Transition Restricted Gray Codes

BETTE BULTENA and FRANK RUSKEY
Department of Computer Science
University of Victoria
Victoria, B.C. V8W 3P6, Canada
abultena@csr.uvic.ca, fruskey@csr.uvic.ca
Submitted: November 16, 1995; Accepted: March 14, 1996.

Abstract

A Gray code is a Hamilton path H on the n-cube, Q_n. By labeling each edge of Q_n with the dimension that changes between its incident vertices, a Gray code can be thought of as a sequence $H = t_1, t_2, \ldots, t_{N-1}$ (with $N = 2^n$ and each t_i satisfying $1 \le t_i \le n$). The sequence H defines an (undirected) *graph of transitions*, G_H, whose vertex set is $\{1, 2, \ldots, n\}$ and whose edge set $E(G_H) = \{[t_i, t_{i+1}] \mid 1 \le i \le N-1\}$. A G-code is a Hamilton path H whose graph of transitions is a subgraph of G; if H is a Hamilton cycle then it is a cyclic G-code. The classic binary reflected Gray code is a cyclic $K_{1,n}$-code. We prove that every tree T of diameter 4 has a T-code, and that no tree T of diameter 3 has a T-code.

Mathematical Reviews Subject Number: 05C45.

1 Introduction

The utility of the ubiquitous binary reflected Gray code is undisputed. See, for example, the books of Nijenhuis and Wilf [5], Reingold, Nievergelt, and Deo [6], and Wilf [8]. For certain applications, however, other Gray codes are desired. Many other Gray codes have been proposed, both for specific values of n and general constructions. For example, Goddyn, Lawrence, and Nemeth [3], motivated by an issue in the design of photon detectors, study the problem of finding a Gray code that maximizes the minimum number of edges between the use of edges of the same dimension. In a recent paper, Savage and Winkler [7] find a Gray code in which all subsets of size k appear before any of size $k+2$, and use this Gray code to improve the best known results on the notorious "middle levels" problem.

The n-cube, Q_n, has vertex set, $V(Q_n)$, consisting of all bitstrings of length n. Its edge set, $E(Q_n)$, consists of all pairs, $[\mathbf{x}, \mathbf{y}]$[1], in which $\mathbf{x}$ and $\mathbf{y}$ have Hamming distance one; i.e., that differ in exactly one position. For $[\mathbf{x}, \mathbf{y}] \in E(Q_n)$, define $\tau(\mathbf{x}, \mathbf{y})$, the *transition* between $\mathbf{x}$ and $\mathbf{y}$, to be the position of the bit that is different. A *Gray code* is a Hamilton path H on the n-cube, Q_n; a *cyclic Gray code* is a Hamilton cycle on Q_n. Given a Gray code $H = \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N$, where $N = 2^n$, its *transition sequence*, $\tau(H)$, is the sequence $\tau(H) = t_1, t_2, \ldots, t_{N-1}$, where $t_i = \tau(\mathbf{b}_i, \mathbf{b}_{i+1})$. If H is a cyclic Gray code, then in a *cyclic transition sequence* we append the additional transition $t_N = \tau(\mathbf{b}_N, \mathbf{b}_1)$. The transitions t_{i-1} and t_i are said to *delimit* the bitstring $\mathbf{b}_i$. The sequence $\tau(H)$ defines an (undirected) *graph of transitions*, G_H, whose vertex set is $[n] = \{1, 2, \ldots, n\}$ and whose edge set is $E(G_H) = \{[t_i, t_{i+1}] \mid 1 \le i \le N-1\}$.

