SOME REMARKS ON INTERVAL GRAPHS

R. HALIN

Dedicated to Tibor Gallai on his seventieth birthday

Received 19 March 1982

Using simplicial decompositions a new and simple proof of Lekkerkerker-Boland's criterion for interval graphs is given. Also the infinite case is considered, and the problem is tackled to what extent the representation of a graph as an interval graph is unique.

1. Introduction

The investigation of interval graphs was suggested and initiated by Gallai, Hajós, and Benzer (see [6], [5], [1]). These graphs have a lot of applications (see e.g. [4]). Among the characterizations of interval graphs [2], [3], [8], that by Lekkerkerker and Boland is the deepest and the most beautiful. In fact Lekkerkerker and Boland prove two criteria for interval graphs: In their main result (Theorem 3 of [8]) they characterize the finite interval graphs as the triangulated graphs without an asteroidal triple of vertices (see section 3 below for a definition); from this theorem they then derive a characterization in terms of forbidden induced subgraphs.

This note provides a new and simple proof of Lekkerkerker—Boland's main theorem whose original proof was very long and complicated. The present proof is based on the theory of simplicial decompositions, which seems, in the opinion of the author, to be the natural tool for the investigation of triangulated graphs. As a variation of Lekkerkerker—Boland's criterion we characterize the interval graphs as the graphs with the following property: Among any three cliques there is always at least one which separates the two others.

Also infinite interval graphs are studied and characterized by a new form of decomposition generalizing consecutive simplicial decompositions. We prove further that a graph G is an interval graph if and only if every finite induced subgraph of G is an interval graph. Finally the problem is solved to what extent the representation of a finite interval graph by a consecutive prime graph decomposition is uniquely determined, and this also throws some light onto the question: Which systems of intervals on R determine the same interval graph?

AMS subject classification (1980): 05 C 99; 05 C 38.

298 R. HALIN

2. Prerequisites

We refer to the theory of triangulated graphs and interval graphs as it is developed in [7], Chap. X, especially § 9 and § 11. However in order to get a base for our discussions the fundamental facts are briefly reported in this section.

All graphs in this paper are undirected and do not contain loops or multiple edges. A *simplex* is a complete graph. A *clique* of a graph G is a maximal simplex contained in G. G is called *prime* if it does not contain a separating simplex. Every finite graph G has a *prime graph decomposition* (pgd), that means G can be represented in the form

$$G = G_0 \cup G_1 \cup ... \cup G_k$$

where each G_i is prime and, for each i=1, ..., k,

$$(G_0 \cup \ldots \cup G_{i-1}) \cap G_i = S_i$$

is a simplex which is properly contained in both $G_0 \cup ... \cup G_{i-1}$ and G_i . In addition it can be stipulated that no G_i is a subgraph of a G_j with $j \neq i$; then the pgd is called reduced and it just contains all maximally prime induced subgraphs of G. The S_i are called the simplices of attachment of the pgd in question: they are exactly those simplices in G which minimally separate some pair of vertices of G. Hence the G_i and the S_i are the same in all reduced pgd's of G; however the order in which they occur is not uniquely determined. We note further that S_i separates (in G) each vertex of $G_0 \cup ... \cup G_{i-1} \setminus S_i$ from each vertex of $G_i \setminus S_i$. As a consequence, each S_i must be contained in some G_i with j < i.

A graph G is called *triangulated* if G does not contain a circuit of length ≥ 4 as an induced subgraph. A (finite or infinite) graph is triangulated if and only if its maximally prime induced subgraphs coincide with its cliques. For finite G we can state

(1) G is triangulated if and only if G has a pgd consisting of simplices.

A finite interval graph is the intersection graph of a finite system of intervals on the real line \mathbf{R} . As "small" changes of the intervals do not change the corresponding interval graph we see that it does not matter whether we restrict ourselves to open or closed (bounded) intervals or whether we admit any convex set on \mathbf{R} as an interval. Also it does not matter whether we allow or forbid that distinct vertices of the interval graph are represented by the same interval on \mathbf{R} .

The following observation is the base of the theory of interval graphs:

(2) Every interval graph is triangulated.

