A CHARACTERIZATION OF COMPARABILITY GRAPHS AND OF INTERVAL GRAPHS

P. C. GILMORE AND A. J. HOFFMAN

1. Introduction. Let $<$ be a non-reflexive partial ordering defined on a set P. Let $G(P, <)$ be the undirected graph whose vertices are the elements of P, and whose edges (a, b) connect vertices for which either $a < b$ or $b < a$. A graph G with vertices P for which there exists a partial ordering $<$ such that $G = G(P, <)$ is called a comparability graph.

In §2 we state and prove a characterization of those graphs, finite or infinite, which are comparability graphs. Another proof of the same characterization has been given in **(2)**, and a related question examined in **(6)**. Our proof of the sufficiency of the characterization yields a very simple algorithm for directing all the edges of a comparability graph in such a way that the resulting graph partially orders its vertices.

Let O be any linearly ordered set. By an interval α of O is meant any subset of O with the same ordering as O and such that, for all a, b, and c, if b is between a and c and a and c are in α, then b is in α. Two intervals of O are said to intersect if and only if they have an element in common.

Let I be any set of intervals on a linearly ordered set O and let $G(O, I)$ be the undirected graph whose vertices are the intervals in I and whose edges (α, β) connect intersecting intervals α and β. A graph G is an interval graph if there exists such an O and I for which $G = G(O, I)$.

In §3 we state and prove a characterization of those graphs, finite or infinite, which are interval graphs. This solves a problem closely related to one first proposed in **(4)**, and independently in **(1)**. A different characterization was given in **(5)**. As a corollary of our result, we are able to determine for any interval graph G the minimum cardinality of a linearly ordered set O for which there is a set of intervals I such that $G = G(O, I)$.

All graphs considered in this paper have no edge joining a vertex to itself.

2. Comparability graphs. By a *cycle* of a graph G is meant here any finite sequence of vertices $a_1, a_2, \ldots, a_k$ of G such that all of the edges (a_i, a_{i+1}), $1 \leqslant i \leqslant k - 1$, and the edge (a_k, a_1) are in G, and for no vertices a and b and integers $i, j < k$, $i \neq j$, is $a = a_i = a_j$, $b = a_{i+1} = a_{j+1}$ or $a = a_i = a_k$, $b = a_{i+1} = a_1$. A cycle is odd or even depending on whether k is odd or even.

Received May 13, 1963. This research was supported in part by the Office of Naval Research under Contract No. Nonr 3775(00), NR 047040. The results were announced, without proofs, in **(3)**.

Note that there can exist cycles in which a vertex appears more than once. For example, in Figure 1, $d, a, b, e, b, c, f, c, a$ is a cycle with nine vertices.

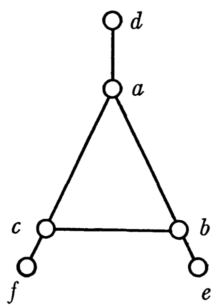

FIGURE 1

By a *triangular chord* of a cycle $a_1, a_2, \ldots, a_k$ of G is meant any one of the edges (a_i, a_{i+2}), $1 \leqslant i \leqslant k-2$, or (a_{k-1}, a_1) or (a_k, a_2). For example, the cycle of nine vertices in Figure 1 has no triangular chords.

THEOREM 1. *A graph G is a comparability graph if and only if each odd cycle has at least one triangular chord.*

Proof. The necessity of the condition is not difficult to establish. For if an odd cycle $a_1, \ldots, a_k$ without a triangular chord occurs in a graph G, then any orientation of the edges of G which is to partially order the vertices of G must give any successive pair of edges of the cycle opposite orientations in the sense that both are directed towards or away from the common vertex of the pair. For if (a, b) and (b, c) are edges of G while (a, c) is not, then if $a \to b$ is the direction given to (a, b), $c \to b$ must be the direction given to (b, c). For the direction $b \to c$ would require, by the transitivity of partial ordering, that (a, c) also be an edge of G. Similarly also, if $b \to a$ is the direction given to (a, b). But only in an even cycle can all successive pairs of edges be given opposite orientations.

Several definitions and lemmas are useful for the argument that the condition of Theorem 1 is also sufficient for G to be a comparability graph.