A motivation for this paper was the construction of a particular type of Hamilton cycle in the cube-connected-cycle, CCC_n. A cube-connected-cycle is a certain cubic Cayley graph of the wreath product of an n-cycle and an n-cube, and is a well-known topology for computer networks. Formally, the vertex set $V(CCC_n)$ consists of all pairs $(d; \mathbf{x})$ where $\mathbf{x}$ is a length n bitstring and $d \in \mathbb{Z}_n$ is an integer mod n. The edge set $E(CCC_n)$ is as shown below.

$$
\begin{aligned}
E(CCC_n) &= \{[(d;\mathbf{x}), (d+1;\mathbf{x})] \mid \mathbf{x} \in V(Q_n) \text{ and } 0 \le d < n\} \ \cup \\
&\quad\ \{[(d;\mathbf{x}), (d;\mathbf{y})] \mid [\mathbf{x}, \mathbf{y}] \in E(Q_n) \text{ and } 0 \le d < n\}
\end{aligned}
$$

A basic fact about cube-connected-cycles is that they are Hamiltonian; an easily understood proof may be found Leighton [4]. We would like to find a particular type of Hamilton cycle — one that contiguously traverses all vertices of the form $(d; \mathbf{x})$, for $\mathbf{x}$ fixed, before moving onto a vertex with a different value of $\mathbf{x}$. Such a Hamilton cycle can be written in the form $\kappa_1, \kappa_2, \ldots, \kappa_N$, with $N = 2^n$, where each κ_i has the form (arithmetic done mod n)

$$(d_i; \mathbf{x}_i), (d_i + 1; \mathbf{x}_i), (d_i + 2; \mathbf{x}_i), \ldots, (d_i + (n-1); \mathbf{x}_i), \text{ or}$$

$$(d_i; \mathbf{x}_i), (d_i - 1; \mathbf{x}_i), (d_i - 2; \mathbf{x}_i), \ldots, (d_i - (n-1); \mathbf{x}_i).$$

The sequence $x_1, x_2, \ldots, x_N$ must be a Hamilton cycle in Q_n, and $d_1, d_2, \ldots, d_N$ is its transition sequence. Clearly, we must have $d_i = d_{i-1} \pm 1$.

The question naturally occurs as to whether a Hamilton cycle H in the n-cube can be "lifted" to a Hamilton cycle in the cube-connected cycle CCC_n by always traversing sucessively all vertices of an n-cycle, either in increasing or decreasing order, as described above. Such a Hamilton cycle exists if and only if the dimension of the successive edges used in H differs by one (mod n). This question is a special case of the problems considered below.

[1] Lower case bold letters always denote bitstrings.

Given any graph G, with n vertices, a G-*code* is a Hamilton path on Q_n whose graph of transitions is a subgraph of G. A *cyclic* G-*code* is a Hamilton cycle whose graph of transitions is a subgraph of G. Every G-code has a unique G-*transition sequence*, whose length is 2^N if the G-code is cyclic, and is $2^N - 1$ otherwise. In either case, we refer to the starting vertex and ending vertex as those numbered 1 and $2^N - 1$ in the sequence. Note that to prove the existence of a G-code for graph G, it is sufficient to produce a G-transition sequence. The question raised in the previous paragraph amounts to that of finding a cyclic C_n-code, where C_n denotes an n-cycle.

The classic binary reflected Gray code (BRGC) has transition sequence T_n defined recursively as follows. The initial condition is $T_0 = 1$ and if $n > 0$, then $T_n = T_{n-1}, n, T_{n-1}$. For example, $T_4 = 1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1$. The cyclic transition sequence for the BRGC is T_n, n. For the BRGC the graph of transitions is $K_{1,n-1}$. The graph $K_{1,n-1}$ for $n \geq 3$ is called a *star*, denoted S_n. It is a tree with one *central* vertex and $n - 1$ *leaves*.

The following useful theorem is due to Gilbert [2].

