On Distance Preserving and Sequentially Distance Preserving Graphs

Jason P Smith[a,1], Emad Zahedi[b,c]

[a]*Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK.*
[b]*Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027, U.S.A.*
[c]*Department of Computer Science and Engineering, Michigan State University East Lansing, MI 48824, U.S.A.*

Abstract

A graph H is an *isometric* subgraph of G if $d_H(u, v) = d_G(u, v)$, for every pair $u, v \in V(H)$. A graph is *distance preserving* if it has an isometric subgraph of every possible order. A graph is *sequentially distance preserving* if its vertices can be ordered such that deleting the first i vertices results in an isometric subgraph, for all $i \geq 1$. We give an equivalent condition to sequentially distance preserving based upon simplicial orderings. Using this condition, we prove that if a graph does not contain any induced cycles of length 5 or greater, then it is sequentially distance preserving and thus distance preserving. Next we consider the distance preserving property on graphs with a cut vertex. Finally, we define a family of non-distance preserving graphs constructed from cycles.

Keywords: Chordal, Cut Vertex, Distance Preserving, Isometric Subgraph, Sequentially Distance Preserving, Simplicial Vertex.

1. Introduction

The distance between two vertices in a graph plays an important role in many areas of graph theory. Moreover, computing distances in graphs is integral to many real-world applications, especially in areas such as network and optimisation theory. Computing distances between vertices in large graphs is extremely expensive, such as in social networks with millions of vertices. It is often desirable to know the distances between vertices in subgraphs of the original graph. However, this requires recomputing all the distances. One solution to this problem would be to find subgraphs where the distances between all vertices is equal to their distance in the original graph. Such a subgraph is called *isometric*. Isometric subgraphs have been used to study network clustering, see [10, 11].

Email addresses: `jason.p.smith@strath.ac.uk` (Jason P Smith), `Zahediem@msu.edu` (Emad Zahedi)

[1]J.P. Smith was supported by the EPSRC Grant EP/M027147/1

In this framework all graphs are finite, non-empty, simple and connected, unless assumed otherwise. A graph G is *distance preserving*, for which we use the abbreviation dp, if G has an isometric subgraph of every possible order. The notion of distance preserving graphs is a generalisation of *distance-hereditary* graphs, where a graph is distance-hereditary if every connected induced subgraph is isometric. Distance-hereditary graphs where first introduced by Howorka in [7] and have since been studied in various papers, see [1, 3, 6]. Distance-hereditary graphs have many nice properties, for example they are known to be perfect graphs, see [4, 5].

The definition of a distance-preserving graph is similar to the one for distance-hereditary graphs, but is less restrictive. Because of this distance preserving graphs can have a more complex structure than distance-hereditary ones. In fact, it is conjectured in [10] that almost all graphs are distance preserving. So far few results have been proven on distance preserving graphs, although many conjectures exist in the literature, see [8, 10, 12].

One way to show that a graph G is dp is to show that there is an ordering of the vertices $v_1, \ldots, v_n$ of G, such that removing $v_1, \ldots, v_i$ results in an isometric subgraph, for all $i \geq 1$. If such an ordering exists we say that G is *sequentially distance preserving*, which we abbreviate to sdp. Clearly an sdp graph is dp. We say a graph is k-*chordal* if the largest induced cycle is of length k. It was shown in [12] that 3-chordal graphs are sdp. This is proved by using the fact that all 3-chordal graphs have a certain type of ordering of the vertices called a simplicial ordering. This property is generalised to k-chordal graphs in [9]. We apply this generalisation in Section 3 to show that 4-chordal graphs are sdp.

A connected graph has a *cut vertex* x if removing x disconnects the graph. In Section 4 we consider graphs of the form $G \cup H$, where G and H have exactly one common vertex x, so x is a cut vertex. We characterise the dp property in $G \cup H$ in terms of G and H, which reduces the complexity of testing if such graphs are dp.

