CHAPTER 8

Interval Graphs

1. How It All Started

In 1957 G. Hajös posed the following problem:

> Given a finite number of intervals on a straight line, a graph associated with this set of intervals can be constructed in the following manner: each interval corresponds to a vertex of the graph, and two vertices are connected by an edge if and only if the corresponding intervals overlap at least partially. The question is whether a given graph is isomorphic to one of the graphs just characterized (Hajös [1957, p. 65, translated by M.C.G.]).

Independently, the well-known molecular biologist, Seymour Benzer, during his investigations of the fine structure of the gene, asked a related question.

> From the classical researches of Morgan and his school, the chromosome is known as a linear arrangement of hereditary elements, the *genes.* These elements must have an internal structure of their own. At this finer level, within the *gene* the question arises again: . . . Are *they* [the subelements within the gene] linked together in a linear order analogous to the higher level of integration of the genes in the chromosome?
>
> A crucial examination of the question should be made from the point of view of *topology,* since it is a matter of how parts of the structure are connected to each other, rather than of the distances between them. Experiments to explore the topology should ask *qualitative* questions (e.g., do two parts of the structure touch each other or not?) rather than *quantitative* ones (how far apart are they?). (Benzer [1959].)

The solution to this question would be found by studying those graphs which represent intersecting intervals on a line, and then verifying whether or not the data that was gathered was consistent with the linear genetic hypothesis.

Our treatment of interval graphs began in Chapter 1. Let us continue looking into the properties of this interesting and useful class of graphs. The reader may wish to review Section 1.3 at this point.

2. Some Characterizations of Interval Graphs

The following theorem and its corrollary will establish where the class of interval graphs belongs in the world of perfect graphs.

Theorem 8.1 (Gilmore and Hoffman [1964]). Let G be an undirected graph. The following statements are equivalent.

(i) G is an interval graph.

(ii) G contains no chordless 4-cycle and its complement $\bar{G}$ is a comparability graph.

(iii) The maximal cliques of G can be linearly ordered such that, for every vertex x of G, the maximal cliques containing x occur consecutively.

Proof. (i) $\Rightarrow$ (ii) This was proved in Chapter 1, Propositions 1.2 and 1.3.

(ii) $\Rightarrow$ (iii) Let us assume that $G = (V, E)$ contains no chordless 4-cycle, and let F be a transitive orientation of the complement $\bar{G}$.

Lemma A. Let A_1 and A_2 be maximal cliques of G.

(a) These exists an edge in F with one endpoint in A_1 and the other endpoint in A_2.

(b) All such edges of $\bar{E}$ connecting A_1 with A_2 have the same orientation in F.

Proof of Lemma A. (a) If no such edge exists in F, then $A_1 \cup A_2$ is a clique of G, contradicting maximality.

(b) Suppose $ab \in F$ and $dc \in F$ with $a,c \in A_1$ and $b,d \in A_2$. We must show a contradiction. If either $a = c$ or $b = d$, then transitivity of F immediately gives a contradiction; otherwise, these four vertices are distinct (Figure 8.1), and ad or bc is in $\bar{E}$, since E may not have a chordless 4-cycle. Assume, without loss of generality, that $ad \in \bar{E}$; which way is it oriented? Using the

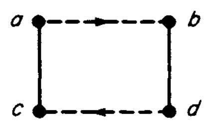

Figure 8.1. Solid edges are in E; broken edges are in $\bar{E}$. Arrows denote the orientation F.

transitivity of F, $ad \in F$ (resp. $da \in F$) would imply $ac \in F$ (resp. $db \in F$), which is impossible, and Lemma A is proved.

Consider the following relation on the collection $\mathscr{C}$ of maximal cliques: $A_1 < A_2$ iff there is an edge of F connecting A_1 with A_2 which is oriented toward A_2. By Lemma A, this defines a tournament (complete orientation) on $\mathscr{C}$. We claim that $(\mathscr{C}, <)$ is a transitive tournament, and hence linearly orders $\mathscr{C}$. For suppose $A_1 < A_2$ and $A_2 < A_3$; then there would be edges $wx \in F$ and $yz \in F$ with $w \in A_1$, $x,y \in A_2$, and $z \in A_3$. If either $xz \notin E$ or $wy \notin E$, then $wz \in F$ and $A_1 < A_3$. Therefore, assume that the edges wy, yx, and xz are all in E (see Figure 8.2). Since G contains no chordless 4-cycle, $wz \notin E$, and the transitivity of F implies $wz \in F$. Thus $A_1 < A_3$, which proves the transitive tournament claim.