Theorem 1.1 *The following statements characterize non-cyclic and cyclic transition sequences:*

- *Let* T *be a sequence of* $2^n - 1$ *integers from* $[n]$. *The sequence* T *is a transition sequence if and only if every non-empty consecutive subsequence of* T *contains some integer an odd number of times.*

- *Let* T *be a sequence of* 2^n *integers from* $[n]$. *The sequence* T *is a cyclic transition sequence if and only if every non-trivial consecutive subsequence of* T *contains some integer an odd number of times, and every integer in* $[n]$ *occurs in* T *an even number of times.*

We will also think in an algebraic sense of transitions as operations on bitstrings, so that $\mathbf{b}t$ denotes the bitstring that results from flipping the t^{th} bit of $\mathbf{b}$. We note the following obvious equalities. The operation is commutative: $\mathbf{b}st = \mathbf{b}ts$. The operation is an involution: $\mathbf{b}tt = \mathbf{b}$ or, equivalently, if $\mathbf{b}t = \mathbf{c}$, then $\mathbf{c}t = \mathbf{b}$.

A graph G is *completely Gray* if there is a G-transition sequence starting from every vertex of G. Observe that if there is a cyclic G-code, then G is completely Gray. A graph G is *Gray-connected* if, for every pair of vertices u and v, there is a G-transition sequence starting at u and ending at v. A bipartite graph G is *Gray-laceable* if, for every pair of vertices u and v in the same partite set, there is a G-transition sequence starting at u and ending at v.

The *diameter* of a graph G is the maximum distance between any two vertices of G. For a tree T the diameter is thus the length of the longest path in T. A *center* of a tree is a vertex of minimum distance from all others. Every tree has one or two centers, and if it has two, then they are adjacent. A tree with one center is *unicentral*, a tree with two is *bicentral*.

We will refer to several standard graphs. The complete graph is denoted K_n. The complete bipartite graph is denoted $K_{n,m}$. The *path* with n vertices (and $n-1$ edges) is denoted P_n. The *cycle* of length n is denoted C_n.

2 Undirected G-codes

Below are two simple lemmata.

Lemma 2.1 *If there is a* G*-code, then* G *is connected.*

Lemma 2.2 *If there is a* G*-transition sequence starting at* u*, then it must end at a vertex* v *of* G *for which there is an even length path from* u *to* v*. In particular, if* G *is bipartite, then a* G*-transition sequence must start and end at vertices in the same partite set.*

Proofs: Every bit must change at least once in a G-code. This means that each vertex is encountered at least once in the G-transition sequence. Therefore, G must be connected.

If a G-transition sequence starts at u, then it lists $2^n - 2$ other vertices, which is an even number. By the property of bipartite graphs, the last vertex is in the same partite set as u. □

Stars have interesting properties and are useful as subgraphs in succeeding constructions.

Lemma 2.3 *(a) There is a cyclic* S_n*-transition sequence which starts and ends at the central vertex. (b) For any leaves* u *and* v *of* S_n*, there is a cyclic* S_n*-transition sequence which starts at* u *and ends at* v *if and only if* $u \neq v$*.*

Proof: (a) Note that the BRGC's graph of transitions is S_n, and that T_n begins and ends on the central vertex 1.

(b) To show the necessity of $u \neq v$, let Σ be the S_n-transition sequence that starts at u and ends at v. Let r be the central vertex. We can create a cyclic transition sequence, $\Sigma_c = t_1, t_2, \ldots, t_N$, with $u = t_1$, $v = t_{N-1}$, and $r = t_N$, where $N = 2^n$, that meets the requirements of Theorem 1.1. Because Σ_c is cyclic, we can start on t_{N-1}, producing a subsequence $t_{N-1}, t_N, t_1, t_2 = vrur$, which is illegal when $u = v$.

Consider S_n, with distinct leaves u and v, and central vertex, r. Relabel r as v and remove the leaf v, creating S_{n-1}. Using the BRGC and starting at u, we can form a cyclic

