ON EMBEDDING GRAPHS IN SQUASHED CUBES

R. L. Graham and H. O. Pollak
Bell Laboratories
Murray Hill, NJ 07974

1. Introduction. For the set of three symbols $S = \{0,1,*\}$, define the function d from $S \times S$ to the nonnegative integers $\mathbb{N}$ by

$$d(s,s') = \begin{cases} 1 & \text{if } \{s,s'\} = \{0,1\}, \\ 0 & \text{otherwise.} \end{cases}$$

For $n \in \mathbb{N}$, d can be extended to a mapping of $S^n \times S^n$ to $\mathbb{N}$ by

$$d((s_1,\ldots,s_n),\ (s_1',\ldots,s_n')) = \sum_{k=1}^{n} d(s_k,s_k').$$

We shall refer to $d((s_1,\ldots,s_n),\ (s_1',\ldots,s_n'))$ as the distance between the two n-tuples $(s_1,\ldots,s_n)$ and $(s_1',\ldots,s_n')$ although, strictly speaking, this is an abuse of terminology since d does not satisfy the triangle inequality.

For a connected graph G, the distance between two vertices v and v' in G, denoted by $d_G(v,v')$, is defined to be the minimum number of edges in any path between v and v'.

The following problem arose recently in connection with a data transmission scheme of J. R. Pierce [4].

Given a connected graph G, find the least integer $N(G)$ for which it is possible to associate, with each vertex v of G, an element $A(v) \in S^{N(G)}$, such that

$$d_G(v,v') = d(A(v),\ A(v')) \tag{1}$$

for all pairs of vertices v and v' in G.

The mapping A will be called an addressing of G; $A(v)$ will be called the address of the vertex v. Of course, it is not a priori clear that addressings exist for all connected graphs G. It will be seen that an addressing of G is equivalent to a distance-preserving embedding of G into the 1-skeleton of an n-dimensional cube in which certain faces have been "squashed".

In the following sections, various bounds on $N(G)$ are established. In addition, $N(G)$ is determined exactly for a number of classes of graphs.

2. Squashed cubes. Let T_n denote the set of 2^n points $\{(\varepsilon_1,\ldots,\varepsilon_n): \varepsilon_k = 0$ or $1\}$ in E^n. Let Q_n denote the graph which has T_n as its set of vertices and an edge between the vertices $(\varepsilon_1,\ldots,\varepsilon_n)$ and $(\varepsilon_1',\ldots,\varepsilon_n')$ iff they differ in exactly one coordinate. Thus, Q_n is just the 1-skeleton of an n-cube.

For a given n-tuple $\overline{s} = (s_1,\ldots,s_n) \in S^n = \{0,1,*\}^n$, associate with $\overline{s}$ the set s^* of vertices of Q_n which can be obtained by replacing all s_k which are $*$'s by either 0 or 1. Thus, if $\overline{s}$ has r $*$'s then s^* has 2^r elements. If one replaces the vertices of Q_n which belong to s^* by a single vertex labeled $\overline{s}$ and an edge is placed between $\overline{s}$ and $(\varepsilon_1,\ldots,\varepsilon_n)$ if and only if some element of s^* and $(\varepsilon_1,\ldots,\varepsilon_n)$ are adjacent in Q_n, one forms a new graph Q_n'. One may think of Q_n' as the 1-skeleton of an n-cube in which a certain r-dimensional face was "squashed" and the 2^r vertices were identified by a single vertex.

More generally, if $\overline{s}_1, \ldots, \overline{s}_t$ all belong to S^n and $d(\overline{s}_i,\overline{s}_j) \geq 1$ for $i \neq j$, then one may form the graph Q_n^* by identifying each of the sets of vertices s_k^* by the corresponding single vertex $\overline{s}_k$ with edges incident to $\overline{s}_k$ as previously indicated. Q_n^* may be thought of as the 1-skeleton of an n-cube in which t disjoint hyperfaces have been squashed to points.

An addressing A of a graph G using elements of S^n can now be seen to be equivalent to the existence of a squashed n-cube Q_n^* which is isomorphic to G. The vertex v_k of G corresponds to the vertex $A(v_k)$ in Q_n^* so that A is a distance-preserving map of G onto Q_n^*. $N(G)$ is the least n for which this is possible.

