

Maximum Cellular Boolean Functions and Perfect Gray Codes¹

A. J. Goldman and B. K. Bender

(January 28, 1963)

I. A Boolean function of n variables, considered as a subset of the discrete unit n -cube B_n , is called *cellular* if each of its connected components is a face of B_n . Hamming's determination of optimal binary single-error-detecting codes is generalized to a characterization of all proper cellular functions with the greatest possible number of elements. II. An analysis is made of a class of Gray codes (Hamiltonian circuits on B_n) with certain special properties, e.g., admitting for $0 \leq d \leq n$ a partition into 2^{n-d} subpaths each forming a d -dimensional face of B_n .

1. Introduction

Let B_n be the set of vertices of the unit cube in n -dimensional space; equivalently, B_n is the set of all binary sequences² of length n . A *Boolean function* of n variables is simply a subset³ of B_n .

Call two vertices of B_n *neighbors* if they differ in precisely one position. A Boolean function f is called *connected* if, given any two vertices X, X' in f , there exists a sequence

$$X = X_1, X_2, X_3, \dots, X_{k-1}, X_k = X'$$

of members of f such that X_i and X_{i+1} are neighbors for $1 \leq i \leq k-1$. It is easy to show that if f is not connected, then it has a unique partition into maximal connected subsets called its *components*;⁴ if f is connected, the unique component consists of f itself.

The relevance of these concepts to the simplification of Boolean functions is as follows. By a *cell* of a Boolean function f , we mean a face⁵ of B_n which lies wholly in f . A *cover* of f is a collection C of cells of f such that each vertex of f lies in at least one cell from C . Each cell of f is assigned a *cost* in a manner which need not be described here, and the cost of a cover C is defined to be the sum of the costs of the cells of C . The problem of finding a minimum-cost cover of f is then identical with the "Problem of Quine"⁶ of finding a simplest "normal disjunctive form" for f , a problem of practical interest (for example) in the design of switching circuits. From the fact that each cell is connected

and therefore lies in a single component of f , it follows that any cover of f is a disjoint union of covers of the components of f and conversely; in particular any minimum-cost cover of f is a disjoint union of minimum-cost covers of the components of f , and conversely. That is, the Boolean simplification problem for a function f which is not connected, can be split into independent "smaller" subproblems by splitting f into its components. Such a decomposition is clearly likely to be helpful in attacking the simplification problem.⁷

In this context it is natural to ask how many components a Boolean function of n variables can possibly have. This question will be answered in section 2, in which we also investigate the structure of the *maximum* disconnected functions of n variables, i.e., those with the greatest possible number of components.

The discussion above also makes it natural to focus attention on classes of Boolean functions f such that the simplification problem for each *component* of f is especially easy. One such class is that of *isolated* functions, those with only single vertices of B_n as components. An isolated Boolean function can be identified with a single-error-detecting binary code, and our characterization in section 2 of *maximum* isolated functions (those with as many vertices as possible), which verifies a conjecture of Chittenden,⁸ is essentially the same as an "optimal code" characterization due to Hamming.⁹

For all methods of "cost" assignment suggested for the Problem of Quine, the faces F of B_n have the property that $\{F\}$ is the unique minimum-cost cover of F . Thus the class of Boolean functions f , such that each component of f is a cell of f , is one of the

¹ Supported in part by the National Science Foundation, Grant No. G-7579.

² A sequence is *binary* if each of its terms is either 0 or 1.

³ The equivalence of this definition with alternative ones is well-known in the field and will not be discussed here.

⁴ Define a binary relation on f ; XX' holds if X and X' are in a connected subset of f . Then R is an equivalence relation and the components of f are the corresponding equivalence classes.

⁵ A d -dimensional *face* of the "discrete cube" B_n is obtained by fixing the (binary) values of $n-d$ of the n coordinates.

⁶ W. V. Quine, The Problem of Simplifying Truth Functions, *Amer. Math. Monthly* **59** (1952), pp. 521-531. There is no reason here to cite the extensive literature on this topic.

⁷ Of course the labor required to determine the components should be taken into account.

⁸ E. W. Chittenden, On the minimal representation of Boolean functions, *Transactions of the Third Conference of Arsenal Mathematicians*, U.S. Army Office of Ordnance Research Rept. No. 58-2, p. 134.

⁹ R. W. Hamming, Error detecting and error correcting codes, *Bell System Tech. J.* **29** (1950), pp. 147-160. Our proof by induction is quite different from Hamming's, and this approach may be of independent interest.

classes mentioned in the last paragraph. Such functions will be called *cellular*, and in section 3 we determine the structure of maximum cellular functions (those other than B_n , with as many vertices as possible). This generalization of Hamming's theorem from isolated functions to cellular functions, is the main new result of the paper.

A *Gray code* on B_n , in the language of graph theory, is a Hamiltonian circuit of the graph formed from B_n by the "neighbor" relation.^{9a} In section 4, the inductive approach introduced earlier (in sec. 2) is employed to establish the existence of Gray codes with certain special properties, e.g., admitting a partition into 2^{n-d} subpaths each forming a d -dimensional face of B_n . These "perfect" Gray codes turn out to coincide with what were called the "conventional" Gray codes by Gilbert,^{9b} who described them in different notation using essentially the same recursive construction. The treatment in section 4, besides giving additional detail on the structure of these Hamiltonian circuits, highlights their geometric aspects in a way which may be better suited to some readers' mathematical intuition than is the strictly arithmetic treatment associated in this context with the phrase "reflected binary."^{9c} It may be of interest that this material was developed from purely geometric considerations, in ignorance of "reflected binary."

2. Maximum Disconnected and Isolated Functions

We begin with a simple lemma:

LEMMA 1. *There is at least one maximum disconnected function of n variables which is isolated.*

For the proof, observe first that because the number of Boolean functions of n variables is finite, a maximum disconnected function f must exist. Now form a function f' by deleting all but one member from every component of f . Clearly f' is isolated, and it must be maximum disconnected since it has the same number of components as does f .