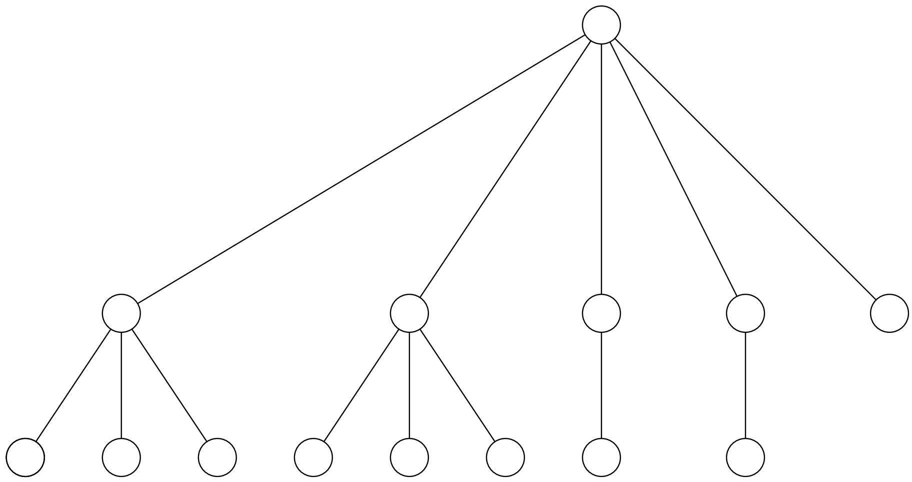

Figure 1: The short tree $T_5(3, 3, 1, 1, 0)$.

transition sequence, $\Sigma_c = uv \cdots vpv$, where p is any leaf in S_{n-1}, and the length of the sequence is 2^{n-1}. We can then place an r between each t_i, t_{i+1} pair, $(1 \leq i \leq 2^{n-1} - 1)$, producing a new sequence $\Sigma = urvr \cdots rvrprv$, with length $2^n - 1$. Note that any proper subsequence of the original sequence contains some vertex an odd number of times. Adding the r's maintains this property, and by appending an r to the end of Σ, we create the condition that every vertex is visited an even number of times. This means that we can create a cyclic transition sequence from Σ. The graph of transitions for Σ is clearly S_n and the starting and ending leaves have been chosen arbitrarily. $\square$

Corollary 2.1 *The complete graph K_n is Gray-connected.*

Proof: Note that K_n contains S_n as a spanning tree. Because you can choose any vertex as the root or any two distinct vertices as leaves, all combinations of start and finish vertices can produce a G-transition sequence. $\square$

Define a *short tree* to be a rooted tree of height at most 2, where the height of a tree is the length of the longest path from the root to a leaf. A short tree may be specified by the number of children of each of the nodes at level 1. Let $T_t(n_1, n_2, \ldots, n_t)$, where $n_1 \geq n_2 \geq \cdots \geq n_t$ be the unique short tree in which the i-th node at level 1 has n_i children. For example, Figure 1 shows the tree $T_5(3, 3, 1, 1, 0)$. The number of nodes in the tree $T_t(n_1, n_2, \ldots, n_t)$ is $1 + t + n_1 + n_2 + \cdots + n_t$. Every free tree of diameter at least 2 and at most 4 can be made into a short tree by taking a center to be the root.

Theorem 2.2 *For every $t \geq 2$ and $n \geq 1$ there is a cyclic $T_t(n, 1, 0, \ldots, 0)$-code.*

