Aequationes Math. 55 (1998) 129–145
0001-9054/98/020129-17 $ 1.50+0.20/0

Aequationes Mathematicae

Graph similarity and distance in graphs

Gary Chartrand, Grzegorz Kubicki and Michelle Schultz

Summary. For connected graphs G_1 and G_2 of order n and a one-to-one mapping $\phi\colon V(G_1) \to V(G_2)$, the ϕ-distance between G_1 and G_2 is

$$d_\phi(G_1, G_2) = \sum \Big| d_{G_1}(u,v) - d_{G_2}(\phi u, \phi v) \Big|,$$

where the sum is taken over all $\binom{n}{2}$ unordered pairs u, v of vertices of G_1. The distance between G_1 and G_2 is defined by $d(G_1, G_2) = \min\{d_\phi(G_1, G_2)\}$, where the minimum is taken over all one-to-one mappings ϕ from $V(G_1)$ to $V(G_2)$. It is shown that this distance is a metric on the space of connected graphs of a fixed order. The maximum distance $D(G_1, G_2) = \max\{d_\phi(G_1, G_2)\}$ for connected graphs G_1 and G_2 of the same order is also explored. The total distance of a connected graph G is $\Sigma d(u,v)$, where the sum is taken over all unordered pairs u, v of vertices of G. Bounds for $d(G_1, G_2)$ are presented both in terms of the total distances of G_1 and G_2 and also in terms of the sizes of G_1, G_2, and a greatest common subgraph of G_1 and G_2. For a set $\mathcal{S}$ of connected graphs having fixed order, the distance graph $\mathcal{D}(\mathcal{S})$ of $\mathcal{S}$ is that graph whose vertex set is $\mathcal{S}$ and such that two vertices G_1 and G_2 of $\mathcal{D}(\mathcal{S})$ are adjacent if and only if $d(G_1, G_2) = 1$. Furthermore, a graph G is a distance graph if there exists a set $\mathcal{S}$ of graphs having fixed order such that $\mathcal{D}(\mathcal{S}) \cong \mathcal{G}$. It is shown that every distance graph is bipartite and, moreover, that all even cycles and all forests are distance graphs. Other bipartite graphs are shown to be distance graphs and it is conjectured that all bipartite graphs are distance graphs.

Mathematics Subject Classification (1991). Primary 05C75; secondary 05C12.

1. The distance between two connected graphs

If two graphs G_1 and G_2 are isomorphic, then there exists a one-to-one mapping $\phi\colon V(G_1) \to V(G_2)$ with the property that vertices u and v are adjacent in G_1 if and only if ϕu and ϕv are adjacent in G_2. Of course, ϕ is an isomorphism. In fact, if G_1 and G_2 are connected, then ϕ preserves the distance between *every* pair of vertices of G_1 (not only pairs of adjacent vertices), that is, if u and v are any two vertices of G_1, then $d_{G_1}(u,v) = d_{G_2}(\phi u, \phi v)$.

Let G_1 and G_2 be connected graphs of order n. There are $n!$ one-to-one mappings from $V(G_1)$ to $V(G_2)$. If G_1 and G_2 are isomorphic, then the number of isomorphisms among these $n!$ mappings is the order of the automorphism group

$\operatorname{Aut} G_1$ of G_1. For a one-to-one mapping ϕ and each pair u, v of vertices of G_1, it is of interest to compare $d_{G_1}(u,v)$ with $d_{G_2}(\phi u, \phi v)$. For this reason, we define the *ϕ-distance* between G_1 and G_2 as

$$d_\phi(G_1, G_2) = \sum \bigl|d_{G_1}(u,v) - d_{G_2}(\phi u, \phi v)\bigr|,$$

where the sum is taken over all $\binom{n}{2}$ unordered pairs u, v of vertices of G_1. Of course, if $d_\phi(G_1, G_2) = 0$, then ϕ is an isomorphism and $G_1 \cong G_2$, while if $G_1 \ncong G_2$, then $d_\phi(G_1, G_2) > 0$ for every one-to-one mapping ϕ. This suggests defining the *distance* $d(G_1, G_2)$ between G_1 and G_2 by

$$d(G_1, G_2) = \min\{d_\phi(G_1, G_2)\},$$

where the minimum is taken over all one-to-one mappings ϕ from $V(G_1)$ to $V(G_2)$. Thus, $d(G_1, G_2) = 0$ if and only if $G_1 \cong G_2$. Hence $d(G_1, G_2)$ can be interpreted as a measure of the similarity of G_1 and G_2, where then the smaller the value of $d(G_1, G_2)$, the more similar the structure of G_1 is to that of G_2.

This distance defined on the space of all connected graphs of a fixed order actually produces a metric space. We have noted that $d(G_1, G_2) = 0$ if and only if $G_1 \cong G_2$, and certainly $d(G_1, G_2) = d(G_2, G_1)$ for every two connected graphs G_1, G_2 of the same order. It remains only to verify the triangle inequality. Let G_1, G_2, and G_3 be connected graphs of the same order. Let $\alpha\colon V(G_1) \to V(G_2)$ and $\beta\colon V(G_2) \to V(G_3)$ be one-to-one mappings such that $d_\alpha(G_1, G_2) = d(G_1, G_2)$ and $d_\beta(G_2, G_3) = d(G_2, G_3)$. Then $\beta \circ \alpha\colon V(G_1) \to V(G_3)$ is a one-to-one mapping. Let S be the set of 2-element subsets of $V(G_1)$. Then

$$\begin{aligned}
d(G_1, G_3) &\le d_{\beta\circ\alpha}(G_1, G_3) \\
&= \sum_{\{u,v\}\in S} \bigl|d_{G_1}(u,v) - d_{G_3}((\beta\circ\alpha)u, (\beta\circ\alpha)v)\bigr| \\
&\le \sum_{\{u,v\}\in S} \bigl|d_{G_1}(u,v) - d_{G_2}(\alpha u, \alpha v)\bigr| \\
&\quad + \sum_{\{u,v\}\in S} \bigl|d_{G_2}(\alpha u, \alpha v) - d_{G_3}((\beta\circ\alpha)u, (\beta\circ\alpha)v)\bigr| \\
&= d_\alpha(G_1, G_2) + d_\beta(G_2, G_3) \\
&= d(G_1, G_2) + d(G_2, G_3).
\end{aligned}$$

Therefore, the triangle inequality holds and the distance function d defined on the space of connected graphs of a fixed order is a metric.

