\
&\quad = \binom{b_{k-1} + 1}{k-1} - 1.
\end{aligned}$$

If $b_k \ge b_{k-1} + 2$, the last expression is less than $\binom{b_k - 1}{k-1}$ and we have a contradiction. The remaining case is when $b_k = b_{k-1} + 1$. Since now

$$\binom{b_k}{k} - \binom{b_{k-1}}{k-1} = \binom{b_{k-1}}{k},$$

the assumption yields

$$\binom{b_{k-2}}{k-2} - \binom{b_{k-3}}{k-3} + \cdots \pm \binom{b_1}{1} \le 0.$$

For odd k, this is obviously a contradiction because $b_{k-2} > b_{k-3} > \cdots > b_1 \ge 1$. For even k, we have also a contradiction, unless the equality sign holds, in which case $b_j = j$, $1 \le j \le k-2$. However, then we have $n = \binom{b_{k-1}}{k}$ and b_k is the smallest possible integer such that $\binom{b_k}{k} > n$. We conclude that in all cases $b_k = a_k$. Similarly we can show that $b_i = a_i$, $k-1 \ge i \ge 1$.

Hence, the representation derived in the first part of the proof is unique. □

The contents of Theorem 1 allow us to represent the positive integers in a unique way, for any fixed value of k. Moreover, if k is even, we can represent 0 as well. We shall call this type of representation the alternating binomial number system (for the chosen k-value) and we shall write

$$n = (a_k a_{k-1} \cdots a_1)_A. \tag{12}$$

We remark that implicit in the proof of Theorem 1, there is an algorithm to determine the digits $a_k, a_{k-1}, \cdots, a_1$.

IV. The Index System for $G(n,k)$

Let g be a codeword of $G(n,k)$ with ones in positions $b_k, b_{k-1}, \cdots, b_1$, and $n-1 \ge b_k > b_{k-1} > \cdots > b_1 \ge 0$. We introduce the following classes of codewords

$$\begin{aligned}
&\qquad\qquad\;\; b_k \qquad b_{k-1} \qquad b_1 \\
G_{b_k} &= \{0 \cdots 0x \cdots \cdots \cdots \cdots \cdots x\}, \\
G_{b_{k-1}} &= \{0 \cdots 010 \cdots 0x \cdots \cdots \cdots x\}, \\
&\;\;\vdots \\
G_{b_1} &= \{0 \cdots 010 \cdots 010 \cdots 0x \cdots \cdots x\}.
\end{aligned}$$

The argument follows that given for the classes L_{b_i} in Section III. One has to choose i crossmarked places in the codewords of G_{b_i} to fill in i ones, $k \ge i \ge 1$. Since all codewords of $G(n,k)$ are also words of $G(n)$ and since their relative order does not change when we restrict ourselves to the subcode $G(n,k)$, we can apply the lemma of Section II. This proves that the number of codewords of $G(n,k)$ preceding g is equal to

$$|G_{b_k}| - |G_{b_{k-1}}| + \cdots \pm |G_{b_1}| + \epsilon_k. \tag{13}$$

Here $\epsilon_k = 0$ if k is even and $\epsilon_k = -1$ if k is odd, since otherwise the codeword g itself would be counted as a word preceding g. It follows that, if the word $0^{n-k}1^k \in G(n,k)$ has index 0, the Gray index $\mathrm{ind}_G(g)$ satisfies

$$\mathrm{ind}_G(g) = \binom{b_k+1}{k} - \binom{b_{k-1}+1}{k-1} + \cdots \pm \binom{b_1+1}{1} + \epsilon_k$$
$$= (b_k+1 b_{k-1}+1 \cdots b_1+1)_A + \epsilon_k. \tag{14}$$

The inverse problem of converting an index n to the corresponding codeword of $G(n,k)$ amounts to expressing $n - \epsilon_k$ in the alternating binomial system by means of the construction of the digits $a_k, a_{k-1}, \cdots, a_1$ in the proof of Theorem 1. The positions $b_k, b_{k-1}, \cdots, b_1$ of the k nonzero entries in the codeword then follow immediately by taking $b_i = a_i - 1$, $k \geq i \geq 1$.