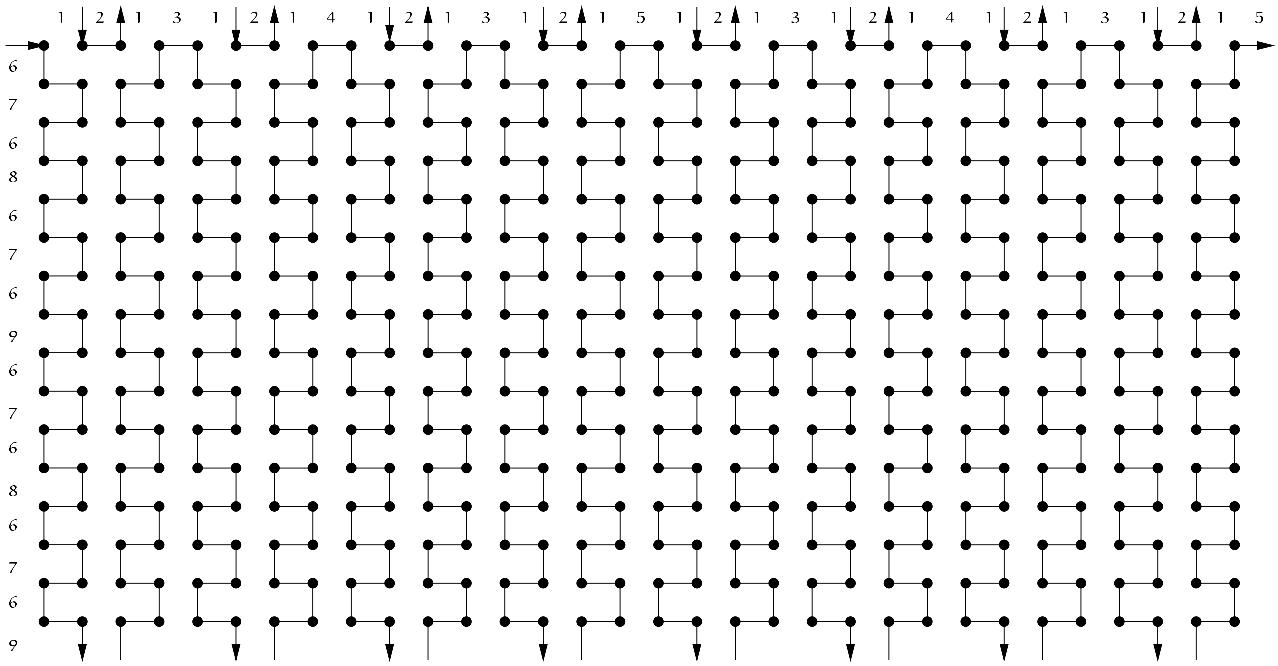

Figure 2: How to weave together the BRGC sequences T_4 and T_5 to get a $T_4(3,1,0,0)$-code.

Proof: We weave together a cyclic BRGC transition sequence over Q_t together with one over Q_{n+2}. Thinking of Q_{t+n+2} as $Q_t \times Q_{n+2}$, one sees that the Hamilton cycles given by the BRGC for t and $n+2$ induce a 2^t by 2^{n+2} toroidal grid structure in Q_{t+n+2}.

Assume that the nodes are labelled so that 1 is the root, that its children are $n+3, \ldots, n+t+2$, that $n+t+2$ has only child 2, and that the children of $n+3$ are $3, 4, \ldots, n+2$. Label the rows of the grid with the cyclic transition sequence for the BRGC, using $n+3, \ldots, n+t+2$, and the columns with $1, 2, \ldots, n+2$. On this toroidal spanning subgraph, we weave a specific Hamilton cycle which starts at the left upper point and follows the patterns illustrated in Figure 2, which uses $t = 4$ and $n+2 = 5$. Beginning with a vertical move, we zigzag down the 2^t rows, wrapping round in column 2, then zigzag back up. This up-down zig-zag is repeated 2^n times and ends at the right upper point, which attaches to the left upper point to create the Hamilton cycle on Q_{t+n+2}.

In the columns, note that if a transition is not 1 (these transitions occur only in the first and last rows), then it is adjacent only to $n+3$ or $n+t+2$; furthermore, it is only 2, which occurs 2^n times as a column, that is adjacent to $n+t+2$; the remaining vertices $3, 4, \ldots, n+2$ are adjacent only to $n+3$. □

Corollary 2.3 *For any complete multipartite graph* G*, there is a cyclic* G*-code.*

Proof: Let the number of vertices in G be n. Consider any 2 of the partite sets, R and P, in G, where $|P| = m$: If either set has only 1 vertex, then G contains S_n as a spanning tree and we can generate a BRGC from it. If R and P each contain 2 vertices, and G is bipartite, then the sequence $1, 2, 3, 2, 1, 2, 3, 4, 3, 2, 1, 2, 3, 2, 1, 4$, using vertices labelled sequentially on the 4-cycle, can be used.