For example, an addressing of K_4, the complete graph on 4 vertices, is given in Figure 1. The associated squashed 3-cube is also shown.

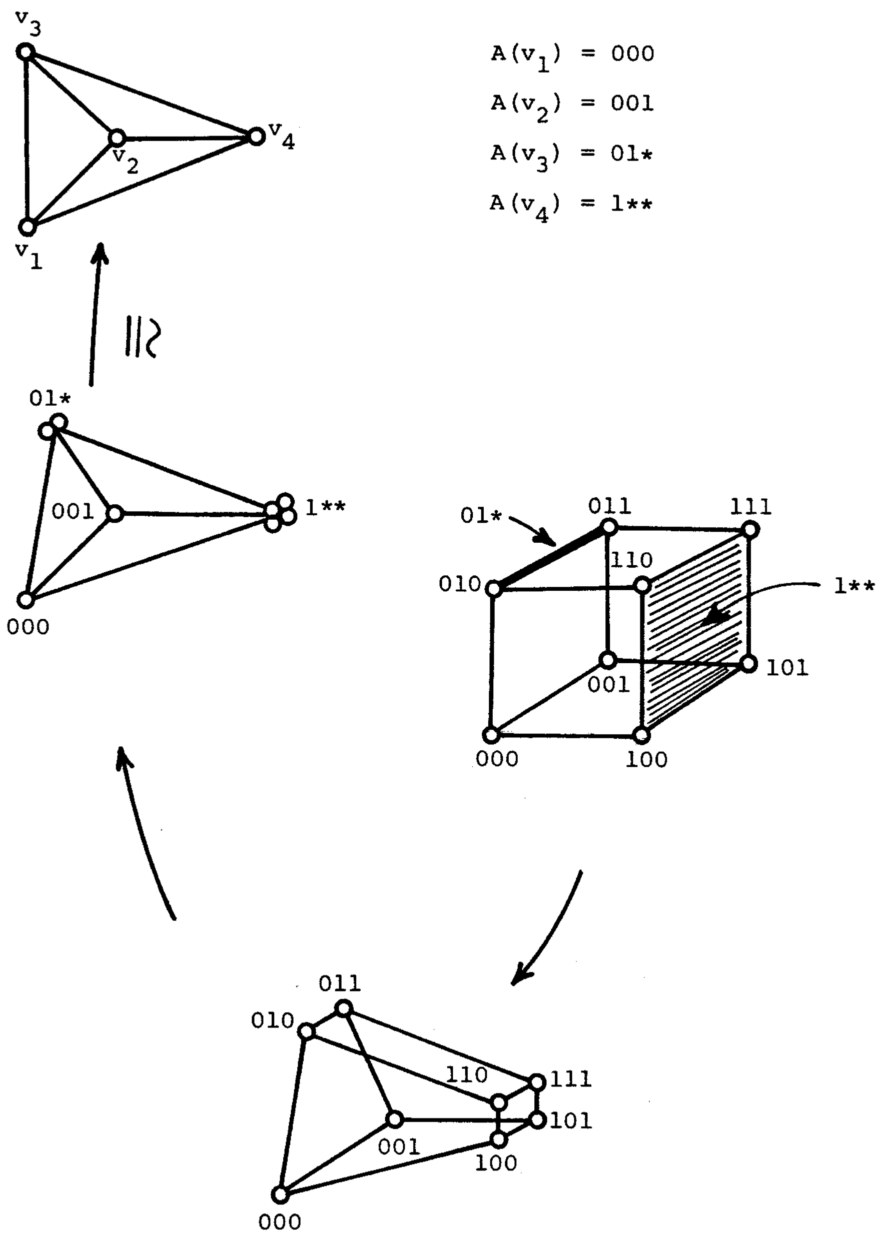
v3
v2
v4
v1
A(v1) = 000
A(v2) = 001
A(v3) = 01*
A(v4) = 1**
01*
001
1**
000
01*
011
111
110
010
1**
001
101
000
100
011
010
110
111
001
101
100
000

Figure 1.