Two edges (a, b) and (b, c) of a graph G are said to be strongly joined if and only if $(a, c) \notin G$. A path $a_1, \ldots, a_k$ in G is a strong path if and only if for all i, $1 \leqslant i \leqslant k-2$, $(a_i, a_{i+2}) \notin G$. Two edges (a, b) and (c, d) are strongly connected with ends a and d if and only if there exists a strong path $a_1, a_2, \ldots, a_k$, where k is odd and where $a_1 = a, a_2 = b, a_{k-1} = c$, and $a_k = d$. Two edges (a, b) and (c, d) are said to be strongly connected if and only if they are strongly connected with ends a and d or strongly connected with ends a and c.

The justification for the apparently restricted definition of "strongly connected with ends" can be seen in the following simple consequences of the definitions. An edge (a, b) is strongly connected to itself with ends a and a

https://doi.org/10.4153/CJM-1964-055-5 Published online by Cambridge University Press

since a, b, a is a strong path. If $a_1, \ldots, a_k$ is a strong path, then so is $a_2, a_1, \ldots, a_k$, or $a_1, \ldots, a_k, a_{k-1}$, or $a_2, a_1, \ldots, a_k, a_{k-1}$. If (a, b) and (c, d) are strongly connected with ends a and d, then they are also strongly connected with ends b and c.

An immediate property of strong connectedness we state as a lemma.

LEMMA 1. *If (a, b) and (e, f) are strongly connected with ends a and f and if (c, d) and (e, f) are strongly connected with ends d and f, then (a, b) and (c, d) are strongly connected with ends a and d.*

Under the assumption that every odd cycle in G has a triangular chord, the following lemmas can be established.

LEMMA 2. *No edges (a, b) and (c, d) of G are both strongly connected with ends a and d and strongly connected with ends a and c.*

Proof. If $a, b_1(=b), b_2, \ldots, b_k(=c), d$ and $a, b_1'(=b), b_2', \ldots, b_m'(=d), c$ were strong paths with k and m odd, then $a, b_1, b_2, \ldots, b_k, b_m', \ldots, b_1'$ would be an odd cycle without any triangular chords.

LEMMA 3. *Let a, b, c be any triangle in G and let (d, e) be any edge strongly connected to (a, b) with ends b and e. Then one of the following three possibilities must occur:*

(1) *(a, b) is strongly connected to (a, c) with ends b and c;*

(2) *(a, b) is strongly connected to (b, c) with ends a and c;*

(3) *(c, d) and (c, e) are both edges of G and (c, d) is strongly connected to (a, c) with ends a and d and (c, e) is strongly connected to (c, b) with ends b and e.*

Proof. Let $a_1 = b$, $a_2 = a$, $a_3, \ldots, a_{k-1} = d$, $a_k = e$ be a strong path with k odd.

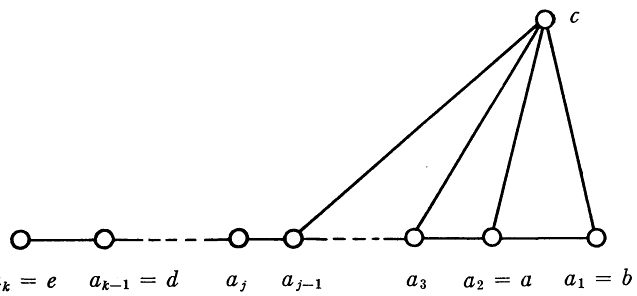

FIGURE 2

Let $j, j \leqslant k$, be such that $(a_i, c) \in G$ for $1 \leqslant i \leqslant j-1$ and $(a_j, c) \notin G$. If j were odd, $a_1, a_2, a_3, \ldots, a_{j-1}, a_j, a_{j-1}, c, a_{j-3}, c, \ldots, c, a_4, c, a_2, c$ would be a strong path with an odd number of vertices and, therefore, (a, b) and (a, c)

would be strongly connected with ends b and c. If j were even, $a, a_1, a_2, a_3, \ldots, a_{j-1}, a_j, a_{j-1}, c, a_{j-3}, c, \ldots, c, a_3, c, a_1, c$ would be a strong path with an odd number of vertices and, therefore, (a, b) and (b, c) would be strongly connected with ends a and c. Thus, if neither (1) nor (2) is to be the case, there can exist no such j. In particular, therefore, no a_i can be identical with c, since that would require that (a_{i-2}, c) not be an edge of G. Hence, we can assume that (a_{k-1}, c) and (a_k, c) are edges of G. But then $a_1, c, a_3, c, \ldots, c, a_k$ and $a_2, c, a_4, c, \ldots, c, a_{k-1}$ are both strong paths with an odd number of vertices. Therefore, (3) must be the case.