Otherwise, choose vertex $r \in R$ and mark its edges that are incident with a vertex in P. Choose another vertex in R and mark its edge to some specific vertex $s \in P$. Choose another vertex $q \neq s \in P$ and mark its incident edges to every remaining unconsidered vertex in R and the other partite sets. We can see that the marked edges of G constitute a spanning $T_m(n - m - 2, 1, 0, \ldots, 0)$ subgraph of G, with r as the root. By Theorem 2.2, this spanning subgraph guarantees a cyclic G-code. □

Theorem 2.4 *For every tree* T *of diameter 4, there is a cyclic* T*-code.*

Proof: Let $m = n_1 + n_2 + \cdots + n_t$ be the total number of leaves in $T = T_t(n_1, n_2, \ldots, n_t)$, and let $n = 1 + t + m$ be the total number of vertices. Consider the multipartite graph G whose partite sets have sizes $n_1, n_2, \ldots, n_t$; label its vertices with the corresponding leaves of $T_t(n_1, n_2, \ldots, n_t)$. By Corollary 2.3, there is a cyclic G-transition sequence, $\Gamma = g_1, g_2, \ldots, g_M$, where $M = 2^m$.

Now, consider S_{t+1}, with central vertex being the root of $T_t(n_1, n_2, \ldots, n_t)$ and the leaves being the children of the root: By Lemma 2.3, we can choose any 2 distinct leaves as a start-finish combination, to produce a S_{t+1}-transition sequence of length $2^{t+1} - 1$. Choose such a sequence, Σ_i, for each g_i, g_{i+1} pair in Γ, such that the starting vertex is adjacent to g_i and the ending vertex is adjacent to g_{i+1}, using $T_t(n_1, n_2, \ldots, n_t)$ to determine adjacencies. Also, choose a sequence Σ_M such that the starting vertex is adjacent to g_M and the finish vertex is adjacent to g_1. Form a new sequence, $\Omega = g_1, \Sigma_1, g_2, \Sigma_2, \ldots, g_M, \Sigma_M$ of length $2^m(2^{t+1} - 1) + 2^m = 2^n$.

We show that Ω is the cyclic transition sequence for a cyclic $T_t(n_1, n_2, \ldots, n_t)$-code: Each g_i bit change is matched with a Gray code corresponding to the S_{t+1}-bits, and the g_i's themselves are a cyclic G-transition sequence. Essentially, we are tracing out a H-path in a 2^m by 2^{t+1} grid graph.

To show that the transition sequence is cyclic, we investigate the parity of all the vertices in the Σ_i's. Each Σ_i has length $2^{t+1} - 1$ and can be made cyclic by appending an r to it, meaning it encounters only the central vertex (the original root) an odd number of times. But since there are an even number of Σ_i's, the total number of r's in Ω is even. Therefore, by Theorem 1.1, we know Ω is a cyclic transition sequence. □

Theorem 2.5 *For every tree* T *of diameter 3, there does not exist a cyclic* T*-code.*

Proof: Given the diameter 3 tree $T = T_t(m, 0, \ldots, 0)$, let r and f be its centers, and let the leaves adjacent to f be $w_1, w_2, \ldots, w_m$. Assume that $H = \mathbf{b}_1, \mathbf{b}_2 \ldots, \mathbf{b}_N$ is a cyclic T-code, whose cyclic transition sequence is $\tau(H)$; we will derive a contradiction.

Divide $\tau(H)$ into all the subsequences which begin with and contain exactly 1 occurence of any w_i; subsequences of three possible forms result: The A-type may exist and has the form $w_i f$. The B-type may exist and has the form $w_i frf$. The C-type is necessary and has the form $w_i fr \cdots rf$.