In what follows, it will be convenient to associate to each Boolean function f of $k+1$ variables, two Boolean functions f_0 and f_1 of k variables each, called the *projections*^{9d} of f . f_0 is obtained by dropping the final 0 from each member of f which ends in 0, while f_1 is obtained by dropping the final 1 from each member of f which ends in 1. Note that

$$|f| = |f_0| + |f_1|, \quad (1)$$

where $|f|$ denotes the number of members of f .

LEMMA 2. *If M_n is the maximum number of components of a Boolean function of n variables, then $M_n \leq 2^{n-1}$.*

^{9a} Some authors use "Gray code" to mean any path or circuit in the linear graph formed by B_n .

^{9b} E. C. Gilbert, Gray codes and paths on the N -cube, Bell System Tech. J. 37 (1958).

^{9c} See for example I. Flores, Reflected number systems, Trans. Inst. Radio Eng. Vol. EC-5 (1956).

^{9d} In the switching-circuit literature f_0 and f_1 would be called the *residues* of f with respect to the "last" of the $k+1$ Boolean variables.

This is clear for $n=1$; assume it true for $n=k$ and consider the case $n=k+1$. Lemma 1 guarantees the existence of a maximum disconnected function f of $k+1$ variables which is isolated; then eq (1) shows that $M_{k+1} = |f_0| + |f_1|$. It is easy to see that f_0 and f_1 are also isolated, so that for $i=0$ and $i=1$, $|f_i|$ is the number of components of f_i . Therefore $|f_i| \leq M_k$, and so we have $M_{k+1} \leq 2M_k$. But $M_k \leq 2^{k-1}$ by the induction hypothesis, so that $M_{k+1} \leq 2^k$ and the induction proof is complete.

We now define two special Boolean functions of n variables; $f(n)$ consists of those members of B_n with an *even* number (possibly zero) of 1's among their entries,^{9e} while $g(n)$ consists of those members of B_n with an *odd* number of 1's among their entries.

THEOREM 1. *The functions $f(n)$ and $g(n)$, which form a partition of B_n , are the only maximum isolated Boolean functions of n variables. $M_n = 2^{n-1}$.*

Part of the proof is trivial; it is clear that $f(n)$ and $g(n)$ are isolated, and that they form a partition of B_n . Since $f(n)$ and $g(n)$ each have 2^{n-1} members and thus the same number of components, we have $M_n \geq 2^{n-1}$ and the equality follows by lemma 2.

It only remains to show that if h is *any* maximum isolated function of n variables, then h is either $f(n)$ or $g(n)$. This is clear for $n=1$; suppose it true for $n=k$ and consider the case $n=k+1$. As in the proof of lemma 2,

$$2^k = M_{k+1} = |h| = |h_0| + |h_1| \leq M_k + M_k = 2^k,$$

where we have used the fact that h_0 and h_1 are isolated because h is. Since equality holds throughout the last display, h_0 and h_1 each have exactly M_k components. The induction hypothesis leads to four possibilities:

$$h_0 = f(k) \text{ and } h_1 = f(k),$$

$$h_0 = g(k) \text{ and } h_1 = g(k),$$

$$h_0 = f(k) \text{ and } h_1 = g(k),$$

$$h_0 = g(k) \text{ and } h_1 = f(k).$$

The first two possibilities are incompatible with the fact that h is isolated, the third yields $h = f(k+1)$, and the fourth yields $h = g(k+1)$; these conclusions follow from the identity¹⁰

$$h = (h_0 \times \{0\}) \cup (h_1 \times \{1\}) \quad (2)$$

valid for Boolean functions of $k+1$ variables. Thus the induction proof is complete.

[Jack Edmonds (NBS Operations Research Section) has observed that B_n can be regarded as a bipartite graph, and has given the following (unpublished) generalization of theorem 1: For a connected bipartite graph which has a connected regular subgraph containing all its vertices, the sets of the

^{9e} That is, a vertex $X = (x_1, x_2, \dots, x_n)$ of B_n is in $f(n)$ if and only if $\sum_{i=1}^n x_i$ is even.

¹⁰ $h_0 \times \{0\}$ is obtained by adjoining a 0 at the end of each member of h_0 . The multiplication sign refers to the Cartesian-product representation of B_{k+1} as $B_k \times \{0, 1\}$.

unique bipartition are the only maximum isolated sets of vertices.]

One might conjecture that a maximum disconnected Boolean function is necessarily isolated. A counter-example for $n=1$ is provided by B_1 itself. The next theorem shows that this is the only counter-example.

THEOREM 2. *For $n>1$, every maximum disconnected function of n variables is isolated.*

To prove this, let f be a maximum disconnected function of $n>1$ variables. As in the proof of lemma 1, form f' by deleting all but one member from each component of f . Since f' is isolated and has the same number of components as f , it follows from theorem 1 that f' is either $f(n)$ or $g(n)$. Without loss of generality assume $f'=f(n)$. If $f \neq f'$, then f includes at least one vertex Y of $g(n)$.

One neighbor X_1 of Y is obtained by changing just the first entry of Y , while a second neighbor X_2 is obtained by changing just the second entry of Y . Since X_1 and X_2 are in $f(n)=f'$ and thus in f , and both are neighbors of the same vertex Y of f , it follows that the elements X_1 and X_2 of f' lie in the same component of f . This however contradicts the construction of f' ; we conclude that $f=f'$, so that f is isolated.

3. Maximum Cellular Functions

Recall that a Boolean function f is called *cellular* if each component of f is a cell of f ; different components may be cells of different dimensions. Cellular functions are a natural generalization of isolated functions.

THEOREM 3. *Let F_d be a d -dimensional face of B_n , $0 \leq d < n$. There exists a Boolean function $f(n; F_d)$ of n variables with F_d as a component, 2^{n-d-1} components in all, and every component a d -dimensional cell. The complement $g(n; F_d)$ of $f(n; F_d)$ in B_n also has 2^{n-d-1} components, each a d -dimensional cell.*

These functions $f(n; F_d)$ and $g(n; F_d)$ can be regarded as the analogs, for cellular functions, of the functions $f(n)$ and $g(n)$ (defined before theorem 1) for isolated functions. For example our later results will imply as a special case that these functions (for various F_d) give the only solutions to the problem of constructing a cellular function with as many components as possible under the restriction that every component is a d -dimensional cell; theorem 1 is the particular instance $d=0$.