Finally, in Section 5 we study the class of non-dp graphs. It is conjectured in [10] that almost all graphs are dp, one way to prove this is to give a full classification of the non-dp graphs. By the results in Section 3, we know that a non-dp graph must contain a cycle of length $k > 4$. We investigate how to add vertices to cycle graphs whilst maintaining the non-dp property. To this end, a family of non-dp graphs is defined.

2. Background

In this section we recall some necessary graph theory concepts. For any definitions and notation not given here, and a general overview of graph theory, we refer the reader to [2]. Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. For ease of notation, we let $|G|$ be the number of vertices of G. A *path* in G is a sequence of distinct vertices $v_0, \ldots, v_k$ such that $v_i v_{i+1} \in E(G)$, for all $i = 0, \ldots, k-1$. The length of a path is k, the number of edges. A path P is *chordless* if there is no edge of G between any non-consecutive pair of vertices of P. The *interior* of a path P is obtained by removing the end points v_0, v_k from P. The *distance* between two vertices u, v in G, denoted $d_G(u, v)$, is the minimum length of a path between these vertices. If G is clear from context, we will use $d(u, v)$, instead of $d_G(u, v)$. A path from u to v with length $d_G(u, v)$ is called a u–v *geodesic* path.

An induced subgraph H of G is called an *isometric* subgraph, denoted $H \leq G$, if $d_H(a,b) = d_G(a,b)$, for every pair of vertices $a, b \in V(H)$. We say that G is *distance preserving (dp)* if there is an i-vertex isometric subgraph, for every $1 \leq i \leq |G|$. Given a set $A \subseteq V(G)$, let $G[A]$ be the graph induced on the set A and $G - A := G[V(G) \setminus A]$. We say that G is *sequentially distance preserving (sdp)* if there is an ordering $v_1, \ldots, v_n$ of $V(G)$ such that deleting the first i vertices results in an isometric subgraph for all $i \geq 1$.

The cycle graph C_k is the graph with vertices $v_1, \ldots, v_k$ and the edge set $E(C_k) := \{v_i v_j : |i-j| = 1\} \cup \{v_1 v_k\}$. If G contains C_k as a subgraph we say it contains a cycle of length k or a *k-cycle*. A vertex $v \in V(G)$ is called a *cut vertex* if $G - \{v\}$ is not connected. The graph G^ℓ is the graph whose vertices are those of G and there is an edge between any two vertices $u, v \in G$ with $d_G(u,v) \leq \ell$. The set of vertices adjacent to $v \in V(G)$ is called its *open neighbourhood* and is denoted $\mathcal{N}_G(v)$. The *closed neighbourhood* of v is $\mathcal{N}_G[v] = \mathcal{N}_G(v) \cup v$. A *clique* in G is an induced subgraph that has an edge between every pair of its vertices.

3. All 4-chordal graphs are distance preserving

It was shown in [12] that 3-chordal graphs, often just called chordal graphs, are sequentially distance preserving. This is shown using the well known property that all chordal graphs have a simplicial ordering. This property is generalised to k-chordal graphs in [9], using the notion of a k-simplicial ordering.

Definition 3.1. *A vertex v of a graph G is* weakly k-simplicial *if $\mathcal{N}_G(v)$ induces a clique in $(G-v)^{k-2}$. Furthermore, v is* k-simplicial *if it is weakly k-simplicial and for each non-adjacent pair x, y in $\mathcal{N}_G(v)$, every chordless x, y-path whose interior is entirely in $G - \mathcal{N}_G[v]$ has at most $k-2$ edges. A vertex ordering $v_1, \ldots, v_n$ of G is a* (weakly) k-simplicial ordering *if v_i is (weakly) k-simplicial in $G[v_i, \ldots, v_n]$.*

We use this generalised simplicial ordering to prove the conjecture in [10] that all 4-chordal graphs are distance preserving. In order to do this we need the main result from [9], which we present next. Note that there is a third equivalent statement in the original theorem which we omit here as we do not require it for our results.

Theorem 3.2. *[9, Theorem 1] Consider a graph G and integer $k \geq 3$. The graph G is k-chordal if and only if G has a k-simplicial ordering.*

Before proving the main result of this section we present the following lemma, which is a generalisation of Lemma 3.1 of [12].