Two corollaries follow immediately from the lemma.

COROLLARY 1. *Let a, b, c be any triangle of G and let d be any vertex for which (c, d) is an edge strongly connected to (a, b). Then one of the possibilities* (1) *or* (2) *of Lemma* 3 *must occur.*

Proof. Since (c, d) is strongly connected to (a, b), it is strongly connected either with ends b and c or with ends b and d. But, in either case, possibility (3) would require that there be an edge joining c to c.

COROLLARY 2. *In a triangle a, b, c of G, if (a, b) and (a, c) are strongly connected with ends b and a, then (a, b) and (b, c) are strongly connected with ends a and c.*

Proof. Let d in Corollary 1 be taken to be a. By hypothesis (a, b) and (a, c) are strongly connected and hence, by Corollary 1, either (a, b) and (b, c) are strongly connected with ends a and c or (a, b) and (a, c) are strongly connected with ends b and c. But, the latter alternative is not possible by Lemma 2 and the hypothesis of the corollary, so that the former alternative is necessarily true.

The proof of the sufficiency for Theorem 1 will provide an algorithm for actually directing all the edges of a comparability graph in such a way that the resulting directed graph partially orders its vertices. The description of the algorithm will require some further definitions involving graphs G' which consist of the same vertices and edges of G but with some of the edges directed.

An edge (a, b) of G' is said to have a strongly determined direction $b \to a$, or $a \to b$, if it is strongly connected with ends a and d to a directed edge (c, d) of G' with direction $c \to d$, or $d \to c$ respectively. Hence, any undirected edge strongly connected to a directed edge has a strongly determined direction which depends upon the direction assigned to the directed edge, and depends upon the ends of the strong path joining the directed edge and the undirected edge.

An edge (a, b) of G' is said to have a transitively determined direction $a \to b$ if there are directed edges (a, c) and (c, b) in G' with directions $a \to c$ and $c \to b$.

G' is consistent if and only if there is no directed cycle; that is, there is no cycle $a_1, \ldots, a_k$ such that $a_1 \to a_2, a_2 \to a_3, \ldots, a_{k-1} \to a_k, a_k \to a_1$ are the directions assigned to its edges. Note that if an edge has two directions in G', then there is a directed cycle in G'.

G' is complete with respect to strong connection if no undirected edge of G' has a direction strongly determined from G'. G' is complete if it is complete with respect to strong connection, and further no undirected edge of G' has a direction transitively determined from G'.

For any edge (a, b) of G let $G(a \to b)$ be the graph obtained from G by giving the edge (a, b) the direction $a \to b$ and by then giving any edge with a determined direction, whether it is already directed or not, that determined direction. By Lemma 2 it follows that no edge of $G(a \to b)$ has two directions assigned to it. For any G' let $G' \cup G(a \to b)$ be the graph obtained from G by giving any of its edges that are directed in either G' or $G(a \to b)$ the directions it has in G' and $G(a \to b)$. For some G' and $G(a \to b)$ it is, therefore, possible for some edges of $G' \cup G(a \to b)$ to receive two directions.

LEMMA 4. *For any edge (e, f) of G, $G(e \to f)$ is consistent and complete.*

Proof. Let $F = G(e \to f)$. We shall show first that F is complete. By definition, F is complete with respect to strong connection. To show that it is complete with respect to transitive connections, let c, a, and b be any three vertices of G for which (c, a) and (a, b) are edges of G which have been assigned the directions $c \to a$ and $a \to b$ in F. Necessarily, (c, b) is also an edge of F, for otherwise (a, c) and (a, b) would be strongly joined and therefore each would have been assigned two directions, which, as we noted above, is not possible. Further, (a, c) and (e, f) are strongly connected with ends a and f, and (a, b) and (e, f) are strongly connected with ends b and f, so that from Lemma 1 it follows that (a, c) and (a, b) are strongly connected with ends a and b. By Corollary 2 to Lemma 3, therefore, (a, b) and (b, c) are strongly connected with ends a and c. Again, by Lemma 1, then (b, c) and (e, f) are strongly connected with ends c and e. Hence, the edge (b, c) must have received the direction $c \to b$ in F.