Several other measures of similarity of graphs have been considered in the literature. Many of these have involved edge transformations, namely, edge rotations, edge slides, and edge jumps, where the class of graphs involved is that of a fixed order and a fixed size or a subclass thereof. Edge rotations were studied in [4], edge slides in [1], and edge jumps in [3]. Some relationships between the first two of these metrics were described in [5].

Let G_1 be the tree shown in Figure 1 and let G_2 be a cycle of order 5. We compute $d(G_1, G_2)$. Of course, there are $5! = 120$ one-to-one mappings from $V(G_1)$ to $V(G_2)$. However, by taking advantage of the symmetries of G_1 and G_2, we find that there are really only six different mappings to consider. These are also depicted in Figure 1, where for each $i = 1, 2, \ldots, 6$, the mapping $\phi_i \colon V(G_1) \to V(G_2)$ is defined by $\phi_i(v_j) = v'_j$ for $j = 1, 2, \ldots, 5$. Each ϕ_i, then, represents a collection of twenty equivalent mappings.

Now in order to obtain the ϕ_i-distance between G_1 and G_2 for each $i = 1, 2, \ldots, 6$, we determine the distances for each pair of vertices in G_1 and the distances for the corresponding images in G_2. After doing so, we find that

$$\begin{aligned} d_{\phi_1}(G_1, G_2) &= d_{\phi_5}(G_1, G_2) = 7, \\ d_{\phi_2}(G_1, G_2) &= d_{\phi_3}(G_1, G_2) = 5, \\ d_{\phi_4}(G_1, G_2) &= d_{\phi_6}(G_1, G_2) = 9. \end{aligned}$$

Thus $d(G_1, G_2) = 5$.

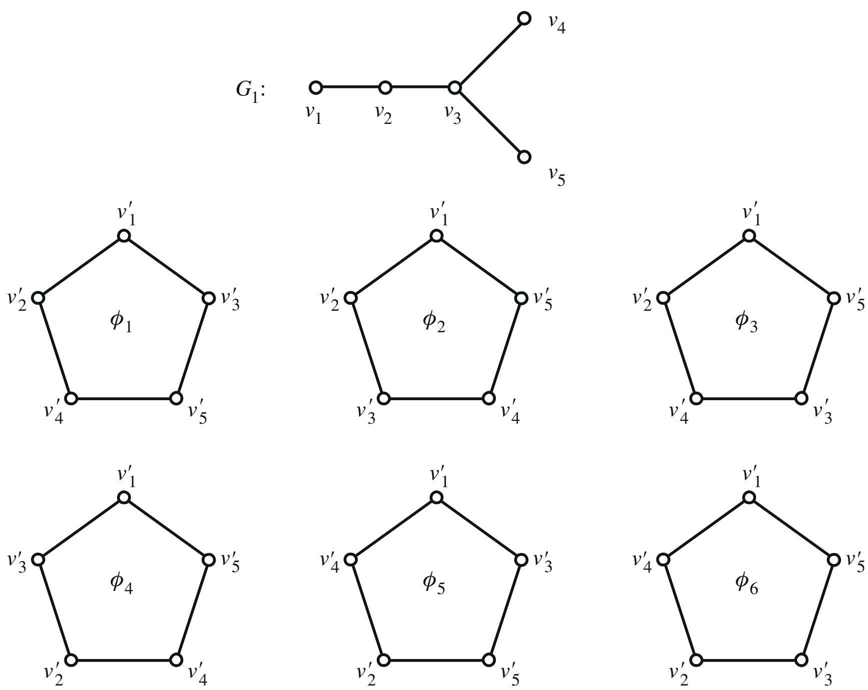

Figure 1

In this example, the ϕ_i-distance between G_1 and G_2 is odd for each i ($1 \leq i \leq 6$). For two connected graphs G_1 and G_2 of the same order and one-to-one mappings α and β from $V(G_1)$ to $V(G_2)$, the numbers $d_\alpha(G_1, G_2)$ and $d_\beta(G_1, G_2)$ are of the same parity, as we show next. For a connected graph G of order n, the *total distance* $\mathrm{td}\,(G)$ of G is defined by

$$\mathrm{td}\,(G) = \sum d(u, v),$$

where the sum is taken over all $\binom{n}{2}$ unordered pairs u, v of distinct vertices of G. This concept has often been studied in the form $\mathrm{td}\,(G) \,/\, \binom{n}{2}$, where it is referred to as the *average distance* of G (see [2, 7], for example). The total distance of a graph has also been studied under the term *Wiener index* after H. Wiener [9] who used this parameter as an aid for determining the boiling point of paraffin. This parameter has been shown to have many other uses in chemistry (see [6]). For additional applications of and motivations for the Wiener index, we refer the reader to [10]. For a connected graph G of order n, the related *Wiener polynomial* (see [4])

$$W(G; x) = \sum x^{d(u,v)}$$

of G, where again the sum is taken over all $\binom{n}{2}$ unordered pairs u, v of distinct vertices of G, has been used to study dendrimers, a class of trees of interest in chemistry. The Wiener polynomial and Wiener index are related by the fact that $W'(G; 1) = \mathrm{td}\,(G)$, that is, the derivative of the Wiener polynomial evaluated at $x = 1$ is the total distance of G.

The parity of the distance between two connected graphs G_1 and G_2 of the same order is determined by $d_\phi(G_1, G_2)$ for any one-to-one mapping $\phi\colon V(G_1) \to V(G_2)$.

Theorem 1. *If G_1 and G_2 are connected graphs of the same order, then*

$$d_\phi(G_1, G_2) \equiv \big[\mathrm{td}\,(G_1) - \mathrm{td}\,(G_2)\big] (\mathrm{mod}\ 2)$$

for every one-to-one mapping $\phi\colon V(G_1) \to V(G_2)$.