Lemma 3.3. *Consider a graph G and vertex $v \in V(G)$. The graph $G - v$ is isometric if and only if v is weakly* 4*-simplicial.*

Proof. Suppose v is weakly 4-simplicial. This implies that $\mathcal{N}_G(v)$ induces a clique in $(G-v)^2$, that is, any pair $x, y \in \mathcal{N}_G(v)$ have a distance of at most 2 in $G - v$. Consider any path P which contains v in its interior. There must be a subpath $x - v - y$ of P, where $x, y \in \mathcal{N}_G(v)$. Because v is 4-simplicial we know that x and y are either neighbours or have a common

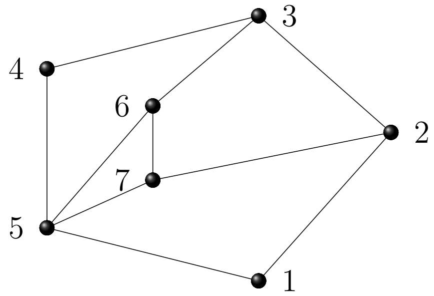

Figure 1: A non-4-chordal graph that is sdp. The vertex labels give an sdp ordering.

neighbour $z \neq v$. Therefore, we can either remove v or replace it with z to get a path that is at least as short as P lying in $G - v$. It follows that $G - v$ is isometric.

Suppose $G - v$ is isometric. Consider any pair $u, w \in \mathcal{N}_G(v)$, then we know that $d_G(u, w) \leq 2$ which implies $d_{G-v}(u, w) \leq 2$. Therefore, $\mathcal{N}_G(v)$ induces a clique in $(G - v)^2$, so v is weakly 4-simplicial. □

The following proposition is an immediately result of Lemma 3.3.

Proposition 3.4. *A graph is sdp if and only if it admits a weakly* 4*-simplicial ordering.*

Proof. Lemma 3.3 implies that a vertex ordering is a weakly 4-simplicial ordering if and only if it is an sdp ordering. □

Now we have all we need to prove Conjecture 5.2 of [10]:

Theorem 3.5. *Any* 4*-chordal graph is sdp, and thus dp.*

Proof. Applying Theorem 3.2 with $k = 4$ shows that for any 4-chordal graph there is a 4-simplicial ordering of the vertices. Moreover, Proposition 3.4 implies this ordering is an sdp ordering. □

The graph in Figure 1 is not 4-chordal, because it contains an induced 5-cycle, so by Theorem 3.2 the graph cannot have a 4-simplicial ordering. However, the ordering given by the vertex labels is a weakly 4-simplicial ordering, so the graph is sdp. To see the ordering is not 4-simplicial, note that the vertex labelled 1 is not 4-simplicial because the path $2 - 3 - 4 - 5$ violates the 4-simplicial condition. Theorem 3.5 implies that a graph that is dp but not sdp cannot contain an induced 4-cycle, combining this with [12, Corollary 3.2] gives the following corollary:

Corollary 3.6. *Any dp graph that is not sdp must contain an induced cycle of length* $k \geq 5$.

4. Separable graphs

A connected graph is said to be *separable* if it can be disconnected by removing a vertex, which we call a *cut vertex*. In this section we consider the distance preserving property in separable graphs. A separable graph can be represented in the following way:

Definition 4.1. *Consider two graphs G and H, with a single common vertex x. Let $G +_x H$ be the union of G and H.*

So $G +_x H$ is a separable graph with a cut vertex x. We characterise the isometric subgraphs of $G +_x H$. To do this we introduce the following lemma.

