DISCRETE MATHEMATICS 9 (1974) 205–212. © North-Holland Publishing Company

A CHARACTERISATION OF RIGID CIRCUIT GRAPHS

Peter BUNEMAN

School of Artificial Intelligence, Edinburgh University, Scotland, UK

Received 26 February 1973

1. Introduction

An interval graph is the intersection graph of a set of intervals of a line; that is, a graph whose points represent these intervals and whose edges join points which represent intervals with a non-empty intersection. One of the reasons for studying rigid circuit graphs lies in the problem, which has various applications [4, 6, 7], of deciding whether a given graph is an interval graph. It turns out that a necessary but insufficient condition for a graph to be an interval graph is that it should have the rigid circuit property. There is, however, another motivation for studying graphs with this property. In the construction of evolutionary trees and in certain other classificatory problems, it is often desirable to say whether a graph is the intersection graph of a set of subtrees of some tree. Here, the rigid circuit property is more naturally relevant because it is both a necessary and sufficient condition for a graph to be such a tree-intersection graph. The main purpose of this paper is to establish this fact, which will also be shown to provide rather more transparent proofs of some existing theorems on rigid circuit graphs. A short discussion of some of the problems of constructing evolutionary trees is also included.

The graphs we shall deal with will all be undirected and finite with no multiple edges. A *path* in a graph G is a sequence $p_1, p_2, ..., p_n$ of distinct points in G such that $p_i p_{i+1}$ is an edge of G for $1 \leq i \leq n$, and a graph is *connected* iff there is a path between any two points in that graph. A *circuit* is a path whose first and last points are joined by an edge and a *tree* is a connected graph with no circuits. A *subtree* of a tree is a connected subgraph of the tree. A *chord* of the circuit $p_1, p_2, ..., p_n$ is an edge which joins points p_i and p_j, where $|i-j| \neq 1 \pmod n$, and a *rigid circuit* graph is a graph in which every circuit through more than three points has a chord. Trees have, trivially, the rigid circuit property; so do *complete* graphs, those graphs in which every possible edge is present. The removal of any number of points from a graph with the rigid circuit property does not affect that property.

2. Main theorems

Theorem 2.1. *Let* $\{T_1, T_2, ..., T_n\}$ *be a set of subtrees of a tree* T. *Then the intersection graph of* $\{T_1, T_2, ..., T_n\}$ *is a rigid circuit graph.*

Suppose otherwise. Then there is a sequence, $T_1, T_2, ..., T_p$ for convenience, such that the intersection of distinct subtrees T_i and T_j is non-empty iff $|i-j| = 1 \pmod p$, and such that $p > 3$. Working mod p in the obvious fashion, choose a point s_i from $T_i \cap T_{i+1}$. By our suppositions, the s_i are all distinct, so let t_i be the last common point of the paths from s_i to s_{i-1} and s_i to s_{i+1}. These paths lie respectively in T_i and T_{i+1} so that t_i lies in $T_i \cap T_{i+1}$ and the t_i are similarly all distinct. Moreover, by this construction, the concatenation of the paths from t_{i-1} to t_i and t_i to t_{i+1} is also a path. If the trees T_i and T_j do not intersect, then the paths from t_{i-1} to t_i and t_{j-1} to t_j cannot intersect since they respectively lie in these trees. By concatenating all such paths in order we obtain a circuit in T which is counter to the definition of a tree.

The proof of the converse of this theorem can be accomplished in various ways; the strategy adopted here is designed to show something more of the structure of rigid circuit graphs. One reason for doing this is that there is not necessarily a canonical set of subtrees of some tree for representing a rigid circuit graph (consider, for example, the various ways of representing as a set of subtrees the graph whose edges are $a_2 a_1$, $a_3 a_1$ and $a_4 a_1$). Another reason is that in a discussion [2] of the metric properties of trees, a characterisation of trees is used which is closely related to some of the properties of rigid circuit graphs which we shall need in order to prove this converse. Following Dirac [3], we introduce some further definitions. If p_1 and p_2 are points in G, a set C of points in $G - \{p_1, p_2\}$ *separates* p_1 and p_2 if every path from p_1 to p_2 meets C. A set with this property is called a *relative cutset* of G

(relative to the separation of p_1, p_2). If no subset of C separates p_1 and p_2, C is called a *relatively minimal cutset.* We shall use the term "clique" to refer to a maximal complete subgraph of a graph, and in what follows the rigid circuit graph under consideration is taken to be connected, though only minor modifications are needed if it is not so.