Proof. Let $\phi\colon V(G_1) \to V(G_2)$ be a one-to-one mapping. Then

$$\begin{aligned} d_\phi(G_1, G_2) &= \sum \big| d_{G_1}(u, v) - d_{G_2}(\phi u, \phi v) \big| \\ &\equiv \Big[\sum d_{G_1}(u, v) - \sum d_{G_2}(\phi u, \phi v) \Big] (\mathrm{mod}\ 2) \\ &\equiv \Big[\sum d_{G_1}(u, v) - \sum d_{G_2}(x, y) \Big] (\mathrm{mod}\ 2) \\ &\equiv [\mathrm{td}\,(G_1) - \mathrm{td}\,(G_2)]\,(\mathrm{mod}\ 2). \end{aligned}$$

□

Of course, since $d(G_1, G_2)$ is the minimum ϕ-distance over all one-to-one functions $\phi\colon V(G_1) \to V(G_2)$, it follows that $d(G_1, G_2) \equiv [\mathrm{td}\,(G_1) - \mathrm{td}\,(G_2)] (\mathrm{mod}\ 2)$.

A lower bound for the distance between two connected graphs of the same order can also be given in terms of total distances.

Theorem 2. *If G_1 and G_2 are connected graphs of the same order, then*

$$d(G_1, G_2) \geq |\mathrm{td}\,(G_1) - \mathrm{td}\,(G_2)|.$$

Proof. Let $\phi\colon V(G_1) \to V(G_2)$ be a one-to-one mapping. Then

$$\begin{aligned} d_\phi(G_1, G_2) &= \sum \bigl| d_{G_1}(u, v) - d_{G_2}(\phi u, \phi v) \bigr| \\ &\geq \Bigl| \sum (d_{G_1}(u, v) - d_{G_2}(\phi u, \phi v)) \Bigr| \\ &= \bigl| \mathrm{td}\,(G_1) - \mathrm{td}\,(G_2) \bigr|. \end{aligned}$$

Since $d(G_1, G_2)$ is the minimum ϕ-distance among all such one-to-one mappings ϕ, it follows that

$$d(G_1, G_2) \geq \bigl| \mathrm{td}\,(G_1) - \mathrm{td}\,(G_2) \bigr|. \qquad \square$$

Using this lower bound, we now obtain the distance between two graphs where one is a spanning subgraph of the other.

Theorem 3. *If G_1 is a connected spanning subgraph of a graph G_2, then*

$$d(G_1, G_2) = \mathrm{td}\,(G_1) - \mathrm{td}\,(G_2).$$

Proof. Let $\alpha\colon V(G_1) \to V(G_2)$ denote a one-to-one mapping that induces an isomorphism between G_1 and a spanning subgraph of G_2. Since $d_{G_1}(u, v) \geq d_{G_2}(\alpha u, \alpha v)$ for each pair u, v of vertices of G_1, we have

$$\begin{aligned} d_\alpha(G_1, G_2) &= \sum \bigl[d_{G_1}(u, v) - d_{G_2}(\alpha u, \alpha v) \bigr] \\ &= \sum d_{G_1}(u, v) - \sum d_{G_2}(\alpha u, \alpha v) \\ &= \mathrm{td}\,(G_1) - \mathrm{td}\,(G_2). \end{aligned}$$

Since $\mathrm{td}\,(G_1) - \mathrm{td}\,(G_2)$ is a lower bound for $d(G_1, G_2)$ and α is a one-to-one mapping attaining this value, $d(G_1, G_2) = \mathrm{td}\,(G_1) - \mathrm{td}\,(G_2)$. $\square$

Hence the distance between a complete graph and any other connected graph having the same order can be determined.

Corollary 4. *If G is a connected graph of order n, where $n \geq 1$, then*

$$d(G, K_n) = \mathrm{td}\,(G) - \binom{n}{2}.$$

For an integer $n \geq 3$, the total distances of P_n and C_n are

$$\operatorname{td}(P_n) = (n^3 - n)/6$$

and

$$\operatorname{td}(C_n) = \begin{cases} n^3/8 & \text{if } n \text{ is even} \\ (n^3 - n)/8 & \text{if } n \text{ is odd.} \end{cases}$$

Hence by Theorem 3, we have the following

Corollary 5. *For an integer $n \geq 3$,*

$$d(P_n, C_n) = \begin{cases} (n^3 - 4n)/24 & \text{if } n \text{ is even} \\ (n^3 - n)/24 & \text{if } n \text{ is odd.} \end{cases}$$

2. The maximum distance between two connected graphs

We noted in Corollary 4 that for every connected graph G of order n, $d(G, K_n) = \operatorname{td}(G) - \binom{n}{2}$. Indeed, for each one-to-one mapping $\phi\colon V(G) \to V(K_n)$,

$$d_\phi(G, K_n) = \operatorname{td}(G) - \binom{n}{2},$$

that is, d_ϕ is independent of ϕ. Of course, this does not occur in general and so it is natural to define the *maximum distance* $D(G_1, G_2)$ of two connected graphs G_1 and G_2 of the same order by

$$D(G_1, G_2) = \max\{d_\phi(G_1, G_2)\},$$

where the maximum is taken over all one-to-one functions $\phi\colon V(G_1) \to V(G_2)$. It will be shown that K_n is the only graph G_1 of order n for which $d(G_1, G_2) = D(G_1, G_2)$ for *every* connected graph G_2 of order n.

In our next result we turn to stars and determine $d_\phi(K_{1,n-1}, G)$ for each one-to-one mapping $\phi\colon V(K_{1,n-1}) \to V(G)$, where G is a connected graph of order n.

Theorem 6. *Let $G_1 \cong K_{1,n-1}$ with central vertex v and let G_2 be a connected graph of order n and size m. If $\phi\colon V(G_1) \to V(G_2)$ is a one-to-one mapping, then*

$$d_\phi(G_1, G_2) = \operatorname{td}(G_2) - 2\binom{n}{2} + 2m + n - 2r - 1,$$

where $r = \deg \phi v$.