Example: In the following, all codewords of the code $G(6,4)$ are listed arranged in Gray order:

001111	110011	111001
011011	110110	101011
011110	110101	101110
011101	111100	101101
010111	111010	100111. (15)

According to (14), the index of the word 110101 is equal to $(6531)_A = \binom{6}{4} - \binom{5}{3} + \binom{3}{2} - \binom{1}{1} = 7$. Conversely, suppose one wants to know the codeword with index 11 in $G(6,4)$. First we choose a_4 as small as possible such that $\binom{a_4}{4} > 11$. We find $a_4 = 6$. Next we choose a_3 as small as possible such that $\binom{a_3}{3} \geq \binom{6}{4} - 11$ and find $a_3 = 4$. Since $\binom{4}{3} - \binom{6}{4} + 11 = 0$ it now follows immediately that $a_2 = 2$ and $a_1 = 1$ (remember that always $a_i \geq i$, $k \geq i \geq 1$, as a consequence of the inequalities that have to be satisfied by the a_i). So $11 = (6421)_A$ that corresponds to the codeword 101011.

By a similar argument, we could derive the index of g in $G(n)$. Instead of the binomial coefficients in (14), we would have powers of 2 since the number of nonzero entries is not fixed any more in a class G_{b_i}. Some elementary manipulations with sequences of powers of 2 would then lead to the expression (5).

V. Bounds for Distances in $G(n,k)$

In this section we present tight lower and upper bounds for the value of $|i-j|$, where i and j are the indexes of g_i and g_j that have a Hamming distance of $2m$ (cf. Section II).

Theorem 2: Let g_i and g_j be codewords of $G(n,k)$, $n > k > 0$, such that $d(g_i, g_j) = 2m$, $0 < m \leq \min\{k, n-k\}$.

1) The value of $|i-j|$ is minimal for the pair of codewords

$$\begin{cases} g_j = 0^{n-k-m}1^{k-m}101001100110\cdots, \\ g_i = 0^{n-k-m}1^{k-m}010110011001\cdots. \end{cases}$$

2) The value of $|i-j|$ is maximal for the pair of codewords

$$\begin{cases} g_j = 10^{n-k-m}0011001100\cdots 1^{k-m}, \\ g_i = 00^{n-k-m}1100110011\cdots 1^{k-m}. \end{cases}$$

We only give the outlines of a proof. Let g_i and g_j be codewords as indicated in Theorem 2. If $g_{il} = g_{jl}$, we say that g_i and g_j have the lth bit in common. Our proof now consists of the following steps.

a) The value of $|i-j|$ does not increase if one shifts common bits to the left in g_i and g_j.
b) Let $k = m$ and $n = 2m$. If $j > i$ and if $j - i$ is minimal, then the codewords have the form $g_j = 10\bar{g}_j$ and $g_i = 01\bar{g}_i$, with $\bar{g}_i$ and $\bar{g}_j \in G(2m-2, m-1)$ and $d(\bar{g}_i, \bar{g}_j) = 2m-2$.
c) Let $k = m$ and $n = 2m$. If $j - i$ is maximal, then the codewords have the form $g_j = 10\bar{g}_j$ and $g_i = 01\bar{g}_i$, with $\bar{g}_i$ and $\bar{g}_j \in G(2m-2, m-1)$ and $d(\bar{g}_i, \bar{g}_j) = 2m-2$.
d) Using b) and c) and applying induction to m, we can now prove that Theorem 2 is true for $G(2m, m)$, $m > 0$. Part 1) of Theorem 2 follows by a).
e) If g_i and g_j are of the type $g_i = f0\bar{g}_i$ and $g_j = f1\bar{g}_j$ and if $g_{i'} = 0f\bar{g}_i$ and $g_{j'} = 1f\bar{g}_j$, then $|i'-j'| > |i-j|$.
f) If $j > i$ and if $g_j = 1\bar{g}_j$ and $g_i = 0\bar{g}_i$, then $j - i$ increases if one shifts common 0-bits in $\bar{g}_i$ and in $\bar{g}_j$ to the left and common 1-bits to the right.
g) Part 2) of Theorem 2 now follows by using d), e) and f), and applying induction to m.