The consistency of F is then immediate. For, if c, a, and b are consecutive vertices of a directed cycle in F, then from $c \to a$ and $a \to b$ will follow that (c, b) is in G and is directed $c \to b$. Hence, for any directed cycle in F there is a smaller one, and since there cannot be one with two vertices, there can be none at all.

LEMMA 5. *If G' is complete and consistent and (e, f) is any undirected edge in G', then $G' \cup G(e \to f)$ is consistent.*

Proof. Let $F = G(e \to f)$. There are certainly no directed cycles of two vertices in $G' \cup F$ since that would require that a directed edge of F be strongly connected to a directed edge of G' and, therefore, that (e, f) be directed in G'.

Let there be a directed cycle of more than two vertices in $G' \cup F$. Since both G' and F are consistent, the cycle must have edges both (directed) in

G' and in F. If any two consecutive edges of a directed cycle are in G', then since G' is complete, necessarily the chord joining their ends is in G' and so directed that a smaller cycle can be found. We can, therefore, assume that there are consecutive vertices a, b, c, and d in a directed cycle such that $a \to b$ and $c \to d$ are directions assigned in F and $b \to c$ is the direction assigned in G'. Then (a, b) and (c, d) are strongly connected, while (a, b) and (b, c) are not. Further, (a, c) must exist; otherwise, (a, b) and (b, c) would be strongly joined, contradicting $a \to b$ in F, $b \to c$ in G'. From Corollary 1 of Lemma 3 it follows that (a, b) and (a, c) are strongly connected with ends b and c. Hence (a, c) is assigned the direction $a \to c$ in F. But this argument permits one to obtain from any directed cycle in $G' \cup F$ a directed cycle in F, which is not possible.

LEMMA 6. *If G' is consistent and complete with respect to strong connections and the undirected edge (a, b) has $a \to b$ as a transitively determined direction, then $G' \cup G(a \to b)$ is consistent.*

Proof. Let $T = G(a \to b)$. We shall show first that every directed edge in T has a transitively determined direction in G' which is the same as the direction given to it in T. For, let (d, e) be any directed edge in T. We can assume without loss in generality that (a, b) and (d, e) are strongly connected with ends b and e. Since (a, b) is undirected in G', it is necessarily not strongly connected to the directed edges (a, c) and (b, c) in G', which gave (a, b) its transitively determined direction. Possibility (3) of Lemma 3 must, therefore, occur. But, since (a, c) and (c, b) have the directions $a \to c$ and $c \to b$ in G', necessarily (c, d) and (c, e) have the directions $d \to c$ and $c \to e$, while (d, e) has the direction $d \to e$ in T.

But, it is therefore possible to replace any directed cycle in $G' \cup T$ by a directed cycle in G' since each edge in the cycle which is in T can be replaced by the two directed edges of G' which transitively determine its direction. This completes the proof of Lemma 6.

Consider now the following algorithm for assigning directions to all the edges of G. Initially in the algorithm G' is G.

(1) Choose any undirected edge (a, b) of G' and a direction $a \to b$ for it; let $G' = G' \cup G(a \to b)$ and go to (2). If there is no undirected edge in G', then stop.

(2) If there is an edge (a, b) of G' with a transitively determined direction $a \to b$, then let $G' = G' \cup G(a \to b)$ and go to (2). If there is no such edge, then go to (1).

It is evident that $G' \cup G(e \to f)$ in Lemma 4 and $G' \cup G(a \to b)$ in Lemma 5 are complete with respect to strong connections. Hence, from Lemmas 4, 5, and 6, one sees that in the finite case the algorithm will produce a partial ordering of the vertices of G consonant with the edges of G. In the infinite case (and the argument embraces the finite case as well), we could partially order all consistent G' with $G' < G''$ if $a \rightarrow b$ in G' implies $a \rightarrow b$ in G''. This partially ordered set has a maximal simply ordered set, by Zorn's lemma, and it is easy to see that the union of the G' in this simply ordered set is a G_0', which is also consistent. If not, every edge in G has been assigned a direction in G_0'; then, using either Lemma 5 or Lemma 6, we would have a contradiction of the maximality of G_0'.

COROLLARY. *Let G' be G with some of its edges directed, where G satisfies the hypothesis of Theorem 1. A necessary and sufficient condition that it be possible to give all edges of G' a direction which partially orders its vertices is that the completion of G' with respect to all strongly determined directions has no directed cycle.*

For, the algorithm given above (or the use of Zorn's lemma) could have begun with any consistent G'.