The interval graphs are distinguished in the class of triangulated graphs by the possibility to paste the cliques together in a specific way. A pgd G_0, \ldots, G_k of a graph G (with simplices of attachment S_1, \ldots, S_k) is called *consecutive* if $S_i \subset G_{i-1}$ for $i=1,\ldots,k$ holds. Then we have

(3) A finite graph is an interval graph if and only if it has a consecutive pgd consisting of simplices.

This result is closely related to the criterion of Fulkerson—Gross [2], which however is primarily formulated in terms of matrices. For a proof and a generalization of (3) see also section 4 below.

We state some properties of consecutive pgd's.

- **(4)** A pgd $G_0, ..., G_k$ is consecutive if and only if for each vertex v holds: If v is in G_i and in G_h and i < j < h, then v is in G_j .
- (5) Let $G_0, G_1, ..., G_k$ be a consecutive pgd of G consisting of the cliques of G (i.e. G is a finite interval graph). Let S_i denote the simplices of attachment. Then we have:
 - i) Also G_k, \ldots, G_1, G_0 is a consecutive pgd of G.
- ii) For $1 \le i \le k$ any vertex of $(G_0 \cup ... \cup G_{i-1}) \setminus S_i$ is separated (in G) from any vertex of $(G_i \cup ... \cup G_k) \setminus S_i$ by S_i .
- iii) Any $T \subset V(G)$ which separates G must contain the vertices of some S_i .
- iv) For $1 \le i < k$, G_i separates G. Especially G_i separates each G_j with j < i from each G_k with h > i.
- v) If G_i does not separate G, then necessarily i=0 or i=k.
- vi) For an induced subgraph H of G we get a consecutive pgd of H by

$$H \cap G_0, H \cap G_1, ..., H \cap G_k$$

(with the simplices of attachment $H \cap S_i$), where only "superfluous" members have to be omitted. We refer to the latter decomposition as the restriction onto H of the given pgd.

3. The Lekkerkerker—Boland criterion

Three vertices x, y, z of a graph G are said to form an asteroidal triple in G if any two of them are connected by a path of G which avoids the third vertex and all its neighbours. (Especially no pair of vertices of an asteroidal triple can be adjacent.) Clearly if x, y, z form an asteroidal triple in an induced subgraph H of G then they are also an asteroidal triple with respect to G.

Let G be an interval graph and G_0, \ldots, G_k a consecutive pgd of G. Let x, y, z be an arbitrary triple of vertices of G; assume $x \in V(G_i)$, $y \in V(G_j)$, $z \in V(G_k)$ and $i \le j \le h$. Then by (5), iv) every x, z-path in G meets G_j , i.e. y or a neighbour of y. Hence an interval graph cannot contain an asteroidal triple. This property, vice versa, characterizes the interval graphs among the triangulated graphs:

(6) Theorem of Lekkerkerker—Boland. A finite graph G is an interval graph if and only if it is triangulated and does not contain an asteroidal triple of vertices.

Proof. We have to show: If G is triangulated then G has an asteroidal triple or G admits a consecutive pgd. We use induction on the number of cliques of G.

If G has only one clique, i.e. G is a simplex, then the assertion is trivial. Assume now that G has a pgd G_0, \ldots, G_{k+1} (with the simplices of attachment S_i) where the G_i are the cliques of G, and that the assertion is true for graphs with less cliques. We briefly write C instead of G_{k+1} and S instead of G_{k+1} . If $H = G_0 \cup \ldots \cup G_k$ contains an asteroidal triple we have such a triple also in G. Thus let us assume H to be an interval graph, and w.l.o.g. assume G_0, \ldots, G_k to be a consecutive pgd of H.

300 R. HALIN

If $S \subset G_0$ or G_k we place C before G_0 or behind G_k (respectively) and get a consecutive pgd of G. Therefore assume $S \subset G_m$, 0 < m < k. S is properly contained in G_m (otherwise G_m were no clique).