Lemma 2.2 (Dirac). *If G is a rigid circuit graph, then any relatively minimal cutset is complete in G* [3].

Suppose C is a relatively minimal cutset separating p_1 and p_2 and suppose that it contains two points q_1 and q_2. There must be a path from p_1 to q_1 which meets C only in q_1, by the minimality of C. Similar paths must exist between q_1 and p_2, q_2 and p_1, and q_2 and p_2. Therefore there are two paths from q_1 to q_2 which, apart from their end points, lie entirely in distinct components of $G-C$. For each component, choose such a path which is shortest and join these paths by their common end points to form a circuit. This circuit must have a chord by the rigid circuit property. This chord cannot link the two components of $G-C$, and since the paths from q_1 to q_2 are both shortest through each component, it must be the chord $q_1 q_2$. Similarly, there is an edge in G between every pair of points in C.

Lemma 2.3. *Any relatively minimal cutset in a rigid circuit graph G is properly contained in at least two distinct cliques in G.*

Suppose C is a minimal cutset relative to the separation of p_1 and p_2. Let G_1 be that component of $G-C$ which contains p_1. Choose a point s of G_1 whose set C_1 of neighbours in C is as large as possible. If $C \neq C_1$, let q be a point of $C-C_1$, and let t be a neighbour of q in G_1 which minimises the length of some path $t, t_1, t_2, ..., s$ in G. Since C is complete, $q_1, q, t, t_1, t_2, ..., s$ is a circuit for any point q_1 of C_1. This circuit has a chord, which must, by construction, be $q_1 t$. t is therefore linked to every point of C_1' and to q, contradicting the maximality of C_1. There is therefore a complete graph with points in G_1 which properly contains C, and a similar complete graph in that component of $G-C$ which contains p_2.

Corollary 2.4. *Each relatively minimal cutset C_i gives rise to an equivalence relation E_i on the cliques of G. Two cliques are related if they have points in the same component of C_i.*

Lemma 2.5. *If E_i and E_j are two such equivalence relations, then at most one equivalence class σ_i of E_i is such that $E_j \mid \sigma_i$ is not the universal relation.*

Suppose otherwise, and suppose that the relatively minimal cutsets C_i and C_j give rise to the equivalence relations E_i and E_j, respectively. Then we can find cliques S_1, S_2, S_3, S_4 such that S_1 and S_2 have points in one component of $G-C_i$ which are separated by the removal of C_j and such that S_3 and S_4 have similar points in another component of $G-C_i$. But this would mean that C_j has points in both components of $G-C_i$ which is impossible since C_j is complete in G.

Lemma 2.6. *If G is a rigid circuit graph and is not itself complete, then at least one of the equivalence relations defined by Corollary* 2.4 *has an equivalence class which contains just one clique.*

Suppose that the class σ_i of E_i contains the cliques S_1 and S_2. There must be two points, one in each of these cliques, which are not adjacent in G. By finding a relatively minimal cutset which separates these points, we can find an equivalence relation E_j with classes σ_j and σ_j' containing S_1 and S_2, respectively. From Lemma 2.5, one of $E_i|\sigma_j$ and $E_i|\sigma_j'$ must be the universal relation so that one of σ_j and σ_j' is properly contained in σ_i. Repeating this we must eventually obtain an equivalence class containing just one clique.

From this we can readily obtain the result noted by Dirac [3] and Fulkerson and Gross [4] that any rigid circuit graph contains at least one point whose neighbours form a complete subgraph of G.