Proof. Let $\phi\colon V(G_1) \to V(G_2)$ be a one-to-one mapping, and let r denote the degree of ϕv, where v is the central vertex of G_1. Let $S = \big\{\{x,y\} \mid xy \notin E(G_2)\big\}$. Since ϕ maps r edges of G_1 into edges of G_2 incident with ϕv and maps the remaining $n-r-1$ edges of G_1 into pairs of nonadjacent vertices of G_2, there is a contribution of

$$\left(\sum_{\{x,y\}\in S}[d(x,y)-2]\right)+n-r-1$$

to $d_\phi(G_1,G_2)$ from the 2-element subsets $\{w,z\}$ of $V(G_1)$ with ϕw and ϕz nonadjacent in G_2. Also there are $m-r$ subsets $\{w,z\}$ of $V(G_1)$ with $\phi w\phi z \in E(G_2)$, and thus

$$\begin{aligned} d_\phi(G_1,G_2) &= \left(\sum_S[d(x,y)-2]\right)+n-r-1+m-r \\ &= \left[\left(\sum_S d(x,y)\right)+m\right]-2\left[\binom{n}{2}-m\right]+n-2r-1 \\ &= \operatorname{td}(G_2)-2\binom{n}{2}+2m+n-2r-1. \end{aligned}$$

□

Thus $D(G_1,G_2)$ (respectively $d(G_1,G_2)$) is attained when ϕ maps the central vertex of G_1 to a vertex of minimum degree (maximum degree) in G_2, and so we obtain the following.

Corollary 7. *Let $G_1 \cong K_{1,n-1}$ and G_2 be a connected graph of order n. Then $D(G_1,G_2)=d(G_1,G_2)$ if and only if G_2 is regular.*

Consider $G_1 \cong K_{1,n-1}+e$, where $e=xy$ is an edge that necessarily joins two end-vertices x and y of $K_{1,n-1}$. Let G_2 be a connected graph of order n and let $\phi\colon V(G_1)\to V(G_2)$ be a one-to-one mapping. Now $d_\phi(G_1,G_2)$ depends not only on the degree of the image of the central vertex of G_1 but also on whether ϕx and ϕy are adjacent. The proof of the next result is similar to the proof of Theorem 6 and hence is omitted.

Theorem 8. *Let $G_1 \cong K_{1,n-1}+e$, where $e=xy$ is an edge joining two end-vertices of $K_{1,n-1}$, and let v denote the central vertex of G_1. Let G_2 be a connected graph of order n and $\phi\colon V(G_1)\to V(G_2)$ a one-to-one mapping. Then*

$$d_\phi(G_1,G_2)=\begin{cases}\operatorname{td}(G_2)-2\binom{n}{2}+2m+n-2r & \text{if } \phi x\phi y\notin E(G_2)\\ \operatorname{td}(G_2)-2\binom{n}{2}+2m+n-2r-2 & \text{if } \phi x\phi y\in E(G_2),\end{cases}$$

where $r=\deg\phi v$.

Thus, if G is any noncomplete connected graph of order n, then, by Theorem 8, $D(K_{1,n-1}+e,G)-d(K_{1,n-1}+e,G)\geq 2$. Hence it follows that K_n is the only graph of order n such that $D(K_n,G)=d(K_n,G)$ for every connected graph G of order n. We restate this below.

Theorem 9. *Let G_1 be a connected graph of order n. Then $D(G_1,G_2)=d(G_1,G_2)$ for every connected graph G_2 of order n if and only if $G_1\cong K_n$.*

We have thus seen graph pairs, namely K_n with any graph and $K_{1,n-1}$ with any regular graph, having equal distance and maximum distance. At the other extreme, two isomorphic graphs can have an arbitrarily large maximum distance.

Theorem 10. *If $G_1\cong G_2\cong K_{n,n}$, where n is even, then $D(G_1,G_2)=n^2$.*

Proof. Let $\phi\colon V(G_1)\to V(G_2)$ be a one-to-one mapping. Let A and B denote the partite sets of G_1 and X and Y the partite set of G_2. Each of these sets is partitioned as follows: $A=A_1\cup A_2$, $B=B_1\cup B_2$, $X=X_1\cup X_2$, and $Y=Y_1\cup Y_2$ such that $\phi(A_1)=X_1$, $\phi(A_2)=Y_2$, $\phi(B_1)=Y_1$, and $\phi(B_2)=X_2$, where $|A_1|=|B_1|=|X_1|=|Y_1|=k$ and $|A_2|=|B_2|=|X_2|=|Y_2|=n-k$. Observe that $\left|d_{G_1}(u,v)-d_{G_2}(\phi u,\phi v)\right|=0$ if and only if $u,v\in A_i\cup B_i$ for $i\in\{1,2\}$ and is 1 otherwise. Thus

$$\begin{aligned} d_\phi(G_1,G_2)=&\sum_{(u,v)\in A_1\times A_2}\left|d_{G_1}(u,v)-d_{G_2}(\phi u,\phi v)\right|\\ &+\sum_{(u,v)\in A_1\times B_2}\left|d_{G_1}(u,v)-d_{G_2}(\phi u,\phi v)\right|\\ &+\sum_{(u,v)\in A_2\times B_1}\left|d_{G_1}(u,v)-d_{G_2}(\phi u,\phi v)\right|\\ &+\sum_{(u,v)\in B_1\times B_2}\left|d_{G_1}(u,v)-d_{G_2}(\phi u,\phi v)\right| \end{aligned}$$

so that

$$d_\phi(G_1,G_2)=1\cdot|A_1\times A_2|+1\cdot|A_1\times B_2|+1\cdot|A_2\times B_1|+1\cdot|B_1\times B_2|=4k(n-k).$$

Hence $d_\phi(G_1,G_2)$ attains its maximum value when $k=n/2$ and so $D(G_1,G_2)=n^2$. □

3. On the distance between graphs with given sizes

First, we present a lower bound for the distance between two connected graphs of the same order in terms of the sizes of the graphs and the size of a greatest common subgraph (a graph of maximum size that is a common subgraph of both graphs).