Case 1. $S \supseteq S_i$ for some i, $1 \le i \le k$. Define $S_0 = S_{k+1} = \emptyset$ and put

$$\begin{split} p &= \max \{i | S \supseteq S_i, \ 0 \le i \le m\}, \\ q &= \min \{i | S \supseteq S_i, \ m+1 \le i \le k+1\}, \\ Z &= G_p \cup \ldots \cup G_m \cup \ldots \cup G_{q-1}. \end{split}$$

By assumption

$$(*) p \ge 1 or q \le k$$

By (5), ii) and iii) $Z \setminus S$ is the component of $H \setminus S$ which contains $G_m \setminus S$. By (4) we have $G_j \supseteq S_p$ for $p \le j \le m$, $G_j \supseteq S_q$ for $m \le j \le q-1$. $Z \cup C$ has less cliques than G (by (*)); therefore we may assume that $Z \cup C$ has a consecutive pgd $Z_0 \cup ... \cup Z_t$ containing its cliques $G_p, ..., G_{q-1}, C$ as its members. As S does not separate Z, by (5), v) we have $C = Z_0$ or $= Z_t$; w.l.o.g. assume $C = Z_0$ (otherwise we reverse the order according to (5), i)). Assume $Z_t = G_s$ ($p \le s \le q-1$). If $s \le m$ then $S_p \subseteq G_s = Z_t$, and $G_0, ..., G_{p-1}, Z_t, Z_{t-1}, ..., Z_1, C, G_q, ..., G_k$ is a consecutive pgd of G. If $s \ge m+1$ then $S_q \subseteq G_s = Z_t$, and symmetrically $G_0, ..., G_{p-1}, C, Z_1, ...$..., $Z_t, G_q, ..., G_k$ is a consecutive pgd of G.

Case 2. $S \not\supseteq S_i$ for all i=1, ..., k. If

(i)
$$S_k \supseteq S$$
, $\supseteq S_i$ for $i = 1, ..., k-1$

and

(ii)
$$S_1 \not\supseteq S$$
, $\not\supseteq S_i$ for $i = 2, ..., k$,

then choose vertices x in $G_0 \setminus S_1$, y in $G_k \setminus S_k$, z in $C \setminus S$; any pair of these vertices are connected by a path in G not meeting the third vertex nor one of its neighbours. Thus x, y, z form an asteroidal triple. If on the other hand, (i) is not true then case 1 applies with G_k , S_k instead of C, S, and if (ii) does not hold then case 1 applies with G_0 , S_1 in the rôle of C, S.

4. The infinite case

We want to extend the criterion (3) to infinite graphs. For this purpose we have to generalize the notions of interval graph and of consecutive pgd.

A graph G is called an *interval graph* if there is a totally ordered set (T, \leq) and a family of intervals of T such that G is isomorphic to the intersection graph of this family. Here an interval of (T, \leq) is any "convex" subset I of T, i.e.:

$$(x \le y \le z \text{ and } x, z \in I) \Rightarrow y \in I.$$

Remark. By appropriate extension of (T, \leq) we see that the notion of interval graph is not changed if we restrict ourselves to intervals which are bounded and contain at least 2 elements.

Let G be a graph, $G_i(i \in I)$ a family of subgraphs of G where I is totally ordered by \leq . This family is called a *generalized simplicial decomposition* of G if the following conditions are fulfilled:

- 1) G is the union of the G_i ($i \in I$);
- 2) For every $i \in I$ which is not the minimum of (I, \leq) we have

$$(\bigcup_{j< i} G_j) \cap G_i = S_i$$

is a simplex.

In view of (4) this generalized simplicial decomposition is called consecutive if

$$x \in V(G_i), x \in V(G_h)$$
 and $i < j < h \Rightarrow x \in V(G_i)$.

We speak of a generalized pgd if all the G_i are prime.

It is clear that every pgd in the old sense (see [7], p. 151 and 160) is also a generalized pgd; however in the latter concept it is not stipulated that S_i be properly contained in G_i and $\bigcup_{j < i} G_j$. For finite G the term (consecutive) generalized pgd coincides with the term (consecutive) pgd, if "superfluous" members are omitted. The following result therefore is an extension of (3):

(7) G is an interval graph if and only if G has a consecutive generalized pgd consisting of simplices.

Remark. The proof will show that these simplices can be chosen as all the cliques of G. However we do not need more than |V(G)| of these cliques. Mind that in general G has more than |V(G)| cliques: If we have a non-empty set S of intervals on \mathbb{Q} such that each interval of S contains two disjoint intervals of S, then the corresponding interval graph is countable and has uncountably many cliques.