Theorem 2.7 (Converse of Theorem 2.1). *A rigid circuit graph G is the intersection graph of a set of subtrees of a tree T whose points correspona to the cliques in G in such a way that if s_i is the point of T corresponding to the clique S_i in G, then the subtree T_p corresponding to p in G contains s_i if and only if S_i contains p.*

We use induction on the number of cliques in G. If G is complete, the result is trivial. Suppose that this theorem holds for all graphs with fewer than N cliques and suppose G has the rigid circuit property and contains just N cliques. By Lemma 2.6, we can find a relatively minimal cuset of C such that one component of $G-C$ consists just of points in a clique S_1 and which, by Lemma 2.3, is properly contained in S_1 and some other clique S_2, say. Let us use D to denote the points of S_1-C and use the inductive hypothesis to construct a tree T' with the properties of this theorem on the rigid circuit graph $G-D$ which has $N-1$ cliques.

There will be a point s_2 in T' corresponding to the clique S_2. We can form a tree T with N points from T' by attaching a point s_1 to T' by the edge $s_2 s_1$. For each point in C, we extend the corresponding subtree along the edge $s_2 s_1$ and for each point of D we form a new subtree of T which contains just the point s_1. s_1 then corresponds in the correct way to S_1 and the induction is complete.

The construction in Theorem 2.7 does not necessarily produce a unique tree. The reason is that there may be, at any stage in the construction, a clique which is the only member of two or more equivalence classes. This would mean that the point s_1 in Theorem 2.7 could be attached in more than one way to the tree T'. The tree produced by Theorem 2.7 is minimal, for if G is the intersection graph of subtrees of a tree T, then there must be a map from a subset of the nodes of T onto the cliques of G, and the map we have produced is one to one.

3. Further results

The results in this section are irrelevant to the discussion in the final section. Suppose Θ is a set of subsets of some set. We say that Θ is *k-chromatic* if k is the order of the smallest set K for which there is a map $f: \Theta \to K$ with the property that for any pair θ_1, θ_2 of sets in Θ, $f(\theta_1) \neq f(\theta_2)$ whenever θ_1 and θ_2 have a non-empty intersection. We also define Θ to be *k-complement chromatic* if k is the order of the smallest set K for which a map $f: \Theta \to K$ has the property that $f(\theta_1) \neq f(\theta_2)$ whenever θ_1 and θ_2 do not intersect. These definitions correspond precisely to the definitions of chromaticity in the intersection graph of Θ and in its complement. The chromatic properties of subtrees of a tree are rather simple, for by choosing an arbitrary root for the tree it is easy to devise colouring algorithms which yield to the following result:

Theorem 3.1. *A set of subtrees of a tree is k-chromatic, where k is the largest number such that a point T lies in k subtrees. Any such set of subtrees is N-complement chromatic, where $N <$ the number of points in T.*

Using the term "chromatic" in its usual sense for graphs, we can readily deduce two more theorems which are proved by Dirac [3]. (Caution: Dirac uses the term "clique" to denote complete subgraphs rather than maximal complete subgraphs.)

Theorem 3.2 (Berge [1]). *A rigid circuit graph G is k-chromatic, where k is the number of points in the largest clique in G.*

This follows directly from the first part of Theorem 3.1.

Theorem 3.3. *The complement of a rigid circuit graph G is k-chromatic, where k is the number of cliques in G.*

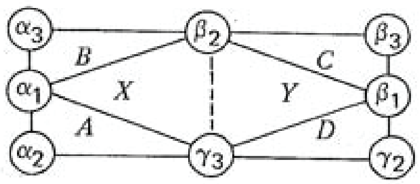
This is a consequence of the second part of Theorem 3.1.

The proof of Theorem 2.7 can also be used to show that the number of cliques in a rigid circuit graph is equal to the maximum number of pairwise unadjacent points in that graph. This fact taken with Theorem 3.3 proves a theorem of Hajnal and Suranyi [3,5].