3. Interval graphs. If G is any graph, then G^c is the complementary graph; that is, G^c has the same vertices as G but has an edge connecting two vertices if and only if that edge does not occur in G.

THEOREM 2. *A graph G is an interval graph if and only if every quadrilateral in G has a diagonal and every odd cycle in G^c has a triangular chord.*

Proof. The necessity of the conditions is readily seen. For, let α, β, and γ be three intervals such that both α and β and β and γ overlap while α and γ do not overlap. Then, any interval overlapping both α and γ must of necessity overlap β. Also, if α and β are any two intervals that do not overlap, i.e. in G^c an edge joins the vertices corresponding to α and β, then we say $\alpha < \beta$ if every element of α precedes (in O) every element of β. This is clearly a partial ordering; hence G^c is a comparability graph.

To prove the sufficiency of the conditions, we shall show how to construct for any G satisfying the conditions a linearly ordered set O and a set of intervals I from O such that $G = G(O, I)$.

Since G^c is a comparability graph, we can by Theorem 1 assume that all of its edges have been directed in such a way as to partially order its vertices. Because G satisfies the characterizing conditions, the directing of the edges of G^c will also be such as to satisfy the following lemma.

LEMMA. *Let a, b, c, and d be any vertices of G for which (a, b) is an edge of G, (c, d) is an edge of G if $c \neq d$, and for which (a, c) and (b, d) are edges of G^c. Then (a, c) and (b, d) are both directed towards or away from (a, b).*

Proof. If $c \rightarrow a$ and $b \rightarrow d$ are the directions assigned to the edges, then necessarily $c \neq d$, since otherwise transitivity would require that (a, b) be an edge of G^c rather than of G. Also, necessarily, either (a, d) or (b, c) is an edge of G^c, since otherwise a, d, c, b would be a quadrilateral of G without a diagonal. But neither (a, d) nor (b, c) can be an edge of G^c, since neither could be assigned a direction which would not require by transitivity that either (a, b) or (c, d) be an edge of G.

Define a complete subgraph of a graph to be a set of vertices each pair of which is joined by an edge, and a maximally complete subgraph to be one properly contained in no other complete subgraph.

Consider now any set of maximally complete subgraphs of G such that every vertex and edge of G is in at least one of them. Form a graph $\bar{G}$ with vertices such a set of maximally complete subgraphs, and with an edge joining each pair of vertices of G. Every pair of maximally complete subgraphs of G necessarily has at least one edge of G^c connecting a vertex of one to a vertex of the other. Hence, each edge of $\bar{G}$ can be given one or more directions depending upon the directions of edges of G^c joining the maximally complete subgraphs of G corresponding to the ends of the edge. But, from the lemma, it is immediate that each edge in $\bar{G}$ receives a unique direction so that $\bar{G}$ can be regarded as a complete graph with every edge directed.

The directed graph $\bar{G}$ is transitive. For, if not, there would exist three maximally complete subgraphs G_1, G_2, and G_3 of G and six vertices (possibly not all distinct) a, b in G_1, c, d in G_2, and e, f in G_3 such that (b, c), (d, e), and (f, a) are all edges in G^c and have the directions $b \to c$, $d \to e$, and $f \to a$ as in Figure 3, where, if $a \neq b$ (a, b) is an edge of G, if $c \neq d$ (c, d) is an edge of G, and if $e \neq f$ (e, f) is an edge of G. But $a = b$, $c = d$, and $e = f$ is not

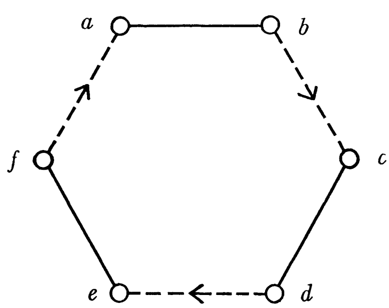

FIGURE 3

possible since transitivity would be violated in G^c; assume, therefore, that $a \neq b$. We may assume that $a \neq d$ and (a, d) is an edge of G^c, since otherwise the vertices a, d, e, and f would contradict the lemma. Again from the lemma it follows that $a \to d$ is the direction assigned to (a, d) in G^c. From transitivity in G^c, therefore, it follows that (a, e) is in G^c and is directed $a \to e$. But then the vertices a, e, f contradict the lemma. Hence, $\bar{G}$ is transitive.