3. <u>Upper bounds on</u> $N(G)$. Given an arbitrary finite connected graph G, we first show $N(G) < \infty$. To see this, let

$$N = \sum_{v_i, v_j} d_G(v_i, v_j)$$

where the sum is over all pairs of vertices v_i, v_j of G. Construct the addresses $A(v_k) \in S^N$ as shown below.

$$\begin{array}{lcccc}
 & \overbrace{\qquad\qquad}^{d_G(v_1,v_2)} & \overbrace{\qquad\qquad}^{d_G(v_1,v_3)} & & \overbrace{\qquad\qquad}^{d_G(v_1,v_j)} \\
A(v_1) = & 0 \;\ldots\ldots\; 0 & 0 \;\ldots\ldots\; 0 & \ldots & * \;\ldots\ldots\; * \\
A(v_2) = & 1 \;\ldots\ldots\; 1 & * \;\ldots\ldots\; * & \ldots & * \;\ldots\ldots\; * \\
A(v_3) = & * \;\ldots\ldots\; * & 1 \;\ldots\ldots\; 1 & \ldots & * \;\ldots\ldots\; * \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
A(v_i) = & * \;\ldots\ldots\; * & * \;\ldots\ldots\; * & \ldots & 0 \;\ldots\ldots\; 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
A(v_j) = & * \;\ldots\ldots\; * & * \;\ldots\ldots\; * & \ldots & 1 \;\ldots\ldots\; 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots
\end{array}$$

For each pair of vertices v_i, v_j, a unique block of $d_G(v_i, v_j)$ coordinate positions is used to achieve $d(A(v_i), A(v_j)) = d_G(v_i, v_j)$ by placing a block of 0's and a block of 1's in these coordinate positions in v_i and v_j and blocks of $*$'s in these coordinate positions for all other v_k. This argument shows

(2) $$N(G) \le \sum_{v_i, v_j} d_G(v_i, v_j) .$$

If G has n vertices and m_G denotes $\max d_G(v_i, v_j)$, where the max is taken over all v_i, v_j in G, then a slight modification of the preceding argument can be used to show

$$N(G) \le m_G(n-1). \qquad (3)$$

No example of a graph G is currently known for which $N(G) > n - 1$. However, it has not yet even been shown that there exists a fixed constant c such that $N(G) \le cn$ for all connected graphs G with n vertices.

4. Lower bounds on $N(G)$. Let $D(G)$ denote the $n \times n$ matrix $(d_{i,j})$, where $d_{i,j} = d_G(v_i,v_j)$, $1 \le i,j \le n$. Let A be an addressing of G using elements of S^N, and write

$$A(v_k) = (s_{k,1},\ldots,s_{k,N}), \quad 1 \le k \le n.$$

We consider the contributions to the various interpoint distances made by a fixed coordinate position in the addresses, say, the m^{th} coordinate position. An important fact to note is that if $s_{i,m} = 0$ and $s_{j,m} = 1$ then these components contribute 1 to the distance $d(A(v_i),A(v_j))$. Of course, if $s_{i,m} = 1$ and $s_{j,m} = 0$ then these components also contribute 1 to $d(A(v_i),A(v_j))$. In all other cases, these components contribute zero to $d(A(v_i,A(v_j))$. Thus, if $C_m(s)$, $s \in S$, denotes the set of t for which $s_{t,m} = s$ then the m^{th} coordinate position contributes 1 to $d(A(v_i),A(v_j))$ iff either $i \in C_m(0)$, $j \in C_m(1)$ or $j \in C_m(0)$, $i \in C_m(1)$.