Lemma 4.2. *Consider a graph $G +_x H$ and two induced subgraphs $H' \subseteq H$, $G' \subseteq G$, with $x \in V(G') \cap V(H')$, then:*

$$G' +_x H' \leq G +_x H \quad \textit{if and only if} \quad H' \leq H \ \textit{and}\ G' \leq G.$$

Proof. First we consider the forward direction. Since x is a cut vertex any geodesic path between a pair of vertices of $H \subseteq G +_x H$ is contained in H, thus $H \leq G +_x H$. The same is true when replacing G and H by G' and H', respectively. Combining this with our assumption we have:

$$d_{H'}(u, v) = d_{G'+_xH'}(u, v) = d_{G+_xH}(u, v) = d_H(u, v),$$

for every pair of vertices $u, v \in V(H')$, so $H' \leq H$. An analogous argument shows that $G' \leq G$.

Now consider the backward direction. Using the fact $H \leq G +_x H$ and the assumption $H' \leq H$, we have

$$d_{G'+_xH'}(u, v) = d_{H'}(u, v) = d_H(u, v) = d_{G+_xH}(u, v), \tag{1}$$

for every pair $(u, v) \in V(H') \times V(H')$. An analogous argument shows that

$$d_{G'+_xH'}(a, b) = d_{G+_xH}(a, b), \tag{2}$$

for every pair $(a, b) \in V(G') \times V(G')$. Next consider a pair $(a, u) \in V(G') \times V(H')$. Any geodesic path from a to u can be considered as the concatenation of an a–x geodesic path in G' and a x–u geodesic path in H'. Applying Equations (1) and (2) implies that:

$$\begin{aligned} d_{G'+_xH'}(a, u) &= d_{G'+_xH'}(a, x) + d_{G'+_xH'}(x, u) \\ &= d_{G'}(a, x) + d_{H'}(x, u) \\ &= d_G(a, x) + d_H(x, u) \\ &= d_{G+_xH}(a, u). \end{aligned}$$

This completes the proof. □

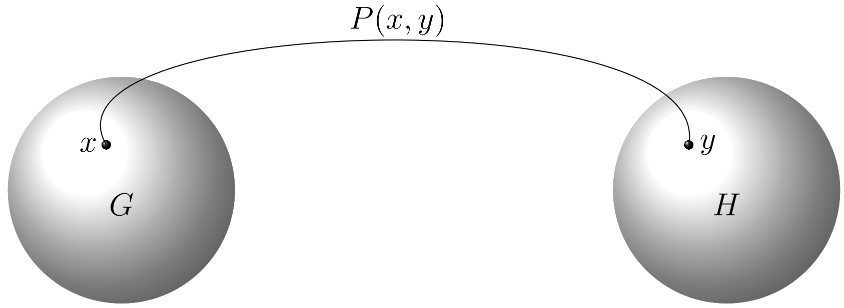

Figure 2: The figure for $G_x \overset{r}{—} H_y$.

To state the main result of this section we use the following definition and notation, which are taken from [8].

Definition 4.3. *For a graph G and sets $X, Y \subseteq V(G)$, let*

$$\mathrm{DP}(G) = \big\{A \subseteq V(G) \ : \ G[A] \leq G\big\} \textit{ and } \mathrm{dp}(G) = \big\{|A| \ : \ A \in \mathrm{DP}(G)\big\}.$$

Also let

$$\begin{aligned}\mathrm{DP}_X^Y(G) &= \big\{A \in \mathrm{DP}(G) : A \cap X = \emptyset, A \cap Y \neq \emptyset\big\} \textit{ and} \\ \mathrm{dp}_X^Y(G) &= \big\{|A| : A \in \mathrm{DP}_X^Y(G)\big\}.\end{aligned}$$

For ease of notation we denote $\mathrm{DP}_{\{v\}}(G)$ by $\mathrm{DP}_v(G)$ and $\mathrm{DP}_X^{\emptyset}$ by DP_X, and similarly for superscripts and dp*.*

Define the set $A + B := \{a + b \ : \ a \in A, \, b \in B\}$, where A and B are sets of integers.