We prove the theorem by explicitly constructing $f(n; F_d)$ and verifying its properties. For simplicity assume F_d consists of all members (x_1, x_2, \dots, x_n) of B_n such that $x_i=0$ for $d+1 \leq i \leq n$; any other d -dimensional face can be treated similarly. The cube B_n can be written as a Cartesian product $B_d \times B_{n-d}$, i.e., a binary sequence of length n is uniquely expressible as a sequence of length d followed by one of length $n-d$. With this notation, we use the functions introduced before theorem 1 to define

$$f(n; F_d) = \cup \{B_d \times \{Z\} : Z \in f(n-d)\} = B_d \times f(n-d).$$

Each set $B_d \times \{Z\}$ is easily shown to be a d -dimensional face of B_n and is therefore connected; F_d itself is obtained as $B_d \times \{0_{n-d}\}$ where 0_{n-d} is a sequence of $n-d$ zeros. Since $f(n-d)$ has 2^{n-d-1} members (see theorem 1), there are 2^{n-d-1} sets $B_d \times \{Z\}$. To see that these sets are the *components* of $f(n; F_d)$, it suffices to observe that any two members Z, Z' of the isolated function $f(n-d)$ differ in at least two positions, so that no member of $B_d \times \{Z\}$ can be a neighbor of any member of $B_d \times \{Z'\}$.

The complement of $f(n-d)$ in B_{n-d} is the isolated set $g(n-d)$ which has 2^{n-d-1} members. Therefore the complement of $f(n; F_d)$ in B_n is

$$g(n; F_d) = \cup \{B_d \times \{Z\} : Z \in g(n-d)\} = B_d \times g(n-d)$$

whose components are the 2^{n-d-1} d -dimensional faces $B_d \times \{Z\}$ of B_n with $Z \in g(n-d)$. Thus theorem 3 is proved.

Before proving the next theorem, it is convenient to collect several elementary facts as a lemma.

LEMMA 3. *Let f_0 and f_1 be the projections of a cellular function f of $k+1$ variables. Then*

- (a) *if h is a component of f , then either $h_0=\phi$ or $h_1=\phi$ or $h_0=h_1$,*
- (b) *for $i=0$ and $i=1$, the components of f_i are the nonvoid projections h_i of the components h of f ,*
- (c) *f_0 and f_1 are cellular if nonempty, and*
- (d) *if component p of f_0 meets component q of f_1 , then $p=q=h_0=h_1$ for some component h of f .*

To begin the proof, let h be any component of f . Since f is cellular, for some d there exist $S_h \subset \{1, 2, \dots, k, k+1\}$ with $k+1-d$ members, and a binary sequence $\{c_i : i \in S_h\}$, such that h consists of all members (x_1, \dots, x_{k+1}) of B_{k+1} obeying

$$x_i = c_i \text{ for all } i \in S_h.$$

If $k+1 \in S_h$ and $c_{k+1}=0$, then $h_1=\phi$ and h_0 consists of all members (x_1, \dots, x_k) of B_k such that

$$x_i = c_i \text{ for all } i \in S_h - \{k+1\}.$$

If $k+1 \in S_h$ and $c_{k+1}=1$, then $h_0=\phi$ and h_1 consists of all members (x_1, \dots, x_k) of B_k such that

$$x_i = c_i \text{ for all } i \in S_h - \{k+1\}.$$

If $k+1 \notin S_h$, then h is the disjoint union of the set of members of B_{k+1} such that

$$x_{k+1}=0, x_i=c_i \text{ for all } i \in S_h$$

and the set of members of B_{k+1} such that

$$x_{k+1}=1, x_i=c_i \text{ for all } i \in S_h;$$

thus $h_0=h_1$ consists of all members of B_k such that

$$x_i = c_i \text{ for all } i \in S_h.$$

This proves (a) of the lemma, and also shows that h_0 and h_1 are faces of B_k and therefore cells of f_0 and f_1 respectively.

To prove (b) of the lemma for $i=0$ (the proof for $i=1$ is similar), it now suffices to show that if h_0 and h_0^* are nonempty projections of *distinct* components h and h^* of f , then no member X of h_0 is a neighbor of any member X^* of h_0^* . If such a "neighboring" occurred, however, then the member $X \times \{0\}$ of h would be a neighbor of the member $X^* \times \{0\}$ of h^* , which is impossible. Thus (b) holds; (c) follows from (b) and the fact (proved incidentally in the last paragraph) that the projections of a face are again faces.

Now suppose a component p of f_0 and a component q of f_1 have a member X in common. By (b), there exist components h_p and h_q of f such that $(h_p)_0 = p$ and $(h_q)_1 = q$. Thus $X \times \{0\}$ is in h_p and its neighbor $X \times \{1\}$ is in h_q , so that $h_p = h_q = h$, say. It follows from (a) that $p = q$, so (d) is proved.

THEOREM 4. *The maximum number of members in a cellular Boolean function of n variables (other than B_n) is 2^{n-1} . For $0 \leq d \leq n$ and each d -dimensional face F_d of B_n , $f(n; F_d)$ is the only cellular function with 2^{n-1} members which has F_d as a component.*

This is clear for $n=1$; assume it true for $n=k$ and consider the case $n=k+1$. Let V_{k+1} be the maximum number of members in a cellular function, other than B_{k+1} , of $k+1$ variables; V_k is defined similarly. It follows from theorem 3 that $V_{k+1} \geq 2^k$. To prove the opposite inequality, let f be a cellular function of $k+1$ variables with V_{k+1} members. If either of f_0 or f_1 is B_k then by (d) of lemma 3 the other must be empty, since otherwise we would have $f_0 = f_1 = B_k$ and thus $f = B_{k+1}$. Thus if $f_0 = B_k$ then $f_1 = \phi$ and $f = B_k \times \{0\}$, while if $f_1 = B_k$ then $f_0 = \phi$ and $f = B_k \times \{1\}$; in either case $V_{k+1} = |f| = 2^k$. If both f_0 and f_1 are proper subsets of B_k then by the inductive hypothesis and (c) of lemma 3 we have

$$V_{k+1} = |f| = |f_0| + |f_1| \leq V_k + V_k = 2^{k-1} + 2^{k-1} = 2^k, \quad (3)$$

completing the proof that $V_n = 2^{n-1}$ for all n .