Theorem 11. *Let G_1 and G_2 be two connected graphs of the same order having sizes m_1 and m_2, respectively, such that the size of a greatest common subgraph is s. Then*

$$d(G_1, G_2) \geq m_1 + m_2 - 2s.$$

Proof. Let $\phi\colon V(G_1) \to V(G_2)$ be a one-to-one mapping. Since the size of a greatest common subgraph of G_1 and G_2 is s, the function ϕ maps at most s edges of G_1 into edges of G_2. Consequently, at least $m_1 - s$ edges of G_1 are mapped into pairs of nonadjacent vertices of G_2. For every such edge uv of G_1, we have $d_{G_1}(u, v) = 1$ while $d_{G_2}(\phi u, \phi v) \geq 2$. Thus

$$\left| d_{G_1}(u, v) - d_{G_2}(\phi u, \phi v) \right| \geq 1.$$

Similarly, there are at least $m_2 - s$ pairs of adjacent vertices of G_2 that are images of pairs of nonadjacent vertices of G_1. For every such edge $\phi x \phi y$ in G_2, it follows that $d_{G_2}(\phi x, \phi y) = 1$ while $d_{G_1}(x, y) \geq 2$. Therefore,

$$\left| d_{G_1}(x, y) - d_{G_2}(\phi x, \phi y) \right| \geq 1.$$

Thus, for every one-to-one mapping $\phi\colon V(G_1) \to V(G_2)$,

$$\begin{aligned} d_\phi(G_1, G_2) &= \sum_{\{u,v\}} \left| d_{G_1}(u, v) - d_{G_2}(\phi u, \phi v) \right| \\ &\geq 1 \cdot (m_1 - s) + 1 \cdot (m_2 - s) = m_1 + m_2 - 2s. \end{aligned}$$

Hence $d(G_1, G_2) \geq m_1 + m_2 - 2s$. □

As immediate consequences of this result, we have the following

Corollary 12. *If G_1 and G_2 are connected nonisomorphic graphs of the same order and same size, then $d(G_1, G_2) \geq 2$.*

Corollary 13. *If G_1 and G_2 are connected graphs of the same order having sizes m_1 and m_2, respectively, such that $d(G_1, G_2) = 1$, then $|m_1 - m_2| = 1$ and one of G_1 and G_2 is a subgraph of the other.*

Although the conditions in Corollary 13 are necessary for the distance between two graphs to be 1, they are not sufficient. For example, let $G_1 = P_4$ and $G_2 = C_4$. Then $G_1 \subseteq G_2$ and $m_2 - m_1 = 1$, but $d(G_1, G_2) = 2$. The obvious necessary and sufficient condition for $d(G_1, G_2) = 1$ is stated next.

Theorem 14. *Let G_1 and G_2 be connected graphs of the same order having sizes m_1 and m_2, respectively, with $m_1 \le m_2$. Then $d(G_1, G_2) = 1$ if and only if $G_1 \subseteq G_2$, $m_2 = m_1 + 1$, and there exists a one-to-one mapping $\phi\colon V(G_1) \to V(G_2)$ such that for some 2-element subset $\{x, y\}$, it follows that $xy \notin E(G_1)$, $\phi x \phi y \in E(G_2)$, $d_{G_1}(x, y) = 2$, and if $\{u, v\} \ne \{x, y\}$, then $d_{G_1}(u, v) = d_{G_2}(\phi u, \phi v)$.*

We now state two immediate corollaries of Theorem 14.

Corollary 15. *Let G_1 and G_2 be connected graphs of the same order having sizes m_1 and m_2, respectively, such that* $\operatorname{diam} G_1 = 2$ *and $m_2 \ge m_1$. Then $d(G_1, G_2) = 1$ if and only if $G_1 \subseteq G_2$ and $m_2 = m_1 + 1$.*

Corollary 16. *Let G_1 and G_2 be connected graphs of the same order having sizes m_1 and m_2, respectively. If $|m_1 - m_2| = k$, then $d(G_1, G_2) \ge k$.*

That the bound for $d(G_1, G_2)$ given in Corollary 16 is sharp is illustrated in Figure 2 for the graphs $G_1 = K_{1,2k}$ and G_2. The function $\phi\colon V(G_1) \to V(G_2)$ defined by $\phi(v_i) = x_i$ $(0 \le i \le 2k)$ has the property that $d_\phi(G_1, G_2) = k$.

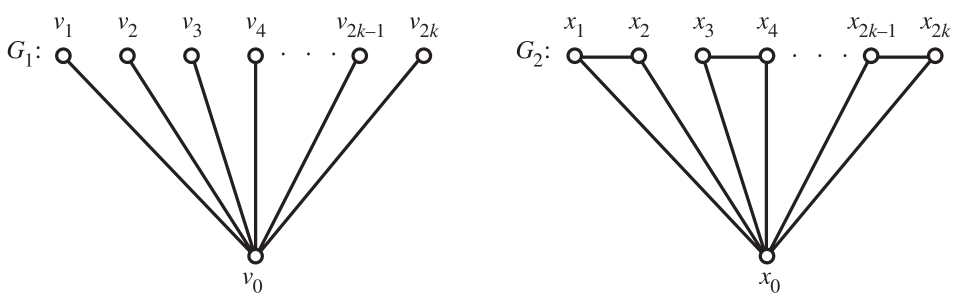

Figure 2

4. Distance graphs

We have already noted that our distance d defined on the space of all connected graphs of a fixed order produces a metric space and now it is desirable to model this metric space by a graph. Precisely, let $\mathcal{S}$ be a set of connected graphs having the same order. Then the *distance graph* $\mathcal{D}(\mathcal{S})$ of $\mathcal{S}$ has vertex set $\mathcal{S}$ and two vertices G_1 and G_2 of $\mathcal{D}(\mathcal{S})$ are adjacent if and only if $d(G_1, G_2) = 1$. Further, we say that a graph G is a distance graph if there exists a set $\mathcal{S}$ of graphs having fixed order such that $\mathcal{D}(\mathcal{S}) \cong G$. The natural problem now is to determine which graphs are distance graphs. The next result narrows our search for distance graphs to the class of bipartite graphs.

Theorem 17. *Every distance graph is bipartite.*

Proof. Let G be a distance graph and let G_1 and G_2 be vertices of G. By Theorem 1, $d(G_1, G_2)$ is odd if and only if $\mathrm{td}\,(G_1)$ and $\mathrm{td}\,(G_2)$ have different parity. In particular, adjacent vertices in G must have total distances of opposite parity. Hence G contains no odd cycles. $\square$

Although no odd cycle is a distance graph, every even cycle is. Let $n \geq 3$ be an integer, and let H be a graph of order $p \leq n$. We use the notation $K_n - H$ to mean $K_n - E(H)$. Before giving the general proof that even cycles are distance graphs, we dispose of two special cases, namely C_4 and C_6. First, consider the graphs F_i $(1 \leq i \leq 4)$ of Figure 3. Let $\mathcal{S} = \{K_5 - F_i \mid 1 \leq i \leq 4\}$. Since the diameter of each graph in $\mathcal{S}$ is 2, Corollary 15 implies that $\mathcal{D}(\mathcal{S}) \cong C_4$. The cycle C_4 is v_1, v_2, v_3, v_4, v_1, where v_i represents the graph $K_5 - F_i$ $(1 \leq i \leq 4)$.