The calculations necessary to prove a)–c), e), and f) are straightforward and only elementary properties of binomial coefficients are involved. However, they are lengthy. For this reason they are omitted here and we refer to [13] for the details.

We remark that, instead of $0^{n-k-m}1^{k-m}$ in part 1) of Theorem 2, we could have taken any other common subword of length $n - 2m$ with $k - m$ ones.

Corollary: Let g_i and g_j be codewords of $G(n,k)$, $n > k > 0$ and let $d(g_i, g_j) = 2m$, $0 < m \leq \min\{k, n-k\}$.

1) The minimal value of $|i-j|$ is equal to

$$\sum_{l=1}^{m-1} \binom{2l}{l-1} + 1.$$

2) The maximal value of $|i-j|$ is equal to

$$\binom{n}{k} - \sum_{l=1}^{m-1} \binom{k-m+2l}{l-1} - 1.$$

Proof: Assume, without restriction of the generality, that $j > i$.

a) From Part 1 of Theorem 2 and from Section IV, it follows immediately that, if $j - i$ is minimal, we have

$$\begin{aligned} j - i &= \binom{2m}{m} - \binom{2m-1}{m} - \binom{2m-2}{m-1} + \binom{2m-3}{m-1} \\ &\quad - \binom{2m-4}{m-2} - \binom{2m-6}{m-3} + \cdots - \binom{2}{1} + \binom{1}{1} \\ &= \binom{2m-1}{m-1} - \binom{2m-2}{m-1} + \binom{2m-4}{m-1} + \binom{2m-5}{m-2} \\ &\quad + \binom{2m-6}{m-2} + \cdots + \binom{2}{2} + 1 \\ &= \binom{2m-2}{m-2} + \binom{2m-4}{m-3} + \binom{2m-6}{m-4} + \cdots + \binom{2}{0} + 1. \end{aligned}$$

b) The proof is analogous to the proof of Part a). □

Example: In the case of $G(6,4)$ with $m = 2$, Part 1) Theorem 2 yields the pair of codewords $g_j = 111010$ and $g_i = 110101$ having the minimal value for $|i-j|$. This value equals 2 (cf. (15)) which is also delivered by Part 2) of the Corollary. Furthermore, Part 2 of Theorem 2 provides us with the pair $g_j = 100111$ and $g_i = 011011$ for which $|i-j|$ is maximal. This maximal value equals 12 according to Part 2) of the Corollary. This result is also obvious from (15).

Part 1) of the Corollary is analogous to Yuen's lower bound for $|i-j|$ where i and j are the indexes of codewords g_i and g_j of $G(n)$, such that $d(g_i, g_j) = m$ (cf. [12]). Part 2) of the Corollary is analogous to the upper bound for $|i-j|$ as given by Cavior in [3].

Remark: The binomial coefficient occurring in Part 1) of the Corollary is close to the Catalan number $C_l = \binom{2l}{l} l+1$. In fact we have

$$\binom{2l}{l-1} = lC_l.$$

References

[1] W. R. Bennett and J. R. Davey, *Data Transmission*. New York: McGraw-Hill, 1965.

[2] J. R. Bitner, G. Ehrlich, and E. M. Reingold, "Efficient generation of the binary reflected Gray code and its applications," *Comm. A.C.M.*, vol. 19, pp. 517–521, 1976.

[3] S. R. Cavior, "An upper bound associated with errors in Gray code," *IEEE Trans. Inform. Theory*, vol. IT-21, p. 596, 1975.

[4] G. Ehrlich, "Loopless algorithms for generating permutations, combinations and other combinatorial configurations, *J.A.C.M.*, vol. 3, pp. 500–513, 1973.

[5] J. A. Fill and E. M. Reingold, *Solutions Manual to Combinatorial Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[6] P. Flajolet and L. Ramshaw, "Gray code and odd–even merge," *SIAM J. Comput.*, vol. 9, pp. 142–158, 1980.

[7] M. Gardner, *Mathematical Puzzles and Diversions*. New York: Simon and Schuster, 1959.

[8] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*. Reading, MA: Addison-Wesley, 1989.

[9] D. E. Knuth, *Fundamental Algorithms*. Reading, MA: Addison-Wesley, 1969.