Now let F_d be a d -dimensional face of B_{k+1} , where $0 \leq d \leq k+1$, and let f be a cellular Boolean function of $k+1$ variables with 2^k members and F_d as a component. For simplicity we again use the representation $B_{k+1} = B_d \times B_{k+1-d}$ and assume $F_d = B_d \times \{0_{k+1-d}\}$, where 0_{k+1-d} is a sequence of $k+1-d$ zeros.^{10a}

If $f_0 = B_k$ then (see the last paragraph)

$$f = B_k \times \{0\} = f(k+1; F_k) = f(k+1; F_d)$$

where the last equation holds since F_d is to be a component and $B_k \times \{0\}$ has F_k as sole component. If $f_1 = B_k$ then (see the last paragraph) $f_0 = \phi$, which is impossible since f_0 contains the nonempty set $^{11}(F_d)_0 = B_d \times \{0_{k-d}\}$. So we can assume f_0 and f_1 are proper cellular functions of k variables; by (3), f_0 and f_1 each have 2^{k-1} members, so that by inductive

hypothesis and (b) and (c) of lemma 3,

$$f_0 = f(k; (F_d)_0).$$

In the next paragraph we will show that $f_1 = B_k - f_0$; it follows that

$$f_1 = g(k; (F_d)_0)$$

and thus by eq (2) that

$$f = (f(k; (F_d)_0) \times \{0\}) \cup (g(k; (F_d)_0) \times \{1\}).$$

From the explicit formulas for $f(n; F_d)$ and $g(n; F_d)$ given in the proof of theorem 3, we have

$$\begin{aligned} f(k; (F_d)_0) \times \{0\} &= [\cup \{B_a \times \{Z\} : Z \in f(k-d)\}] \times \{0\} \\ &= \cup \{B_a \times \{Z \times \{0\}\} : Z \in f(k-d)\} \\ &= B_d \times (f(k-d) \times \{0\}), \end{aligned}$$

where we have used the fact $(F_d)_0 = B_d \times \{0_{k-d}\}$, and similarly

$$g(k; (F_d)_0) \times \{1\} = B_d \times (g(k-d) \times \{1\}).$$

Thus

$$\begin{aligned} f &= B_d \times [(f(k-d) \times \{0\}) \cup (g(k-d) \times \{1\})] \\ &= B_d \times f(k+1-d) = f(k+1; F_d), \end{aligned}$$

as was to be proved.

It remains to prove that $f_1 = B_k - f_0$. Since f_0 and f_1 each have 2^{k-1} members, it suffices to prove $f_1 \subset B_k - f_0$, i.e., that no component p of f_0 meets any component q of f_1 . If such a meeting occurred, then by (d) of lemma 3 we have $p = q = h_0 = h_1$ for some component h of f , and by (b) of lemma 3 p is a component of f_1 as well as f_0 . By (c) of lemma 3, f_0 and f_1 are both cellular, and by induction hypothesis there is only one cellular function of k variables with 2^{k-1} members having p as a component. Therefore $f_1 = f_0$, so that $(F_d)_0$ is a component of f_1 as well as of $f_0 = f(k; (F_d)_0)$. Thus

$$\begin{aligned} B_d \times \{0_{k-d}\} \times B_1 &= [B_d \times \{0_{k-d}\} \times \{0\}] \cup [B_d \times \{0_{k-d}\} \times \{1\}] \\ &= [(F_d)_0 \times \{0\}] \cup [(F_d)_0 \times \{1\}] \\ &\subset [f_0 \times \{0\}] \cup [f_1 \times \{1\}] = f; \end{aligned}$$

since the first set is a face of B_{k+1} which properly contains $F_d = B_d \times \{0_{k+1-d}\}$, we have a contradiction to the hypothesis that F_d is a component of f . This completes the proof of theorem 4.

The last two theorems yield the following statement:

COROLLARY. *A cellular Boolean function of n variables (other than B_n) with the maximum number of members has all its component cells of the same dimension, d , and its complement is again such a function (with the same "d"). There are $2^{\binom{n}{d}}$ such functions*

^{10a} For consistency it is important that this representation and assumption match the corresponding ones in the proof of theorem 3, so that the explicit formula for $f(n; F_d)$ introduced in that proof can validly be used below.

¹¹ For $k=d$, the Cartesian product reduces to B_k .

for each $d < n$, and thus in all there are $2(2^n - 1)$ cellular functions of n variables with the maximum number of members.

The number $2^{\binom{n}{d}}$ is obtained by dividing the number $2^{n-d} \binom{n}{d}$ of d -dimensional faces of B_n by the number 2^{n-d-1} of these faces appearing in a maximum cellular function.

4. Perfect Gray Codes

An even simpler proof of lemma 2 and theorem 1, in section 2, can be given if we take as known the existence of a Gray code, i.e., an enumeration $\{X_i : 1 \leq i \leq 2^n\}$ of the members of B_n such that X_i and X_{i+1} are neighbors for $1 \leq i \leq 2^n = N$, and X_N and X_1 are neighbors. From this it is obvious that an isolated set cannot contain both X_i and X_{i+1} , or both X_N and X_1 , and so has at most 2^{n-1} members. It is also clear that the sets

$$F(n) = \{X_i : i \text{ is even}\}, \quad G(n) = \{X_i : i \text{ is odd}\},$$

if they are isolated, are the unique maximum isolated functions of n variables. But the defining property of a Gray code shows that the parity of the number of 1's in X_i is opposite to that for X_{i+1} , so that the number of 1's in every member of $F(n)$ has the same parity, while the opposite parity holds for all members of $G(n)$. Thus either $F(n) = f(n)$ and $G(n) = g(n)$, or $F(n) = g(n)$ and $G(n) = f(n)$; since $f(n)$ and $g(n)$ are easily proved to be isolated the same is true of $F(n)$ and $G(n)$, so that the conclusions of lemma 2 and theorem 1 have been established.