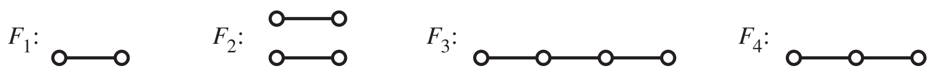

Figure 3

Similarly, the collection $\{K_{10} - H_i \mid 1 \leq i \leq 6\}$ of graphs, where each H_i is shown in Figure 4, has distance graph C_6, where the cycle is v_1, v_2, v_3, v_4, v_5, v_6, v_1 and v_i represents the graph $K_{10} - H_i$ $(1 \leq i \leq 6)$.

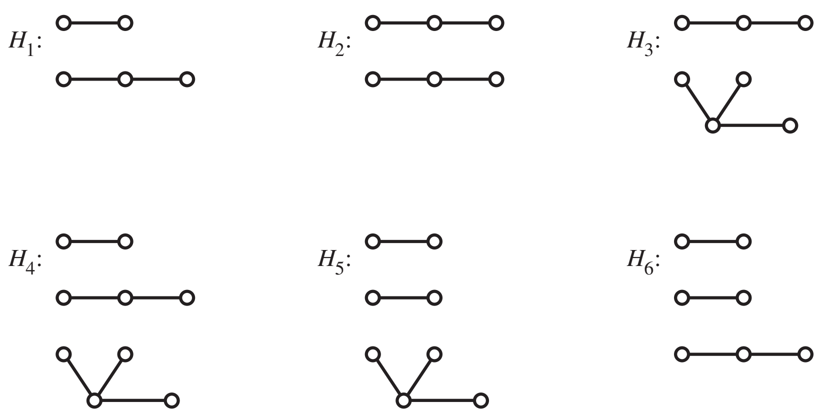

Figure 4

Now we are ready for the general result.

Theorem 18. *Every even cycle is a distance graph.*

Proof. We have already seen that C_4 and C_6 are distance graphs. So let C be a cycle of length $2p \geq 8$. We define a collection $\mathcal{S}$ of $2p$ graphs, each of order $n = p+5$, such that $\mathcal{D}(\mathcal{S}) \cong C$. For each $i = 1, 2, \ldots, p-1$, let $G_i = K_n - K_{1,i}$; let $G_p = K_n - (K_2 \cup K_{1,p-1})$; let $G'_2 = K_n - 2K_2$; and for each $i = 3, 4, \ldots, p+1$, let $G'_i = K_n - (2K_2 \cup K_{1,i-2})$. Now let $\mathcal{S} = \{G_1, G_2, \ldots, G_p, G'_2, G'_3, \ldots, G'_{p+1}\}$. See Figure 5 for $\mathcal{S}$ when $p = 4$. We show that $\mathcal{D}(\mathcal{S}) \cong C$. First note that the size of G_i is $\binom{n}{2} - i$ for $i = 1, 2, \ldots, p$ while the size of G'_i is $\binom{n}{2} - i$ for $i = 2, 3, \ldots, p+1$. Further observe that since $n = p+5$, each graph in $\mathcal{S}$ has diameter 2. Thus, using Corollary 15, we see that

$$G_1, G_2, \ldots, G_p, G'_{p+1}, G'_p, \ldots, G'_3, G'_2, G_1$$

is a cycle of length $2p$ in $\mathcal{D}(\mathcal{S})$. It remains to show that $d(G_i, G'_{i+1}) \geq 2$ and $d(G'_i, G_{i+1}) \geq 2$ for every $i = 2, 3, \ldots, p-2$. Since it is impossible to obtain G'_{i+1} from G_i by removing a single edge from G_i, it follows that G'_{i+1} is not a subgraph of G_i and hence, by Corollary 15, $d(G_i, G'_{i+1}) \geq 2$. Similarly, since it is impossible to obtain G_{i+1} from G'_i by removing a single edge of G'_i, we have $d(G_i, G'_{i+1}) \geq 2$. Thus $\mathcal{D}(\mathcal{S}) \cong C$. $\square$

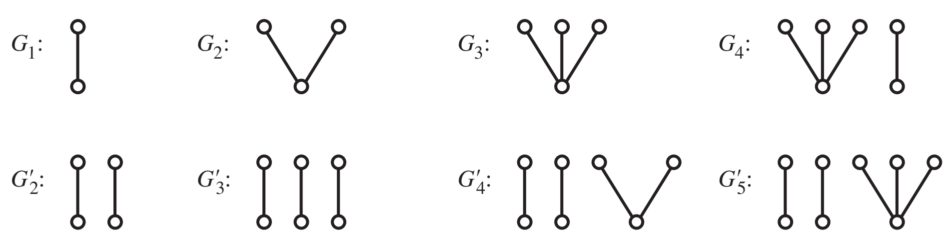

Figure 5

Next, we show that every tree is a distance graph, but first we illustrate the idea of the proof with an example. Consider the tree T shown in Figure 6, where r is the root of T. Perform a Depth-First Search of T, where the vertices are visited in the order $r = v_1, v_2, \ldots, v_6$.

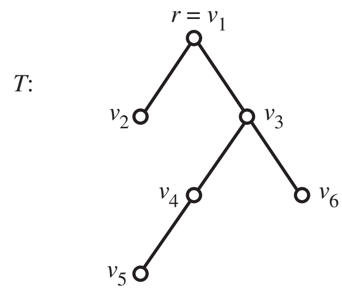

Figure 6

For each vertex v_i $(1 \leq i \leq 6)$, we define a graph G_i. To begin, let $G_1 = K_{1,2,\ldots,6}$ with partite sets $V_1, V_2, \ldots, V_6$, where $|V_i| = i$ for $1 \leq i \leq 6$. Let G_2 be the graph obtained from G_1 by joining the two vertices of the partite set V_2 by an edge. We proceed inductively. Suppose that $G_1, G_2, \ldots, G_k$ $(k < 6)$ have been defined to correspond with v_1, v_2, $\ldots$, v_k, respectively. Then define G_{k+1} as the graph obtained from G_j, where v_j is the parent of v_{k+1}, by joining two vertices in the partite set V_{k+1}. Figure 7 shows the graphs $G_1, G_2, \ldots, G_6$; however, since each of the graphs G_i $(1 \leq i \leq 6)$ contains $K_{1,2,\ldots,6}$ as a subgraph, Figure 7 suppresses these edges and only includes edges added during the construction. Using Corollary 15, we can verify that if $\mathcal{S} = \{G_1, G_2, \ldots, G_6\}$, then $\mathcal{D}(\mathcal{S}) \cong T$.