This remark can be restated in the following terms. Let $Q(G)$ denote the quadratic form defined by

$$Q(G) = \sum_{1 \le i,j \le n} d_{i,j}x_ix_j ,$$

and let

$$Q'(G) = \sum_{1 \le i < j \le n} d_{i,j}x_ix_j = (1/2)Q(G).$$

Then

$$Q'(G) = \sum_{m=1}^{N} \Big(\sum_{i \in C_m(0)} x_i \Big)\Big(\sum_{j \in C_m(1)} x_j \Big) . \tag{4}$$

Hence, the existence of an addressing for G using elements of S^N is equivalent to a decomposition of $Q'(G)$ of the type given by (4). A simple algebraic transformation allows (4) to be rewritten as

$$Q'(G) = \frac{1}{4} \sum_{m=1}^{N} \Big[\Big(\sum_{i \in C_m(0)} x_i + \sum_{j \in C_m(1)} x_j \Big)^2 - \Big(\sum_{i \in C_m(0)} x_i - \sum_{j \in C_m(1)} x_j \Big)^2 \Big] . \tag{4'}$$

This shows that $Q'(G)$ is congruent to a form which has at most N positive squares and at most N negative squares. However, results from matrix theory [3] allow us to conclude from this that

$N \geq$ index $Q'(G) = n_+ =$ number of positive eigenvalues of $D(G)$

and

$N \geq$ index $Q'(G)$ - rank $Q'(G) = n_- =$ number of negative eigenvalues of $D(G)$.

We summarize this in the following theorem.*

THEOREM. <u>A lower bound for</u> $N(G)$ <u>is given by</u>:

$$N(G) \geq \max\{n_+, n_-\} . \tag{5}$$

Since the sum of the eigenvalues of $D(G)$ equals the trace of $D(G)$, which is 0, then $\max\{n_+, n_-\} \leq n-1$. Hence, this theorem can never be used to find a counterexample to the inequality $N(G) \leq n-1$.

---

* First established in a somewhat different way by H.S. Witsenhausen.

It should be noted that if $\{s_1,\ldots,s_{2^n+1}\} \subseteq S^n$, then for some $i \neq j$, $d(s_i,s_j) = 0$. This implies the bound

$$N(G) \geq \log_2 n \tag{6}$$

for a graph G with n vertices.

5. Some special cases. We shall apply (5) to determine $N(G)$ for a number of classes of graphs.

(i). $G = K_n$ - the complete graph on n vertices. In this case $d_{i,j} = 1$ for all $i \neq j$ and $n_+ = 1$, $n_- = n-1$. By (5) this implies $N(G) \geq n-1$. However, it is easy to see that $N(G) \leq n-1$ by considering the decomposition of $Q'(G)$ given by

$$Q'(G) = \sum_{1 \leq i<j \leq n} x_i x_j = \sum_{i=1}^{n-1} \left(x_i \sum_{j=i+1}^{n} x_j\right).$$

(This decomposition corresponds to squashing one hyperface of each dimension in the (n-1)-cube.) The two inequalities imply

$$N(K_n) = n - 1. \tag{7}$$

Equation (7) has an interesting graph-theoretic interpretation obtained by associating complete bipartite subgraphs of G with terms of the decomposition of $Q'(G)$ given in (4).

COROLLARY. *If K_n is decomposed into t edge-disjoint complete bipartite subgraphs, then $t \geq n - 1$.*

No proof of this fact is known which does not use an eigenvalue argument.

(ii). $G = T_n$ - a tree on n vertices. Suppose the vertices of T_n are labeled $v_1, v_2, \ldots, v_n$ so that for $1 \leq k \leq n$, the subgraph of T_n determined by the vertices $v_1, \ldots, v_k$ is a subtree of T_n. In [2], it is shown that it is possible to transform the matrix $D(T_n)$ using elementary row and column operations to a matrix of the form

$$D^*(T_n) = \begin{bmatrix} 0 & 1 & 1 & 1 & \cdots\cdots & 1 & 1 \\ 1 & -2 & 0 & 0 & \cdots\cdots & 0 & 0 \\ 1 & 0 & -2 & 0 & \cdots\cdots & 0 & 0 \\ 1 & 0 & 0 & -2 & \cdots\cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ 1 & 0 & 0 & 0 & \cdots\cdots & -2 & 0 \\ 1 & 0 & 0 & 0 & \cdots\cdots & 0 & -2 \end{bmatrix}.$$

Since $D^*(T_n)$ and $D(T_n)$ have the same determinant, then this implies

(8) $$\det D(T_n) = (-1)^{n-1}(n-1)2^{n-2}, \quad n \geq 1,$$

independent of the structure of T_n.