The inductive technique used in section 2 to prove lemma 2 and theorem 1, and in section 3 to prove theorem 4, can also be employed to establish the existence of a Gray code on B_n for all $n \geq 1$. We shall actually prove a stronger result, namely the existence of a class of Gray codes (to be called *perfect*) with certain additional properties which have been found helpful in aiding visualization. For example ((a) of theorem 5), for each d with $1 \leq d \leq n$ such a code partitions into 2^{n-d} successive subpaths, each essentially a perfect Gray code on some d -dimensional face of B_n . Also a rather explicit description can be given of the positions (in a perfect Gray code) of the n neighbors of any member of B_n whose position in the code is given; see corollary 1 to theorem 7.

For $0 \leq d \leq n$ and $1 \leq j \leq 2^{n-d} = N(n-d)$, it is convenient to define

$$S_j^d = \{i : (j-1)2^d < i \leq j2^d\},$$

a subset of $\{1, 2, 3, \dots, 2^n = N(n)\}$. For any subset S of $S_{2\nu-1}^d$, we call the subset

$$S' = \{i : (2\nu-1)2^{d+1} + 1 - i \in S\}$$

of $S_{2\nu}^d$ the *reflection* of S in $(2\nu-1)2^d$, while for any subset S of $S_{2\nu}^d$ the *reflection* S' of S in $(2\nu-1)2^d$ is the subset of $S_{2\nu-1}^d$ defined by the same formula. Here $0 \leq d < n$ and $1 \leq \nu \leq N(n-d-1)$. Familiarity with the reflected binary number system should provide partial motivation for this definition.

For any enumeration $\{X_i : 1 \leq i \leq 2^n = N(n)\}$ of the members of B_n , and any subset T of $\{1, 2, 3, \dots, N(n)\}$, let

$$X(T) = \{X_i : i \in T\}.$$

The enumeration will be called a *perfect Gray code* if, for each d with $0 \leq d < n$, for each ν with $1 \leq \nu \leq N(n-d-1)$, and for each subset S of either $S_{2\nu-1}^d$ or $S_{2\nu}^d$ such that $X(S)$ is a δ -dimensional face of B_n for some $\delta \leq d$, the set $X(S')$ corresponding to the reflection S' of S in $(2\nu-1)2^d$ is also a δ -dimensional face, and furthermore $X(S) \cup X(S')$ is a $(\delta+1)$ -dimensional face.

To indicate the implications of this fairly complicated definition, we collect some properties of perfect Gray codes in the next theorem, in which the notation

$$F_j^d = X(S_j^d)$$

is employed.

THEOREM 5. *If $\{X_i : 1 \leq i \leq N(n)\}$ is a perfect Gray code on B_n , then*

- (a) F_j^d is a d -dimensional face,
- (b) $F_j^d \cup F_{j+1}^d$ is a $(d+1)$ -dimensional face,
- (c) $F_j^d \cup F_{N(n-d)+1-j}^d$ is a $(d+1)$ -dimensional face.

First (a) will be proved for $0 \leq d \leq n$. It is clearly true for $d=0$; assume it true for $d=k < n$, and consider the case $d=k+1$. Observe that

$$S_j^{k+1} = S_{2j-1}^k \cup S_{2j}^k$$

and that each disjunct is the other's reflection in $(2j-1)2^k$. Since by induction hypothesis F_{2j-1}^k is a k -dimensional face, it follows from the definition of "perfectness" that F_j^{k+1} is a $(k+1)$ -dimensional face. This completes the inductive proof of (a).

To prove (c), it now suffices to observe that $S_{N(n-d)+1-j}^d$ is the reflection of S_j^d in $N(n-1)$. To prove (b), factor j as

$$j = 2^a b < N(n-d) \quad (a \geq 0, b \text{ odd})$$

and observe that S_j^d is a subset of S_b^{a+d} whose reflection in $b2^{a+d}$ is S_{j+1}^d .

By setting $d=0$ in (b) of theorem 5, and also in (c) with $j=1$, we see that a perfect Gray code is indeed a Gray code. An imperfect Gray code on B_3 , where for example F_1^2 is not a face, is given by

$$\begin{array}{cccc} (0,0,0), & (1,0,0), & (1,0,1), & (1,1,1) \\ (1,1,0), & (0,1,0), & (0,1,1), & (0,0,1). \end{array}$$

Next we prove the existence of a perfect Gray code on B_n ; in analogy with theorem 3 it will in fact be shown that such a code can be chosen to "begin" with any preassigned face F of B_n . Because the formal proof which follows is somewhat forbidding, we summarize the geometric idea first: B_n is considered as made up of two copies of B_{n-1} (for simplicity, $B_{n-1} \times \{0\}$ and $B_{n-1} \times \{1\}$), the first copy containing F ; the perfect Gray code on B_n is constructed by choosing an appropriate perfect Gray code on B_{n-1} (which exists by induction hypothesis), tracing it out in $B_{n-1} \times \{0\}$ and then tracing it out in

reverse order in $B_{n-1} \times \{1\}$. When $F = \phi$ this construction agrees with one given by Gilbert (op. cit. in footnote 9b), and if iterated for $n > 1$ it gives the particular perfect Gray codes which Gilbert termed the "conventional" ones.¹²

THEOREM 6. *Let F be a $d(0)$ -dimensional face of B_n , with $0 \leq d(0) \leq n$. There exists a perfect Gray code on B_n for which $F_1^{(0)} = F$.*

This is clearly true for $n=1$; assume it true for $n=k$ and consider the case $n=k+1$. If $d(0)=k+1$, so that $F=B_{k+1}$, then we are simply to prove the existence of a perfect Gray code on B_{k+1} , which will follow from the analysis for $d(0) \leq k$. So suppose $d(0) \leq k$; to simplify notation we can assume $F \subset B_k \times \{0\}$, so that F is given in terms of its projection F_0 as $F = F_0 \times \{0\}$.