4. Constructing trees

The author has been involved in two practical problems which involve evolutionary trees. One is the reconstruction of the evolutionary tree of a set of organisms; the other is the recovery of the genealogy of a set of medieval manuscripts which are all directly or indirectly copied from a common source. In each case we observe among these objects certain characteristics which are believed to be hereditary; that is, characteristics that derive from a subtree of the tree which we are trying to construct. In one of these problems these characteristics are introduced by genetic mutation, in the other by scribal error. There are several complications to this, one is that certain characteristics can appear independently in different parts of the tree (two scribes making the same error independently, for example). Another difficulty is that of "conflation"

	f_1	f_2	f_3
A	α_1	α_2	γ_3
B	α_1	β_2	α_3
C	β_1	β_2	β_3
D	β_1	γ_2	γ_3



Example 1.

where one scribe copies from two or more sources. In the latter case the underlying structure does not even have a tree form. But even if we neglect these complications, there is another obstacle which presents itself; this is the problem of dealing with missing objects. In either the manuscript or the biological situation we may, and usually do, have only a subset of the points in the tree we wish to construct and we will want to infer the existence of objects from those that are available to us.

Suppose the descriptions of the objects are presented to us in the form of an attribute table as in Example 1. Here the left-hand column lists the set of objects and the other columns list the values that the various attributes $f_1, f_2, \ldots$ take on these objects. It is the attribute values which we hope will correspond to inherited characteristics and therefore each derive from a subtree of some tree. Is this possible? The overlap graph of the attribute values is also given in Example 1, but we see that it has a non-rigid circuit. In order to make this graph into a rigid circuit graph we would have to add a chord to this circuit and this would mean that two attribute values overlap although there is nothing in the attribute table to say they do. In view of this, only one chord β_2, γ_3 can be drawn, for were we to draw the other chord α_1, β_1, we would be suggesting that two values of the same attribute overlap. Having added this chord, Theorem 2.7 shows us how to build the tree and we notice that the two cliques created by the addition of this chord give us the "missing" objects. We have, of course, found an unrooted tree and the root would have to be inferred from other evidence than that present in the attribute table.

This raises the general question: Supposing there is a "rigidification" of an attribute table, is there a simple method for finding it? Example 2 shows a table of six objects and three attributes which has two very different rigidifications. One would want such a method either to produce

	f_1	f_2	f_3
A	α_1	δ_2	β_3
B	α_1	α_2	δ_3
C	γ_1	α_2	α_3
D	β_1	γ_2	α_3
E	β_1	β_2	γ_3
F	δ_1	β_2	β_3

Example 2.

both trees or to manifest the ambiguity in some interesting way. The problem of detecting conflations poses another problem in graph theory. Given a graph G, what is the simplest graph H for which G is the intersection graph of a set $H_1, H_2, \ldots$ of subgraphs of H? The term "simplest" could be taken to mean "minimum connectivity", but there may be another use of the word which is relevant to this situation.

Acknowledgments

The author would like to thank Mr. S.D. Isard, Dr. D.J. Willshaw and Professor H.C. Longuet-Higgins for their help and criticism. This work was financed by a grant from the Royal Society.

References

[1] C. Berge, Färbung von Graphen, deren sämlichte bzw. deren ungerade Kreise starr sind. Beiträge zur Graphentheorie, Proc. Colloq. held at Halle, March 1960, Wiss. Z. Univ. Halle 10 (1961) 114.

[2] O.P. Buneman, The recovery of trees from measures of dissimilarity, in: Mathematics in the Archaeological and Historical Sciences (Edinburgh Univ. Press, Edinburgh, 1972) 387.

[3] G.A. Dirac, On rigid circuit graphs, Abh. Math. Sem. Univ. Hamburg 25 (1974) 71.

[4] D.R. Fulkerson and O.A. Gross, Incidence matrices and internal graphs, Pacific J. Math. 15(3) (1965) 835.

[5] A. Hajnal and J. Surányi, Über die Ausflosung von Graphen in vollständige Teilgraphen, Ann. Univ. Sci. Budapest. Eötvös, Sect. Math. 1 (1958) 113.

[6] F. Harary, A graph-theoretic approach to similarity relations, Psychometrica 29 (2) (June 1964).

[7] F.S. Roberts, Indifference graphs, in: F. Harary, ed., Proof Techniques in Graph Theory (Academic Press, New York, 1969) 139.