Theorem 4.4. *Consider a graph $G +_x H$. Then:*

$$\mathrm{dp}(G +_x H) \ = \ \big(\mathrm{dp}^x(G) + \mathrm{dp}^x(H) + \{-1\}\big) \ \cup \ \mathrm{dp}_x(G) \ \cup \ \mathrm{dp}_x(H).$$

Proof. We consider two cases based upon whether $A \in \mathrm{DP}(G +_x H)$ contains x. If A does not contain x, then A is fully contained in either G or H, so $\mathrm{dp}_x(G +_x H) = \mathrm{dp}_x(G) \cup \mathrm{dp}_x(H)$. If A does contain x, then Lemma 4.2 implies $A = G' +_x H'$, where $G' \leq G$, $H' \leq H$ and both contain x. Therefore, $\mathrm{dp}^x(G +_x H) = \big(\mathrm{dp}^x(G) + \mathrm{dp}^x(H) + \{-1\}\big)$ where the minus 1 accounts for the common vertex x in G' and H'. Combining these two cases with the formula $\mathrm{dp}(G +_x H) = \mathrm{dp}^x(G +_x H) \cup \mathrm{dp}_x(G +_x H)$ completes the proof. $\square$

Given two disjoint graphs we can connect the graphs by a path of length r, for any $r > 0$.

Definition 4.5. *Consider two disjoint graphs G and H. Let $G_x \overset{r}{—} H_y$ be the graph obtained by connecting $x \in V(G)$ and $y \in V(H)$ with a path $P(x, y)$ of length r.*

These graphs are separable, so applying a simple iteration of Theorem 4.4 gives the following corollary:

Corollary 4.6. *Consider two disjoint graphs G and H. If $r > 0$ then:*

$$\begin{aligned}\mathrm{dp}(G_x \overset{r}{—} H_y) &= \big(\mathrm{dp}^x(G) + \mathrm{dp}^y(H) + \{-1, \ldots, r-1\}\big)\\ &\quad \cup \mathrm{dp}_x(G) \cup \mathrm{dp}_y(H).\end{aligned}$$

5. Maintaining the non-dp property

In this section we investigate the class of non-dp graphs. It is conjectured in [10] that almost all graphs are non-dp. So understanding this class is a logical step towards a full classification of the class of dp graphs. The simplest non-dp graphs are the cycle graphs C_k, for all $k \geq 5$. We investigate how we can add vertices to the cycle graphs and preserve the non-dp property. To this end we introduce the following class of graphs.

Consider the cycle C_k and a set of vertices A, with $|A| = \ell$, such that $A \cap V(C_k) = \emptyset$. For each $a \in A$, select three consecutive vertices of C_k and join a to at least one of the three selected vertices. Let $\mathcal{C}_{k,\ell}$ denote the family of graphs that can be constructed in this way. Given a graph $G \in \mathcal{C}_{k,\ell}$, let $C(G)$ be the original cycle vertices of G and $A(G)$ the added vertices. Note that the addition of the vertices to the cycle graph cannot change the distance between any pair of vertices in $C(G)$, so $C_k \leq G$.

Recall that we label the vertices of C_k as $v_1, \ldots, v_k$, and let $v_{k+1} := v_1$ and $v_0 := v_k$. So there is an edge between two vertices v_i and v_j if and only if $i = j \pm 1$.

Theorem 5.1. *If $k > 2(\ell + 2)$, then any graph in $\mathcal{C}_{k,\ell}$ is non-dp.*

Proof. Consider a graph $G \in \mathcal{C}_{k,\ell}$. If an added vertex a is connected to two cycle vertices c_{i-1} and c_{i+1}, then the removal of either a or c_i results in isomorphic subgraphs. Therefore, when constructing an isometric subgraph of G, by removing a set of vertices of G, we can assume that a is always removed before c_i. Also recall that the added vertices do not alter the distance between any of the cycle vertices. Combining these two points implies that given a graph $H \leq G$ there is a geodesic path in H between any two elements of $C(H)$ that is entirely contained in $H[C(H)]$. Therefore, if $H \leq C_{k,\ell}$, then $H[C(H)] \leq C_k$.