It is easy to prove that F_0 is a $d(0)$ -dimensional face of B_k . By inductive hypothesis, there is a perfect Gray code $\{Y_i: 1 \leq i \leq N(k)\}$ on B_k for which

$$F_0 = \{Y_i: 1 \leq i \leq 2^{d(0)}\}.$$

Now define an enumeration $\{X_i: 1 \leq i \leq N(k+1)\}$ of B_{k+1} by

$$X_i = Y_i \times \{0\} \text{ if } 1 \leq i \leq N(k),$$

$$X_i = Y_{N(k+1)+1-i} \times \{1\} \text{ if } N(k) < i \leq N(k+1).$$

We shall show that this enumeration is the desired perfect Gray code. The following notation will be used:

$$X(S) = \{X_i: i \in S\},$$

$$Y(R) = \{Y_i: i \in R\},$$

$$S_j^d = \{i: (j-1)2^d < i \leq j2^d\} \quad (0 \leq d \leq k+1, \\ 1 \leq j \leq N(k+1-d)),$$

$$R_j^d = \{i: (j-1)2^d < i \leq j2^d\} \quad (0 \leq d \leq k, \\ 1 \leq j \leq N(k-d)),$$

$$F_j^d = X(S_j^d),$$

$$G_j^d = Y(R_j^d).$$

First, it is clear that

$$F = F_0 \times \{0\} = G_1^{d(0)} \times \{0\} = F_1^{d(0)},$$

as desired. Second, consider any d with $0 \leq d \leq k$, any ν with $1 \leq \nu \leq N(k-d)$, and any subset S of either $S_{2\nu-1}^d$ or $S_{2\nu}^d$ such that $X(S)$ is a δ -dimensional face of B_{k+1} for some $\delta \leq d$. Let S' be the reflection of S in $(2\nu-1)2^d$. The proof will be completed by showing that $X(S')$ is a δ -dimensional face, and $X(S) \cup X(S')$ a $(\delta+1)$ -dimensional face. It will be assumed that S is a subset of $S_{2\nu-1}^d$ rather than $S_{2\nu}^d$; the details for the other case are analogous.

Suppose first that $\nu \leq N(k-d-1)$. Then $X_i = Y_i \times \{0\}$ for each $i \in S \cup S'$. Therefore

$$X(S) = Y(S) \times \{0\}, \quad X(S') = Y(S') \times \{0\}, \\ X(S) \cup X(S') = (Y(S) \cup Y(S')) \times \{0\}.$$

The first of these equations shows that $Y(S)$ is a δ -dimensional face of B_k , so that $Y(S')$ is a δ -dimensional face and $Y(S) \cup Y(S')$ a $(\delta+1)$ -dimensional face; the second and third equations then yield the same results for $X(S)$ and $X(S) \cup X(S')$.

Next suppose $\nu > N(k-d-1)$ and $d < k$. Then

$$X_i = Y_{N(k+1)+1-i} \times \{1\}$$

for each $i \in S \cup S'$. Let R and R' be the respective reflections of S and S' in $N(k)$, so that

$$X(S) = Y(R) \times \{1\}, \quad X(S') = Y(R') \times \{1\} \\ X(S) \cup X(S') = (Y(R) \cup Y(R')) \times \{1\}.$$

The first of these equations shows that $Y(R)$ is a δ -dimensional face of B_k . It is readily verified that R' is the reflection of R in

$$[2(2^{k-d} - (\nu-1)) - 1]2^d,$$

so that $Y(R')$ is a δ -dimensional face and $Y(R) \cup Y(R')$ a $(\delta+1)$ -dimensional face; the second and third equations then yield the same results for $X(S')$ and $X(S) \cup X(S')$.

The final situation is $d=k$, so that $\nu=1$. Here

$$X_i = Y_i \times \{0\} \text{ for } i \in S,$$

$$X_i = Y_{N(k+1)+1-i} \times \{1\} \text{ for } i \in S'$$

so that

$$X(S) = Y(S) \times \{0\} \quad X(S') = Y(S) \times \{1\}, \\ X(S) \cup X(S') = Y(S) \times \{0, 1\},$$

and the result follows.

The next theorem gives more detail on the structure of perfect Gray codes.

THEOREM 7. *Let $\{X_i: 1 \leq i \leq N(n)\}$ be a perfect Gray code on B_n , and T a set of $2^{\delta+1}$ elements of $\{1, 2, \dots, N(n)\}$. If $X(T)$ is a $(\delta+1)$ -dimensional face of B_n , then there exists d with $\delta \leq d < n$, ν with $1 \leq \nu \leq N(n-d-1)$, and a partition $T = S \cup S'$ such that S is a subset of either $S_{2\nu-1}^d$ or $S_{2\nu}^d$, S' is the reflection of S in $(2\nu-1)2^d$, and $X(S)$ is a δ -dimensional face.*

This is clear for $n=1$; assume it true for $n=k$, and consider the case $n=k+1$. By (a) of Theorem 5, F_1^k is a k -dimensional face of B_{k+1} , and to simplify notation we may assume $F_1^k = B_k \times \{0\}$. Thus $F_2^k = B_k \times \{1\}$. If we define Y_i to be the projection $(X_i)_0$ for $1 \leq i \leq N(k)$, then by (c) of theorem 5 $X_{N(k+1)+1-i} \in B_k \times \{1\}$ is a neighbor of $X_i \in B_k \times \{0\}$, which implies that $Y_i = (X_{N(k+1)+1-i})_1$. That is,

$$X_i = Y_i \times \{0\} \text{ for } 1 \leq i \leq N(k) \\ X_i = Y_{N(k+1)+1-i} \times \{1\} \text{ for } N(k) < i \leq N(k+1).$$

That $\{Y_i: 1 \leq i \leq N(k)\}$ is a perfect Gray code on B_k follows easily from the corresponding fact for $\{X_i: 1 \leq i \leq N(k+1)\}$ and the first of the last two equations.