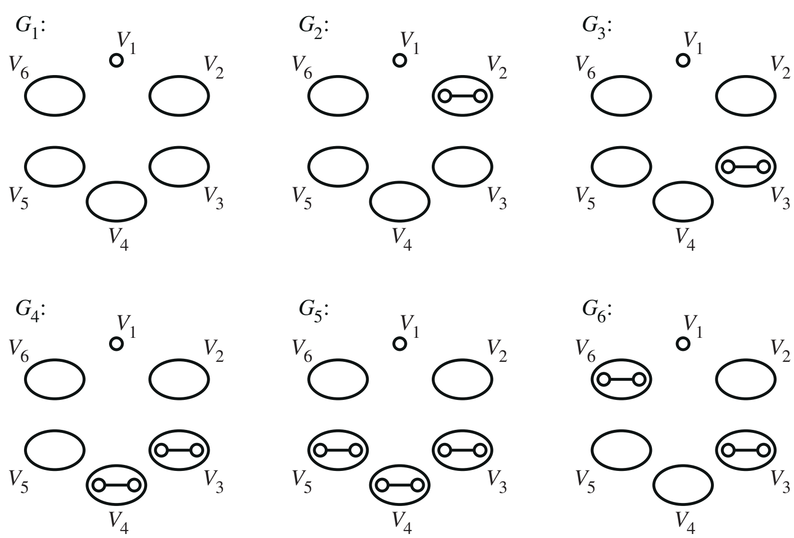

Figure 7

We now describe this procedure for an arbitrary tree.

Theorem 19. *Every tree is a distance graph.*

Proof. Let T be a tree of order p with root r and perform a Depth-First Search of T, where the vertices of T are visited in the order $v_1 = r$, $v_2, v_3, \ldots, v_p$. Let $G_1 = K_{1,2,\ldots,p}$ with partite sets $V_1, V_2, \ldots, V_p$ such that $|V_i| = i$ for $1 \le i \le p$. Next, let G_2 be the graph obtained from G_1 by adding an edge in the partite set V_2. Proceeding inductively, we assume that $G_1, G_2, \ldots, G_k$, where $k < p$, have been defined so that G_i corresponds to v_i for each i $(1 \le i \le k)$. Now if v_j is the parent of v_{k+1}, then G_{k+1} is the graph obtained from G_j by adding an edge in the partite set V_{k+1}.

Let $\mathcal{S} = \{G_1, G_2, \ldots, G_p\}$, and consider $1 \le a, b \le p$ with $a \ne b$. Observe that G_a is a subgraph of G_b if and only if v_a lies on the $v_1 - v_b$ path in T. Indeed, if $G_a \subseteq G_b$, then $d(G_a, G_b) = d_T(v_a, v_b)$. Hence $\phi\colon V(\mathcal{D}(\mathcal{S})) \to V(T)$ defined by $\phi(G_i) = v_i$ induces an isomorphism between a subgraph of $\mathcal{D}(\mathcal{S})$ and T. Now if v_a does not lie on the $v_1 - v_b$ path, then neither G_a nor G_b is a subgraph of the other and so $d(G_a, G_b) \ge 2$. Thus $\mathcal{D}(\mathcal{S}) \cong T$. $\square$

By varying the sizes of the complete multipartite graphs in the proof of the preceding theorem, we can show that several trees are distance graphs simultaneously.

Corollary 20. *Every forest is a distance graph.*

We have seen already that both $K_{2,1}$ and $K_{2,2}$ are distance graphs. Now we show that the graphs $K_{2,n}$ $(n \ge 3)$ are distance graphs as well. It is useful to consider $K_{2,3}$ separately, however. For $1 \le i \le 5$, consider the graphs $K_6 - F_i$, where each F_i is shown in Figure 8. Each graph $K_6 - F_i$ has diameter 2, and so, by Corollary 15, this set of graphs has distance graph isomorphic to $K_{2,3}$.

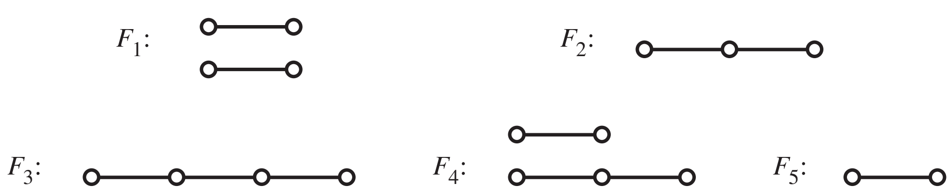

Figure 8

We are now prepared to show that all other graphs $K_{2,n}$ are distance graphs.

Theorem 21. *The graph $K_{2,n}$ is a distance graph for every positive integer n.*

Proof. Since the result has been verified previously for $n = 1, 2, 3$, we assume that $n \geq 4$. Let C be a cycle of length $2n-2$ with vertices labelled C: $v_1, v_2, \ldots, v_{2n-2}, v_1$. Define $G'_1 = C + v_2v_{2n-2}$ and G'_2 as that graph obtained from C by adding a pendant edge incident with v_1. Let $G_i = K_{2n} - G'_i$ for $i = 1, 2$. Since neither G_1 nor G_2 is a subgraph of the other, $d(G_1, G_2) \geq 2$. For each i $(1 \leq i \leq n)$, let F'_i be the graph obtained from G'_1 by adding a pendant edge incident with v_i. See Figure 9 for an illustration of the graphs G'_1, G'_2, and F'_i $(1 \leq i \leq n)$. Define $F_i = K_{2n} - F'_i$ for each i $(1 \leq i \leq n)$. Since F_i is not a subgraph of F_j for every pair i, j of distinct integers with $1 \leq i, j \leq n$, we have $d(F_i, F_j) \geq 2$. On the other hand, $F_i \subseteq G_1$ and $F_i \subseteq G_2$ for every i $(1 \leq i \leq n)$ and so $d(G_1, F_i) = d(G_2, F_i) = 1$ for every i. Thus $\mathcal{D}(\{G_1, G_2, F_1, F_2, \ldots, F_n\}) \cong K_{2,n}$. $\square$