[10] I. I. Mansour, "Gray code weighting system," *IEEE Trans. Inform. Theory*, vol. IT-21, p. 596, 1975.

[11] E. M. Reingold, J. Nievergelt, and N. Deo, *Combinatorial Algorithms: Theory and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[12] C. K. Yuen, "The separability of Gray code," *IEEE Trans. Inform. Theory*, vol. IT-20, p. 668, 1974.

[13] A. J. van Zanten, "Binary Gray codes and index systems," Rep. 89-29, Fac. of Math. and Informatics, Univ. of Technology, Delft, The Netherlands, 1989.

Note on "The Calculation of the Probability of Detection and the Generalized Marcum Q-Function"

David A. Shnidman

In the above paper,[1] computational results for $P_N(X,Y)$ are given in Table I. Professor Carl W. Helstrom[2] provided me with corresponding results using steepest descent integration [1].

Manuscript received December 20, 1990.

The author is with the Advanced Techniques Group, Massachusetts Institute of Technology, Lincoln Laboratory, 244 Wood Street, Lexington, MA 02173-9108.

IEEE Log Number 9143033.

[1]D. A. Shnidman, *IEEE Trans. Inform. Theory*, vol. 35, no. 3, pp. 389–400, Mar. 1989.

[2]Private communication.

These results indicate that, for parameters near 10^9, errors in the tabulated values are in the order of 10^{-7}. This is much too large to be accounted for by accumulated roundoff error. With NX and Y near 10^7 the error is more reasonable, in the order of 10^{-10}, but still larger than expected. The problem is in the large parameter calculations of the two exponents "A" in Figs. 1–4. A is calculated as a difference between $M^*\ln(Y)$ and $(Y+C)$, in one case, and $K\ \ln(Nx)$ and $(Nx+C)$, in the other. It turns out that, for parameters in the order of 10^9, each pair of terms is large and about equal so that A is a small difference of two large numbers. The resulting loss in significant digits noticeably affects the accuracy in the final answers in these cases. This problem can be largely overcome by combining terms differently. We can replace the original terms used to calculate A,

$$A = M\ln(y) - (y+C)$$

where, with $z = M+1$, we have

$$C = (z-1/2)\ln(z) - z + \ln(\sqrt{2\pi}) + J(z)$$

and

$$J(z) \approx \frac{1}{12z+}\,\frac{2}{5z+}\,\frac{53}{42z+}\,\frac{1170}{53z+}\,\frac{53}{z},$$

by the following rearrangement,

$$A = \left(z+\frac{1}{2}\right)\left[\frac{(1-y/z)}{1+\frac{1}{(2z)}} + \ln\left(\frac{y}{z}\right)\right] - \frac{1}{2}\ln(2\pi y) - J(z).$$

This substantially reduces the loss in significant digits for A. Alternatively one could compute A using quadruple precision for even more accurate results. The errors with the adjusted calculations for A are in the order of 10^{-12} for parameters near 10^9 and 10^{-14} for parameters near 10^7. Using quadruple precision for the calculation of A, we obtain yet smaller errors, in the order of 10^{-15} or better even for parameters as large as 10^9. This level of error is the limit of accuracy with the double precision arithmetic used throughout (except for the calculation of A). Since there is little or no noticeable effect on the error when parameters are below 10^7 and virtually all cases of practical interest would have values below this, there is little practical reason why one should implement these changes if the earlier version is already installed.

Some corrections are as follows. Line 3 of Fig. 4 should read $XK \leftarrow e^{-Nx}$. The word "be" on the line below (35) should read "by." Equation (47) should read

$$Y_S \approx \frac{1}{2}\left[\left(N-\frac{1}{2}\right) + \left(\sqrt{-\frac{8}{5}\ln\left[4P_{FA}(1-P_{FA})\right]} + \sqrt{\left(N-\frac{1}{2}\right)}\right)^2\right].$$

References

[1] C. W. Helstrom and J. A. Ritcey, "Evaluating radar detection probabilities by steepest descent integration," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-20, pp. 624–634, Sept. 1984.

0018-9448/91/0700-1233$01.00 ©1991 IEEE