By the way T_n was labeled, the upper left k^{th} order principal submatrices $D_k(T_n)$ of $D(T_n)$ are also distance matrices of trees and, hence, $\det D_k(T_n) = (-1)^{k-1}(k-1)2^{k-2}$, $1 \leq k \leq n$. However, a theorem from linear algebra [3] asserts that, in this case, the number of permanences in sign* of the sequence

$$1, \det D_1(T_n), \det D_2(T_n), \ldots, \det D_n(T_n)$$

is just equal to n_+, the number of positive eigenvalues of $D(T_n)$. But there is just one permanence in sign in the above sequence, so that $n_+ = 1$. Since, for $n > 1$, $\det D(T_n) \neq 0$ then $D(T_n)$ is nonsingular and $n_- = n-1$. Therefore by (5), $N(T_n) \geq n-1$.

It is easy to show that $N(T_n) \leq n-1$ by inductively addressing the v_k with increasing k, using the fact that if v_i is an exterior vertex of a tree T (i.e., v_i has degree 1) and v_i is adjacent to v_j in T then $d_T(v_i,v_k) = 1 + d(v_j,v_k)$ for all $k \neq i$.

Thus, it follows that

(9) $$N(T_n) = n - 1.$$

In fact, any addressing of T_n with elements of S^{n-1} can use no d's.

---

* Where the sign of 0 may be fixed arbitrarily.

(iii). $G = C_n$ - a cycle on n vertices. Again, (5) can be used directly to show

$$N(C_n) \geq \begin{cases} n-1 & \text{if } n \text{ is odd,} \\ \frac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

It is not difficult to construct addressings (cf. [2]) which achieve these bounds so that we have

$$N(C_n) = \begin{cases} n-1 & \text{if } n \text{ is odd} \\ \frac{n}{2} & \text{if } n \text{ is even.} \end{cases} \qquad (10)$$

(iv). $G = Q_n$ - the 1-skeleton of an n-cube. The labeling of Q_n described previously produces an addressing of Q_n using n-tuples of 0's and 1's. On the other hand, by (6), $N(Q_n) \geq \log_2 |Q_n| = n$. This implies $N(Q_n) = n$ (which is not surprising).

(v). $G = K_{n_1,n_2}$ - the complete bipartite graph on n_1 and n_2 vertices. Rearrange the rows and columns of $D(K_{n_1,n_2})$ so that it has the form

$$D(K_{n_1,n_2}) = \left[\begin{array}{cccc|cccc} \multicolumn{4}{c|}{\overbrace{\hphantom{0\ 2\ \cdots\ 2}}^{n_1}} & \multicolumn{4}{c}{\overbrace{\hphantom{1\ 1\ \cdots\ 1}}^{n_2}} \\ 0 & 2 & \cdots & 2 & 1 & 1 & \cdots & 1 \\ 2 & 0 & \cdots & 2 & 1 & 1 & \cdots & 1 \\ \vdots & & & \vdots & \vdots & \vdots & & \vdots \\ 2 & 2 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ \hline 1 & 1 & \cdots & 1 & 0 & 2 & \cdots & 2 \\ 1 & 1 & \cdots & 1 & 2 & 0 & \cdots & 2 \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ 1 & 1 & \cdots & 1 & 2 & 2 & \cdots & 2 \end{array}\right].$$

A straightforward induction argument shows that

(11) $$\det D(K_{n_1,n_2}) = (-1)^{n_1+n_2} 2^{n_1+n_2-2}(3n_1n_2-4n_1-4n_2+4)$$

for $n_1,n_2 \geq 1$. By the result mentioned in (ii) on the signs of the determinants of the principal submatrices of $D(G)$, and the fact

$$\det D(K_n) = (-1)^{n-1}(n-1)2^n,$$

it follows that for $D(K_{n_1,n_2})$,

$$n_- = \begin{cases} n_1 + n_2 - 2 & \text{if } n_1 \geq 2,\ n_2 \geq 2, \\ n_1 + n_2 - 1 & \text{otherwise.} \end{cases}$$

Thus*, by (5),

$$N(K_{n_1,n_2}) \geq \begin{cases} n_1 + n_2 - 2 & \text{if } n_1 \geq 2,\ n_2 \geq 2, \\ n_1 + n_2 - 1 & \text{otherwise.} \end{cases}$$

In the other direction, the following general result applies.