Since $\tau(H) = t_1 t_2 \cdots t_N$ is cyclic we may assume that it begins with a subsequence that is a C-type. Let k be the index of the last f in this subsequence. Let $\mathbf{x}' = O^n t_1 t_2 \cdots t_{k-2}$, $\mathbf{x} = \mathbf{x}' t_{k-1} = \mathbf{x}' r$, $\mathbf{y} = \mathbf{x} t_k = \mathbf{x} f$, and $\mathbf{y}' = \mathbf{y} r = \mathbf{x}' f$.

There must be a unique index j for which $\mathbf{y}' = \mathbf{b}_j$. Denote $\mathbf{u} = \mathbf{b}_{j-1}$ and $\mathbf{v} = \mathbf{b}_{j+1}$. There are four mutually exclusive cases to consider since either r or a w_i, but not both, must delimit $\mathbf{y}'$, either on the right or on the left. These cases are (a) $\tau(\mathbf{u}, \mathbf{y}') = r$, (b) $\tau(\mathbf{y}', \mathbf{v}) = r$, (c) $\tau(\mathbf{u}, \mathbf{y}') = w_i$, and (d) $\tau(\mathbf{y}', \mathbf{v}) = w_i$. In case (a), we must have $\mathbf{u} = \mathbf{y}' r = \mathbf{y}$, and in case (b), we must have $\mathbf{v} = \mathbf{y}' r = \mathbf{y}$. But both these cases cannot occur as $\mathbf{y}$ by definition is delimited by f and w_i, and not by r. In case (c), we must have $\mathbf{b}_{j+1} = \mathbf{y}' f = \mathbf{x}'$ which means that the subsequence, $w_i frf$ is contained within this subsequence. But by definition, a C-type subsequence cannot contain a B-type subsequence. In case (d), we conclude that $\mathbf{b}_{j-1} = \mathbf{y}' f = \mathbf{x}'$. This means that $\mathbf{x}'$ is delimited on the right by f, which is not true. □

Conjecture 2.1 *If T is a tree then there is a cyclic T-code if and only if the diameter of T is 2 or 4.*

If the conjecture is true, then for $n \geq 6$, P_n-codes do not exist, where P_n is a path of length n. A natural way to add edges is to take the square of the graph.

Question 2.2 *Does a P_n^2-code exist for all $n \geq 1$?*

The question asked in the introduction about Gray codes being lifted to Hamilton cycles on the cube-connected cycle is equivalent to the following question.

Question 2.3 *Does a cyclic C_n-code for $n \geq 5$ exist?*

2.1 Experimental results

We have obtained a complete classification of all G-codes for graphs on at most 6 vertices, and for some graphs on 7 vertices. Refer to Figure 3. Any graph G of order at most six not illustrated contains, as a spanning subgraph, one or more of these graphs having a cyclic Gray Code, and thus itself possesses a cyclic G-code. For example, any 6 vertex connected graph with 6 edges, which is not a 6-cycle, must contain (6.2) or (6.3) as a spanning subgraph.