Proof of (7). a) Assume $G_i(i \in I)$ to be a consecutive generalized pgd of simplices for G, with respect to a total order \leq in I. For every $x \in V(G)$ put $I(x) = \{i | i \in I, v \in V(G_i)\}$. For $x, y \in V(G)$ we have: $xy \in E(G) \Leftrightarrow \exists i \text{ with } x, y \in V(G_i) \Leftrightarrow I(x) \cap I(y) \neq \emptyset$. Hence G is an interval graph.

b) On the other hand suppose G to be an interval graph and let $x \to I(x)$ be a representation of G in a totally ordered set (T, \leq) . Let G_i $(i \in J)$ be the family of cliques of G. Put i < j if there exist $x \in V(G_i \setminus G_j)$, $y \in V(G_j \setminus G_i)$ with $xy \notin E(G)$ such that I(x) < I(y) (this means that $\alpha < \beta$ for all $\alpha \in I(x)$, $\beta \in I(y)$). It follows by straightforward considerations that for any pair of distinct cliques G_i , G_j exactly one of the relations i < j, j < i holds and that < is transitive, hence forms a total order on J. We assert that the G_i in this order form a consecutive generalized pgd of G. Of course G is the union of the G_i . Suppose $x \in V(G_i) \cap V(G_h)$ and i < j < h. If $x \notin V(G_j)$ there exists $y \in V(G_j)$ which is not adjacent to x (because G_j is a clique). Then we had I(x) < I(y) by i < j and I(y) < I(x) by j < h. This contradiction shows that the G_i , with respect to the given order on J, have the consecutive property. Further if vertices $x \ne y$ are in G_i and in $\bigcup_{j < l} G_j$ then $x \in V(G_j)$, $y \in V(G_h)$ for some j, h < i;

if (w.l.o.g.) $j \le h$ then the edge xy lies in G_h . Thus $(\bigcup_{j < i} G_j) \cap G_i$ is a simplex.

302 R. HALIN

Next we prove:

(8) G is an interval graph if and only if every finite induced subgraph of G is an interval graph.

Proof. The "only if" part being trivial we assume that every finite induced subgraph of G is an interval graph. Let G_i $(i \in I)$ be the family of all cliques of G. We consider pairs (J, τ) with $J \subseteq I$, τ a total order on J. We put $(J', \tau') \preceq (J, \tau)$ if $J' \subseteq J$ and $\tau' = \tau|_{I'}$.

Let $\vartheta: F_0, \ldots, F_k$ be a consecutive pgd of some finite induced subgraph F of G. We say ϑ induces i before j (for some $i, j \in I$) if there are non-adjacent vertices x in $F_v \cap G_i$, y in $F_\mu \cap G_j$ with $0 \le v < \mu \le k$. Clearly if ϑ induces i before j and j before k then ϑ induces i before k. Given some (J, τ) , $i, j \in J$ are said to be in discord with ϑ if $j\tau i$ but ϑ induces i before j. If no pair of vertices of J are in discord with ϑ , we say that (J, τ) and ϑ agree. In that case (J, τ) also agrees with the restriction of ϑ onto any induced subgraph of F. (J, τ) is called admissable if for every finite induced subgraph F of G there is a consecutive pgd ϑ of F such that (J, τ) and ϑ agree. If (J, τ) is admissable and $i\tau j$, then for non-adjacent $x \in V(G_i \setminus G_j)$, $y \in V(G_j \setminus G_i)$ any consecutive pgd of an F containing x, y which agrees with (J, τ) necessarily induces i before j. (J, τ) is admissable if and only if $(J', \tau|_{J'})$ is admissable for each finite $J' \subseteq J$. If (J, τ) is admissable then the G_i with $i \in J$, ordered by τ , form a consecutive generalized pgd of their union. (For, if x is in $V(G_i \cap G_h)$, $i\tau j\tau h$ and $x \notin V(G_j)$, choose $y \in V(G_j)$ non-adjacent to x. Then i, k are in discord with each consecutive pgd of any finite F containing x, y.) Of course every (J, τ) with |J| = 1 is admissable.

It follows from our observations that the union of any chain (with respect to \leq) of admissable pairs is again an admissable pair. Hence by Zorn's Lemma there is a maximal admissable pair (J, τ) . We assert J=I.