Since $\bar{G}$ is directed and transitive and since every pair of vertices in $\bar{G}$ has an edge joining them, it linearly orders its vertices. Let O be the vertices of $\bar{G}$ linearly ordered by $\bar{G}$.

We shall say that a vertex of G is a member of an element of O if and only if it is a vertex of the maximally complete subgraph of G corresponding to the element of O. If a vertex a of G is a member of two elements G_1 and G_3 of O, then it is a member of every element G_2 of O lying between G_1 and G_3. For, if not, there would be a vertex b of G_2 not connected to a in G and the edge (a, b) of G^c would have to receive two different directions since G_2 lies between G_1 and G_3. Hence, for any vertex a of G, the set $\alpha(a)$ of all elements of O of which a is a member is an interval of O. Let I be the set of all such intervals of O.

It is immediate that $G = G(O, I)$, for the elements of O correspond to a set of maximally complete subgraphs of G which cover all the vertices and edges of G. Hence, two intervals $\alpha(a)$ and $\alpha(b)$ of I overlap if and only if (a, b) is an edge of G.

COROLLARY. *There is a set O of cardinality equal to the least cardinality of a set of maximally complete subgraphs that contain all the vertices and edges of G. This is the set O of least cardinality.*

Proof. If $G = G(O, I)$, then the set of intervals in I containing a given element of O is a maximally complete subgraph of G.

When G is finite and an interval graph, the only set of maximally complete subgraphs containing all the vertices and edges of G is the set of all maximally complete subgraphs. For, let O and I be as constructed in the proof of the theorem and let G_1 be a maximally complete subgraph which does not correspond to an element of O. The directed edges of G^c, as above, will linearly order $O \cup \{G_1\}$ in such a way that if a vertex of G is a member of any two elements of $O \cup \{G_1\}$, then it is a member of all elements lying between the two. Hence, G_1 cannot be an end-point of $O \cup \{G_1\}$ since it would be necessary that its immediate neighbour in $O \cup \{G_1\}$ contain all of its vertices. But also G_1 cannot be between two other elements G_2 and G_3 since there must be a vertex a of G_1 which is not in G_3 and a vertex b of G_1 which is not in G_2. Since the vertices of G_1 must be contained in its immediate neighbours if they are to be contained in any elements of $O \cup \{G_1\}$, it follows that a is in G_2 and b is in G_3. But the edge (a, b) is in G_1 and, hence, must be in some member G_4 of $O \cup \{G_1\}$, which, therefore, necessarily contains both a and b. G_4 cannot be between G_2 and G_3, since we assumed G_2 and G_3 to be immediate neighbours of G_1. Yet, neither can G_2 lie between G_4 and G_3, nor can G_3 lie between G_2 and G_4, since the first case would imply that b is in G_2, while the second case would imply that a is in G_3.

When G is infinite, however, and an interval graph, then a proper subset of the set of all maximally complete subgraphs may cover all edges and vertices of G. For example, consider the interval graph R arising from the set of all open intervals on the real line. Let S be the set of all maximally complete subgraphs of R, each of which is generated by the intervals containing a rational point. Then S covers all the vertices and edges of R even though the cardinality of A is strictly less than the cardinality of R.

References

1. S. Benzer, *On the topology of the genetic fine structure*, Proc. Natl. Acad. Sci. U.S., *45* (1959), 1607–1620.

2. A. Ghouilà-Houri, *Caractérisation des graphes nonorientés dont on peut orienter les arêtes de manière à obtenir le graphe d'une relation d'ordre*, C. R. Acad. Sci. Paris, *254* (1962), 1370–1371.

3. P. C. Gilmore and A. J. Hoffman, *Characterizations of Comparability and Interval Graphs*, Abstract, Internat. Congress Mathematicians (Stockholm, 1962), p. 29.

4. G. Hajos, *Über eine Art von Graphen*, Intern. Math. Nachr., *11* (1957), Sondernummer 65.

5. C. G. Lekkerkerker and J. Ch. Bohland, *Representation of a finite graph by a set of intervals in the real line*, Fund. Math., *51* (1962), 45–64.

6. E. S. Wolk, *The comparability graph of a tree*, Proc. Am. Math. Soc., 13 (1962), 789–795.

IBM Research Center