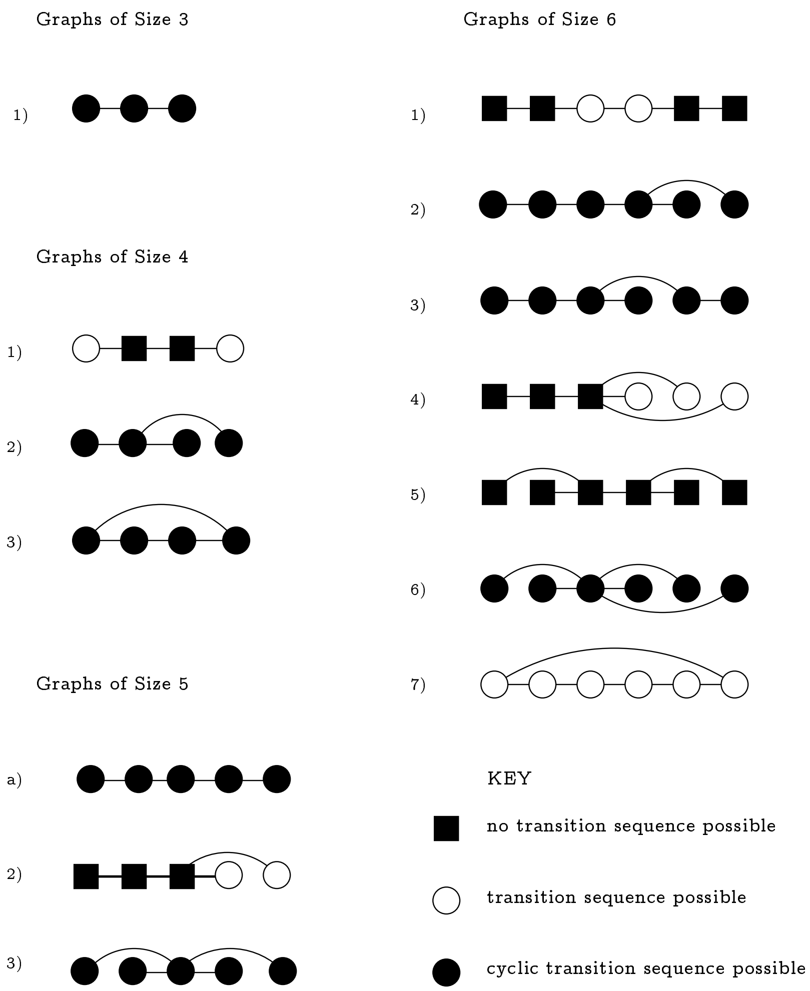

Figure 3: Minimal subgraphs of order 3 to 6. Every connected graph contains one of these as a spanning subgraph.

3 Directed G-codes

Let G be a directed graph. We make a make here a few small observations about directed G-codes All the definitions in the Introduction carry over to directed G-codes.

Lemma 3.1 *If there is a cyclic* G*-code, then* G *is strongly connected.*

Proof: Choose any 2 vertices, u and v in G. Let $H = \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N$ be a cyclic G-code such that $u = \tau(\mathbf{b}_1, \mathbf{b}_2)$. Since the bit in position v must change, there must be a path from u to v in the graph of transitions, which is a subgraph of G. □

That a strongly connected graph is not sufficient to produce a G-code is demonstrated by considering any directed cycle of n vertices. This graph forces the transition sequence to repeat the sequence after $2n$ which is less than 2^n for $n \geq 3$.

Question 3.1 *Does a cyclic transition sequence require that its graph of transitions have at least one bi-directional edge?*

4 Acknowledgement

We thank Stephen Chen of the Department of Computer Science of Griffiths University, Australia, for motivating this paper by posting the essence of Question 2.3 to the newsgroup `sci.math.research`. We also thank Luis Goddyn, Wendy Myrvold, and Glenn Hurlbert for helpful comments on an earlier draft of this paper.

References

[1] J. Bitner, G. Ehrlich, and E. Reingold, "Efficient Generation of the Binary Reflected Gray Code and its Applications," *Communications of the ACM*, 19 (1976) 517-521.

[2] E.N. Gilbert, *Gray Codes and paths on the* n*-cube*, Bell Systems Technical Journal, 37 (1958) 815-826.

[3] L. Goddyn, G.M. Lawrence, and E. Nemeth, *Gray Codes with optimized run lengths*, Utilitas Mathematica, 34 (1988) 179-192.

[4] F. Thomson Leighton, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes*, Morgan Kaufmann, 1992.

[5] A. Nijenhuis and H. S. Wilf, *Combinatorial Algorithms*, Second Edition, Academic Press, 1978.

[6] Reingold, Nievergelt, and Deo, *Combinatorial Algorithms: Theory and Practice*, Prentice Hall, 1977.

[7] Carla D. Savage and Peter Winkler, *Monotone Gray Codes and the Middle Levels Problems*, J. Combinatorial Theory, Series A, 70 (1995) 230-248.

[8] H.S. Wilf, *Combinatorial Algorithms: An Update*, SIAM, CBMS 55.