Otherwise there is an $i \in I \setminus J$. We extend τ to a total order $\hat{\tau}$ on $J \cup \{i\}$ as follows. Let j be an arbitrary element of J.

a) If for each finite induced $F \subseteq G$ there is a consecutive pgd of F which agrees with (J, τ) and does not induce i before j, then put $j\hat{\tau}i$.

b) If there is a finite induced $\tilde{F} \subseteq G$ such that each consecutive pgd of \tilde{F} which agrees with (J, τ) induces i before j, then put $i \hat{\tau} j$.

Assume $i\hat{\tau}j$, $j\tau h$. Choose \tilde{F} according to b) and non-adjacent vertices x in $G_j \setminus G_i$, y in $G_h \setminus G_j$. Let F' be the subgraph induced by \tilde{F} and x, y. Then any consecutive pgd of F' which agrees with (J, τ) induces i before h; hence $i\hat{\tau}h$. Therefore (with $\hat{\tau}|_{J}=\tau$) $\hat{\tau}$ is a total order on $J \cup \{i\}$.

Let F be any finite induced subgraph of G and $\vartheta_1, \ldots, \vartheta_r$ the consecutive pgd's of F which agree with (J, τ) . If $(J \cup \{i\}, \hat{\tau})$ agrees with none of them then there exist j_1, \ldots, j_r in J such that i, j_q are in discord with ϑ_q for $q = 1, \ldots, r$; w.l.o.g. $j_1 \tau j_2 \tau \ldots \tau j_r$. We extend F to an F' such that each consecutive pgd of F' agreeing with (J, τ) induces $j_r \tau j_{\bar{r}}$ for $0 \le v < \lambda \le r$.

Suppose $j_v \hat{\tau} i \hat{\tau} j_{v+1}$. By b) there is \tilde{F} such that each consecutive pgd of \tilde{F} agreeing with (J, τ) induces i before j_{v+1} . Let F'' be the subgraph induced by $F' \cup \tilde{F}$. By a) there is a consecutive pgd ϑ of F'' agreeing with (J, τ) which does not induce i before j_v (hence it does not induce i before j_μ for $1 \le \mu \le \nu$), but by b) it induces i before $j_{\nu+1}$ (hence i before j_μ for $\mu = \nu + 1, \ldots, r$). Thus ϑ agrees with $(J \cup \{i\}, \hat{\tau})$, and so does the restriction of ϑ onto F, a contradiction.

Analogously we come to a contradiction if i is not between two of the j_{ϱ} . It follows J=I, which completes the proof.

As an immediate consequence we have

(9) The Lekkerkerker—Boland criterion also holds for infinite graphs.

Further we get the following new characterization of interval graphs:

- (10) A graph G is an interval graph if and only if among any three cliques of G there is (at least) one which separates the two others.
- **Proof.** 1) Suppose G is an interval graph and C_1 , C_2 , C_3 are distinct cliques such that no pair of these C_i is separated by the third one. Then to each C_i there is a path P_i avoiding C_i and connecting the two other cliques. (Possibly P_i consists of a single vertex.) For any $x \in V(P_i)$ choose a $v_{xi} \in V(C_i)$ such that v_{xi} is not adjacent to x. Then the P_1 and v_{xi} are in a finite induced subgraph H of G such that $H \cap C_i$ are distinct cliques of H no two of which are separated in H by the third one. By (3) H has a consecutive pgd H_0, \ldots, H_k containing the cliques of H. Let $H \cap C_i = H_{v_i}$; w.l.o.g. $v_1 < v_2 < v_3$. But then H_{v_2} separates H_{v_1} from H_{v_3} (by (5), iv)); a contradiction.
- 2) Assume G not to be an interval graph. If G contains an induced circuit C of length ≥ 4 , then choose any three edges e_1, e_2, e_3 of C and cliques G_i containing e_i ; if G contains an asteroidal triple x_1, x_2, x_3 , then choose cliques G_i with $x_i \in V(G_i)$. In both cases the G_i form a triple of cliques no pair of which is separated by the third clique.

5. A uniqueness theorem

In the light of (7) and its proof the possible realizations of an interval graph G by a system of intervals on the real line (or any totally ordered set) correspond, in a certain sense, to the orderings of the cliques fulfilling the consecutivity condition, i.e. to the consecutive (generalized) pgd's which G admits. It is therefore of interest to know to what extent a consecutive pgd of G is uniquely determined. In this last section we solve this problem for the finite case.