THEOREM. <u>Suppose</u> G <u>is a graph on</u> n <u>vertices such that for some edge</u> $\{v_i,v_j\}$,

(12) $$\min\{d_G(v_i,v_k),\ d_G(v_j,v_k)\} \leq 1$$

<u>for all vertices</u> v_k <u>of</u> G. <u>Then</u> $N(G) \leq n - 1$.

<u>Proof</u>. Assume without loss of generality that (12) holds for $i = 1$, $j = 2$. The quadratic form $Q'(G)$ has the following decomposition:

---

* $K_{2,2}$ is the only K_{n_1,n_2} for which $D(K_{n_1,n_2})$ is singular.

$$Q'(G) = \sum_{1\le i<j\le n} d_{ij}x_ix_j$$

$$= (x_1 + \sum_{d_{2i}=2} x_i)(x_2 + \sum_{d_{1j}=2} x_j) + x_3(\ldots) + x_4(\ldots) + \ldots + x_n(\ldots),$$

where it is not difficult to check that the appropriate choices can be made in the parenthetical expressions.

Since the complete bipartite graph K_{n_1,n_2} satisfies the hypothesis of the theorem then

$$N(K_{n_1,n_2}) \le n_1 + n_2 - 1.$$

This bound for $N(K_{n_1,n_2})$ has also been obtained in [1]. However, W. T. Trotter has recently shown [5] that

$$N(K_{3n_1+2,3n_2+2}) \le 3n_1 + 3n_2 - 2.$$

We summarize the preceding results on K_{n_1,n_2}.

$$N(K_{n_1,n_2}) = \begin{cases} n_1 + n_2 - 1 & \text{for } 1 = n_1 \le n_2 \\ 2 & \text{for } n_1 = n_2 = 2 \\ n_1 + n_2 - 2 & \text{for } n_1 \equiv n_2 \equiv 2 \pmod 3. \end{cases}$$

In general,

$$n_1 + n_2 - 2 \le N(K_{n_1,n_2}) \le n_1 + n_2 - 1.$$

6. <u>Concluding remarks</u>. Many of the results of the preceding section can also be derived from the recent interesting work of Brandenburg, Gopinath and Kurshan [1]. In particular, they establish the following theorem.

THEOREM. *A graph* G *with* n *vertices can be addressed with elements of* S^N *if and only if there exist binary-valued* $n \times N$ *matrices* P *and* Q *such that* $D(G) = PQ^t + QP^t$.

As previously mentioned, no counterexamples are known to the inequality $N(G) \leq n - 1$. However, this inequality has not even been established for the class of graphs G satisfying $d_G(v_i, v_j) \leq 2$ for all v_i, v_j in G. The example of $K_{2,3}$ in the preceding section shows that the stronger assertion $N(G) = \max\{n_+, n_-\}$ does not always hold.

The fact that $\det D(T_n)$ is independent of the structure of the tree T_n (cf. Eq. (8)) was initially unexpected. It is true, though, that the eigenvalues of $D(T_n)$ do depend on the structure of T_n. It seems likely that as the number of edges in a connected graph G increases the possible range of $\det D(G)$ increases, at least for a while. It would be of interest to know what this range is as a function of the number of edges of G. Perhaps $\det D(G)$ has a simple enumerative interpretation which would make these relationships clear.

References

1. L. H. Brandenburg, B. Gopinath and R. P. Kurshan, On the addressing problem of loop switching, *Bell System Tech. Jour.*, to appear.

2. R. L. Graham and H. O. Pollak, On the addressing problem for loop switching, *Bell System Tech. Jour.* 50 (1971), 2495-2519.

3. B. W. Jones, The arithmetic theory of quadratic forms, *CARUS Mathematical Monograph No.* 10, Math. Assoc. of Amer.,1960, Providence.

4. J. R. Pierce, Network for block switching of data, *Bell System Tech. Jour.*, to appear.

5. W. T. Trotter (personal communication).