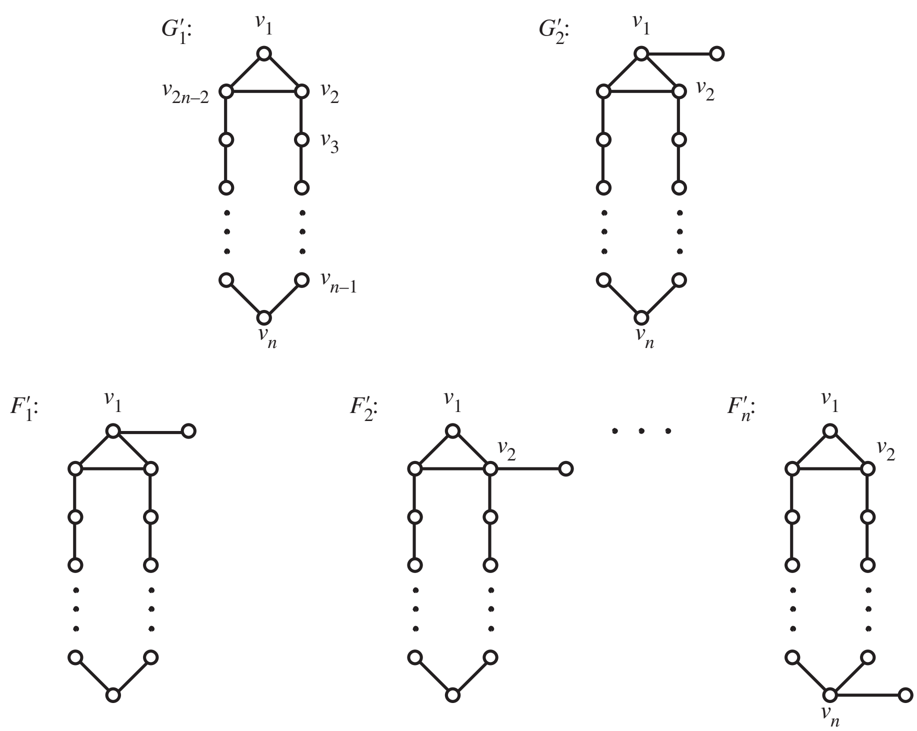

Figure 9

Although we do not know whether all complete bipartite graphs are distance graphs, we do show that $K_{3,3}$ is a distance graph. For each i $(1 \leq i \leq 6)$, let $G_i = K_8 - F_i$, where F_i is given in Figure 10. Then $\operatorname{diam} G_i = 2$ for each i and by Corollary 15, it is easily seen that if $\mathcal{S} = \{G_1, G_2, \ldots, G_6\}$, then $\mathcal{D}(\mathcal{S}) \cong K_{3,3}$, where the partite sets of $\mathcal{D}(\mathcal{S})$ are $\{G_1, G_2, G_3\}$ and $\{G_4, G_5, G_6\}$.

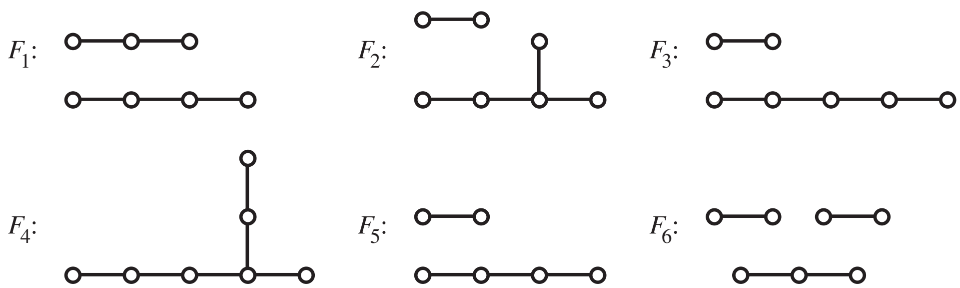

Figure 10

Whether every bipartite graph is a distance graph is not known; however, we conjecture that this, in fact, is the case.

Conjecture. *A graph G is a distance graph if and only if G is bipartite.*

Acknowledgement

The research of the first author was supported in part by Office of Naval Research Grant N00014-91-J-1060.

References

[1] G. Benadé, W. Goddard, T. A. McKee and P. Winter, *On distances between isomorphism classes of graphs*, Math. Bohem. *116* (1991), 160–169.

[2] D. Bienstock and E. Győri, *Average distance in graphs with removed elements*, J. Graph Theory *12* (1988), 375–390.

[3] G. Chartrand, H. Gavlas, H. Hevia and M. A. Johnson, *Rotation and jump distances between graphs*, Discuss. Math. Graph Theory, to appear.

[4] G. Chartrand, F. Saba and H. B. Zou, *Edge rotations and distance between graphs*, Časopis Pro Pěst. Mat. *113* (1985), 87–91.

[5] W. Goddard and H. C. Swart, *Distance between graphs under edge operations*, Discrete Math. *161* (1996), 121–132.

[6] I. Gutman and O. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin, 1986.

[7] G. R. T. Hendry, *Existence of graphs with prescribed mean distance*, J. Graph Theory *10* (1986), 173–175.

[8] B. E. Sagan, Y.-N. Yeh and P. Zhang, *The Wiener polynomial of a graph*, Int. J. Quantum Chem. *60* (1996), 959–969.
[9] H. Wiener, *Structural determination of paraffin boiling points*, J. Amer. Chem. Soc. *69* (1947), 17–20.
[10] B. Y. Yang and Y. N. Yeh, *Zigging and zagging in pentachains*, Adv. in Appl. Math. *16* (1994), 72–94.

Gary Chartrand
Department of Mathematics and Statistics
Western Michigan University
Kalamazoo MI 49008-5152
USA

Grzegorz Kubicki
Department of Mathematics
University of Louisville
Louisville KY 40292
USA

Michelle Schultz
Department of Mathematical Sciences
University of Nevada, Las Vegas
Las Vegas NV 89154-4020
USA

Manuscript received: July 31, 1995 and, in final form, June 18, 1997.