We show that there is no isometric subgraph of G with order $\lfloor \frac{k}{2} \rfloor + 2$. Suppose for a contradiction that such a subgraph does exist, we denote it H. We know that $\ell < \frac{k}{2} - 2$, so to obtain H we must remove a set of s cycle vertices, where $\lceil \frac{k}{2} \rceil - 2 > s > 0$. However, this implies that $C(H)$ has t vertices, where $k > t > \lfloor \frac{k}{2} \rfloor + 2$, and it is straightforward to see that there is no isometric subgraph of C_k with t vertices. Therefore, H is not isometric, so G is non-dp. □

Note that the converse of Theorem 5.1 is not true. For example in Figure 3, the graph G is not dp but $k = 10 \not> 10 = 2(\ell + 2)$. An interesting question, which we leave open, is:

Open Problem 5.2. *How can we add vertices to non-cycle non-dp graphs to get further results on the class of non-dp graphs?*

We end this paper with an interesting conjecture about distance preserving graphs.

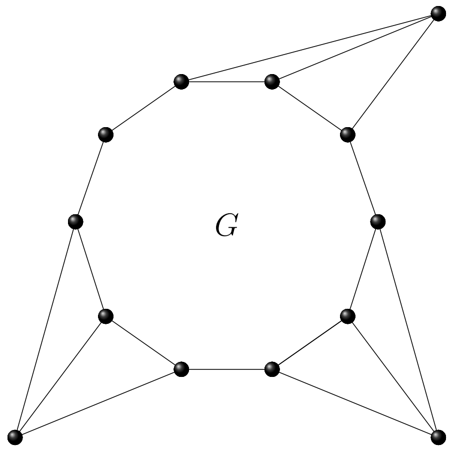

Figure 3: A counterexample to the converse of the Theorem 5.1

Conjecture 5.3. *If G is an n-vertex graph with minimum degree $\delta(G) \geq \frac{n}{2}$, then G is dp.*

Nussbaum and Esfahanian [10] have shown that $\delta(G) \geq \frac{2n}{3} - 1$ forces G to be dp. It is not clear whether this bound is tight or not. If $\delta(G) \geq \lfloor \frac{n}{2} \rfloor$ then G has diameter at most 2. Since the possible distances in such a graph are so limited, one might be able to find the required isometric subgraphs.

References

[1] H.-J. Bandelt and H. M. Mulder. Distance-hereditary graphs. *Journal of Combinatorial Theory, Series B*, 41(2):182–208, 1986.

[2] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*, volume 6. Macmillan London, 1976.

[3] G. Damiand, M. Habib, and C. Paul. A simple paradigm for graph recognition: application to cographs and distance hereditary graphs. *Theoretical Computer Science*, 263(1):99–111, 2001.

[4] A. D'Atri and M. Moscarini. Distance-hereditary graphs, steiner trees, and connected domination. *SIAM Journal on Computing*, 17(3):521–538, 1988.

[5] M. C. Golumbic and U. Rotics. On the clique-width of some perfect graph classes. *International Journal of Foundations of Computer Science*, 11(03):423–443, 2000.

[6] P. L. Hammer and F. Maffray. Completely separable graphs. *Discrete applied mathematics*, 27(1):85–99, 1990.

[7] E. Howorka. A characterization of distance-hereditary graphs. *The quarterly journal of mathematics*, 28(4):417–420, 1977.

[8] M. Khalifeh, B. E. Sagan, and E. Zahedi. Distance preserving graphs and graph products. *arXiv preprint arXiv:1507.04800*, 2015.

[9] R. Krithika, R. Mathew, N. Narayanaswamy, and N. Sadagopan. A Dirac-type characterization of k-chordal graphs. *Discrete Mathematics*, 313(24):2865–2867, 2013.

[10] R. Nussbaum and A.-H. Esfahanian. Preliminary results on distance-preserving graphs. *Congressus Numerantium*, 211:141–149, 2012.

[11] R. Nussbaum, A.-H. Esfahanian, and P.-N. Tan. Clustering social networks using distance-preserving subgraphs. In *The Influence of Technology on Social Network Analysis and Mining*, pages 331–349. Springer, 2013.

[12] E. Zahedi. Distance preserving graphs. *arXiv preprint arXiv:1507.03615*, 2015.