Suppose first that $T \subset S^k$. Then

$$X(T) = Y(T) \times \{0\}$$

¹² In theorem 6 below we write $d(0)$ instead of d_0 for typographical convenience.

and so $Y(T)$ is a $(\delta+1)$ -dimensional face of B_k . By induction hypothesis there exist d with $\delta \leq d < k$, ν with $1 \leq \nu \leq N(k-d-1)$, and a partition $T = S \cup S'$ such that S is a subset of either $S_{2\nu-1}^d$ or S_{22}^d , S' is the reflection of S in $(2\nu-1)2^d$, and $Y(S)$ is a δ -dimensional face of B_k . Thus

$$X(S) = Y(S) \times \{0\}$$

is a δ -dimensional face of B_{k+1} , as desired.

Next suppose $T \subset S_2^k$. Let T' be the reflection of T in $N(k)$; then

$$X(T) = Y(T') \times \{1\}$$

and so $Y(T')$ is a $(\delta+1)$ -dimensional face of B_k . By induction hypothesis there exist d with $\delta \leq d < k$, ν with $1 \leq \nu \leq N(k-d-1)$, and a partition $T' = S \cup S'$ such that S is a subset of either $S_{2\nu-1}^d$ or $S_{2\nu}^d$, S' is the reflection of S in $(2\nu-1)2^d$, and $Y(S)$ is a δ -dimensional face of B_k . Let R and R' be the respective reflections of S and S' in $N(k)$, so that we have a partition $T = R \cup R'$, and

$$X(R) = Y(S) \times \{1\}$$

is a δ -dimensional face of B_{k+1} . Since R is a subset of either $S_{2\mu-1}^d$ or $S_{2\mu}^d$ where

$$\mu = 2^{k-d} - \nu + 1,$$

and R' is the reflection of R in $(2\mu-1)2^d$, again matters are as desired.

Finally, suppose T meets both S_1^k and S_2^k . Then $X(T)$, as a $(\delta+1)$ -dimensional face of B_{k+1} which meets both $B_k \times \{0\}$ and $B_k \times \{1\}$, can easily be shown to have the form

$$X(T) = F \times \{0, 1\}$$

where F is a δ -dimensional face of B_k . Say $F = Y(S)$ where $S \subset S_1^k$. It follows that

$$T \cap S_1^k = S, \quad T \cap S_2^k = S'$$

where S' is the reflection of S in $N(k)$. Since $X(S) = F \times \{0\}$ is a δ -dimensional face of B_{k+1} , the proof is complete.

COROLLARY 1. *Let $\{X_i : 1 \leq i \leq N(n)\}$ be a perfect Gray code on B_n . Then the n neighbors of X_1 are the $\{X_{k(d)} : 0 \leq d < n\}$ defined by*

$$k(d) = (2\nu(d) - 1)2^{d+1} + 1 - i$$

where $\nu(d)$ is defined by

$$(2\nu(d) - 2)2^d < i \leq 2\nu(d)2^d.$$

This follows from theorem 7 with $\delta = 0$.

COROLLARY 2. *Let $\{X_i : 1 \leq i \leq N(n)\}$ be a perfect Gray code on B_n . If $j(1)$ and $j(2)$ are distinct but have the same parity, then no member of $F_{j(1)}^\delta$ is a neighbor of a member of $F_{j(2)}^\delta$.*

For, suppose a member X_i of $F_{j(1)}^\delta$ and a member X_k of $F_{j(2)}^\delta$ were neighbors, so that

$$(j(1) - 1)2^\delta + 1 \leq i \leq j(1)2^\delta,$$

$$(j(2) - 1)2^\delta + 1 \leq k \leq j(2)2^\delta,$$

and by corollary 1 there exist d with $0 \leq d < n$, and $\nu = \nu(d)$, such that

$$i + k = (2\nu - 1)2^{d+1} + 1$$

and

$$(\nu - 1)2^{d+1} + 1 \leq i \leq \nu 2^{d+1},$$

which imply that

$$(\nu - 1)2^{d+1} + 1 \leq k \leq \nu 2^{d+1}.$$

On the one hand, we have

$$\begin{aligned} (j(2) - 1)2^\delta + 1 - j(1)2^\delta &\leq k - i \\ &= (i + k) - 2i \\ &\leq [(2\nu - 1)2^{d+1} + 1] \\ &\quad - 2[(\nu - 1)2^{d+1} + 1], \end{aligned}$$

so that

$$[j(2) - j(1) - 1]2^\delta + 1 \leq 2^{d+1} - 1$$

and therefore

$$j(2) - j(1) < 2^{d+1-\delta} + 1.$$

On the other hand by interchanging i with k and $j(1)$ with $j(2)$ in this argument, we obtain

$$j(1) - j(2) < 2^{d+1-\delta} + 1.$$

Since the hypotheses on $j(1)$ and $j(2)$ imply $|j(2) - j(1)| \geq 2$, the last two displays imply that $d+1-\delta > 0$. But then we would have

$$\begin{aligned} [(j(1) - 1)2^\delta + 1] + [(j(2) - 1)2^\delta + 1] &\leq i + k \\ &= (2\nu - 1)2^{d+1} + 1 \\ &= i + k \\ &\leq j(1)2^\delta + j(2)2^\delta, \end{aligned}$$

so that

$$\begin{aligned} [j(1) + j(2) - 2]2^\delta + 2 &\leq (2\nu - 1)2^{d+1} + 1 \leq [j(1) + j(2)]2^\delta, \\ [j(1) + j(2) - 2]2^\delta &< (2\nu - 1)2^{d+1} < [j(1) + j(2)]2^\delta, \\ j(1) + j(2) - 2 &< (2\nu - 1)2^{d+1-\delta} < j(1) + j(2), \end{aligned}$$

with the hypotheses implying that the middle term (which lies between the consecutive even integers constituting the extreme terms) is an even integer. Since this is impossible, the corollary is proved.