Let G be finite and

$$(*)$$
 $G_0, ..., G_{i-1}, G_i, ..., G_j, G_{j+1}, ..., G_k$

be a consecutive pgd of G consisting of its cliques, S_1, \ldots, S_k denoting the simplices of attachment. Assume that $G_i, G_{i+1}, \ldots, G_j$ all contain S_i and S_{j+1} . Then also

$$G_0, ..., G_{i-1}, G_j, G_{j-1}, ..., G_i, G_{j+1}, ..., G_k$$

is a consecutive pgd; we say that the latter arises from (*) by an elementary inversion. If we put $S_0 = S_{k+1} = \emptyset$, then also G_k , G_{k-1} , ..., G_0 is an elementary inversion of (*). We overlook the consecutive pgd's of G by the following result.

(11) Every reduced consecutive pgd of G arises from a fixed one (*) by a sequence of elementary inversions.

Proof. By induction on the number k of cliques of G. If k=1 the assertion is trivial. Let k>1 and (**) be any other consecutive pgd of G. Let S_i be a simplex of attachment which has minimal order. Denote the components of $G\setminus S_i$ by C_1, \ldots, C_t $(t\geq 2)$, and let D_{τ} be the subgraph induced by $C_{\tau}\cup S_i$ $(\tau=1,\ldots,t)$. No D_{τ} is separated by S_i or a subgraph of S_i , and $G=\bigcup_{\tau=1}^t D_{\tau}$. Each D_{τ} has certain $G_{\tau 1},\ldots,G_{\tau s_{\tau}}$ among the G_j as its cliques. In any consecutive pgd of G the $G_{\tau\sigma}$, for fixed τ , form a period because each simplex of attachment of D_{τ} has a vertex which is not in S_i and there fore in no D_{τ} with $\varrho \neq \tau$.

If D_{τ} has a simplex of attachment not containing S_i , then any consecutive pgd of G starts or ends with a pgd of D_{τ} . By inverting if necessary we can assume that (*) and (**) starts with D_{τ} . Then (*) and (**) both consist of a consecutive pgd of D_{τ} followed by one of $\bigcup_{\mu \neq \tau} D_{\mu}$; by application of the induction hypothesis onto D_{τ} and $\bigcup_{\mu \neq \tau} D_{\mu}$ the assertion follows.

It remains to consider the case that in each D_{τ} all the simplices of attachment contain S. Let (*) start with D_{τ} , (**) with D_{ϱ} . In (**) we reverse the section from D_{ϱ} to D_{τ} , to get a consecutive pgd starting with D_{τ} . Then, applying the induction hypothesis for D_{τ} and $\bigcup_{\mu \neq \tau} D_{\mu}$, we complete the proof.

References

- [1] S. BENZER, On the topology of the genetic fine structure, *Proc*, *Nat. Acad. Sci. U.S.A.* 45 (1959), 1607—1620.
- [2] D. R. FULKERSON, and O. A. GROSS, Incidence matrices and interval graphs, Pacific J. Math. 15 (1965), 835—855.
- [3] P. C. GILMORE and A. J. HOFFMAN, A characterization of comparability graphs and interval graphs, Canad. J. Math. 16 (1964), 539—548.
- [4] M. C. GOLUMBIC, Algorithmic graph theory and perfect graphs, Academic Press, New York, 1980.
- [5] G. HAJÓS, Über eine Art von Graphen, Intern. Math. Nachr. 11, Problem 65 (1957).
- [6] A. HAJNAL and J. SURÁNYI, Über die Auflösung von Graphen in vollständige Teilgraphen, Ann. Univ. Sci. Budapestinensis 1 (1958), 113—121.
- [7] R. Halin, Graphentheorie II, Wiss. Buchges., Darmstadt, 1981.
- [8] C. Lekkerkerker and J. Boland, Representation of a finite graph by a set of intervals on the real line, Fundam. Math. 51 (1962), 45—64.

R. Halin

Universität Hamburg Mathematisches Seminar Bundesstr. 55 2000 Hamburg 13 F.R.G.