Just as the existence of ordinary Gray codes provides an alternative approach to the material of section 2 (see the beginning of the present section), so the existence and properties of perfect Gray codes

provide an alternative approach to the material of section 3. Suppose for example that F_d is a d -dimensional face of B_n , with $0 \leq d < n$. Then by theorem 6 there is a perfect Gray code on B_n with $F_d = F_1^d$, and theorem 3 can be quickly proved by setting

$$f(n; F_d) = \cup \{F_j^d : 1 \leq j \leq N(n-d), j \text{ is odd}\},$$

$$g(n; F_d) = \cup \{F_j^d : 1 \leq j \leq N(n-d), j \text{ is even}\}$$

and invoking the last corollary (with d as δ) together with (a) of theorem 5.

We conclude by showing that "essentially" there is only one perfect Gray code on B_n , which must therefore be the "conventional" one mentioned earlier. To make this precise, observe that as a "discrete cube," B_n has as its natural group of symmetries the hyperoctahedral group \mathcal{O}_n whose $(n!)2^n$ elements can be represented as the ordered pairs (σ, π) where σ is a subset of $\{1, 2, 3, \dots, n\}$ and π is a permutation of $\{1, 2, 3, \dots, n\}$. The operation of (σ, π) on a member $X = (x_1, \dots, x_n)$ of B_n is the member X' defined by

$$x'_j = 1 - x_{\pi(j)} \text{ if } j \in \sigma,$$

$$x'_j = x_{\pi(j)} \text{ if } j \in \{1, 2, 3, \dots, n\} - \sigma.$$

If $\{X_i : 1 \leq i \leq N(n)\}$ and $\{X'_i : 1 \leq i \leq N(n)\}$ are perfect Gray codes on B_n , we call them *equivalent* in case (with an obvious notation)

$$(\sigma, \pi)(X_i) = X'_i \text{ for } 1 \leq i \leq N(n)$$

for some (σ, π) in \mathcal{O}_n .

THEOREM 8. Any two perfect Gray codes on B_n are equivalent.

This is clear for $n=1$; assume it true for $n=k$, and consider the case $n=k+1$. Let $\{X_i : 1 \leq i \leq N(k+1)\}$ and $\{X'_i : 1 \leq i \leq N(k+1)\}$ be perfect Gray codes on B_{k+1} . By theorem 5 $X(S_1^k)$ and $X'(S_1^k)$ are k -dimensional faces of B_{k+1} , so that there exist elements r and t of $\{1, 2, \dots, k+1\}$, and numbers c and c' each either zero or one, for which

$$X(S_1^k) = \{X : X \epsilon B_{k+1}, x_r = c\},$$

$$X'(S_1^k) = \{X : X \epsilon B_{k+1}, x_t = c'\}.$$

Therefore one-to-one correspondences between B_k and each of $X(S_1^k)$ and $X'(S_1^k)$, are set up by the functions

$$g : X(S_1^k) \rightarrow B_k, \quad g' : X'(S_1^k) \rightarrow B_k$$

defined by

$$g(X) = (x_{\rho(1)}, x_{\rho(2)}, \dots, x_{\rho(k)}),$$

$$g'(X) = (x_{\tau(1)}, x_{\tau(2)}, \dots, x_{\tau(k)})$$

where $X = (x_1, \dots, x_{k+1})$ and the permutations ρ and τ of $\{1, 2, \dots, k+1\}$ are given by

$$\rho(j) = j \text{ if } j < r, \quad \rho(j) = j+1 \text{ if } r \leq j \leq k, \quad \rho(k+1) = r,$$

$$\tau(j) = j \text{ if } j < t, \quad \tau(j) = j+1 \text{ if } t \leq j \leq k, \quad \tau(k+1) = t.$$

It is trivial to verify that g transforms each d -dimensional face of B_{k+1} which lies wholly in $X(S_1^k)$ into a d -dimensional face of B_k , while g' does the same for those faces of B_{k+1} which lie wholly in $X'(S_1^k)$. Therefore

$$\{g(X_i) : 1 \leq i \leq N(k)\} \text{ and } \{g'(X'_i) : 1 \leq i \leq N(k)\}$$

are perfect Gray codes on B_k , so by induction hypothesis there exist a subset σ of $\{1, 2, \dots, k\}$ and a permutation π of $\{1, 2, \dots, k\}$ such that

$$(\sigma, \pi)(g(X_i)) = g'(X'_i) \text{ for } 1 \leq i \leq N(k).$$

If x_{ij} and x'_{ij} denote the j th entries of X_i and X'_i respectively, then this can be written as

$$x'_{i, \tau(j)} = 1 - x_{i, \rho\pi(j)} \text{ if } j \in \sigma$$

$$x'_{i, \tau(j)} = x_{i, \rho\pi(j)} \text{ if } j \in \{1, 2, \dots, k\} - \sigma$$

for $1 \leq j \leq k$ and $1 \leq i \leq N(k)$.

Now form the permutation π' of $\{1, 2, \dots, k+1\}$ defined by $\pi'(t) = r$ and

$$\pi'(j) = \rho\pi\tau^{-1}(j) \text{ for } j \neq t,$$

as well as the subset σ' of $\{1, 2, \dots, k+1\}$ defined by

$$\sigma' = \tau(\sigma) \cup \{t\} \text{ if } c \neq c',$$

$$\sigma' = \tau(\sigma) \text{ if } c = c'.$$

Then it follows easily that

$$(\sigma', \pi')(X_i) = X'_i \text{ for } 1 \leq i \leq N(k),$$

and this remains true for $N(k) < i \leq N(k+1)$ because

$$x_{N(k+1)+1-i, j} = x_{ij} \text{ if } j \neq r,$$

$$x_{N(k+1)+1-i, r} = 1 - x_{ir},$$

$$x'_{N(k+1)+1-i, j} = x'_{ij} \text{ if } j \neq t,$$

$$x'_{N(k+1)+1-i, t} = 1 - x'_{it}$$

for $1 \leq i \leq N(k)$. Thus the two given perfect Gray codes on B_{k+1} are equivalent.

It is a pleasure to acknowledge helpful suggestions by K. Goldberg and E. Johnson (NBS Numerical Analysis Section) and R. Kirsch (NBS Components and Techniques Section).

(Paper 67B2-94)