Next, assume that $\mathscr{C}$ has been linearly ordered $A_1, A_2, \ldots, A_m$ according to the relation above (i.e., $i < j$ iff $A_i < A_j$). Suppose there exist cliques $A_i < A_j < A_k$ with $x \in A_i$, $x \notin A_j$, and $x \in A_k$. Since $x \notin A_j$, there is a vertex $y \in A_j$ such that $xy \notin E$. But $A_i < A_j$ implies $xy \in F$, whereas $A_j < A_k$ implies $yx \in F$, contradiction. This proves (iii).

(iii) $\Rightarrow$ (i) For each vertex $x \in V$, let $I(x)$ denote the set of all maximal cliques of G which contain x. The sets $I(x)$, for $x \in V$, are intervals of the linearly ordered set $(\mathscr{C}, <)$. It remains to be shown that

$$xy \in E \Leftrightarrow I(x) \cap I(y) \neq \varnothing \qquad (x, y \in V).$$

This obviously holds, since two vertices are connected if and only if they are both contained in some maximal clique. ∎

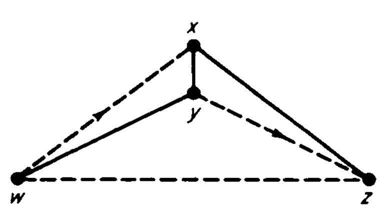

Figure 8.2.

Corollary 8.2. An undirected graph G is an interval graph if and only if G is a triangulated graph and its complement $\bar{G}$ is a comparability graph.

Remark. The coloring, clique, stable set, and clique cover problems can be solved in polynomial time for interval graphs by using the algorithms of Chapters 4 and 5. A recognition algorithm can be obtained by combining Algorithms 4.1 and 5.2, although the recognition algorithm to be presented in Section 8.3 will be asymptotically more efficient.

Statement (iii) of the Gilmore–Hoffman theorem has an interesting matrix formulation. A matrix whose entries are zeros and ones, is said to have the *consecutive 1's property for columns* if its rows can be permuted in such a way that the 1's in each column occur consecutively. In Figure 8.3 the matrix $\mathbf{M}_1$ has the consecutive 1's property for columns since its rows can be permuted to obtain $\mathbf{M}_2$. Matrix $\mathbf{M}_3$ does not possess the property. Consider the *clique matrix* $\mathbf{M}$ (maximal cliques-versus-vertices incidence matrix) of a graph G. The following corollary to Theorem 8.1 is immediate.

Theorem 8.3 (Fulkerson and Gross [1965]). An undirected graph G is an interval graph if and only if its clique matrix $\mathbf{M}$ has the consecutive 1's property for columns.

Proof. An ordering of the maximal cliques of G corresponds to a permutation of the rows of $\mathbf{M}$. This theorem follows from Theorem 8.1. ∎

The earliest characterization of interval graphs was obtained by Lekerkerker and Boland. Their result embodies the notion that an interval graph cannot branch into more than two directions, nor can it circle back onto itself.

Theorem 8.4 (Lekkerkerker and Boland [1962]). An undirected graph G is an interval graph if and only if the following two conditions are satisfied:

(i) G is a triangulated graph, and
(ii) any three vertices of G can be ordered in such a way that every path from the first vertex to the third vertex passes through a neighbor of the second vertex.

Three vertices which fail to satisfy (ii) are called an *astroidal triple.* They would have to be pairwise nonadjacent, but any two of them would have to be connected by a path which avoids the neighborhood of the remaining vertex. Thus, G is an interval graph if and only if G is triangulated and contains no astroidal triple. Condition (ii) illustrates a well-known law of the business world: Every shipment from a supplier to the consumer must pass by the middle man.

$$\begin{array}{c} ① \\ ② \\ ③ \\ ④ \\ ⑤ \end{array}\!\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \rightarrow \begin{array}{c} ① \\ ④ \\ ② \\ ⑤ \\ ③ \end{array}\!\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{M}_1 \qquad\qquad\qquad \mathbf{M}_2 \qquad\qquad\qquad \mathbf{M}_3$$

Figure 8.3. Matrix $\mathbf{M}_1$ has the consecutive 1's property for columns since it can be transformed into $\mathbf{M}_2$. Matrix $\mathbf{M}_3$ does not have the consecutive 1's property for columns since it cannot be suitably transformed.

3. The Complexity of Consecutive 1's Testing

Interval graphs were characterized as those graphs whose clique matrices satisfy the consecutive 1's property for columns (Theorem 8.3). We may apply this characterization to a recognition algorithm for interval graphs $G = (V, E)$ in a two-step process. First, verify that G is triangulated and, if so, enumerate its maximal cliques. This can be executed in time proportional to $|V| + |E|$ (Corollary 4.6, Theorem 4.17) and will produce at most $n = |V|$ maximal cliques (Proposition 4.16). Second, test whether or not the cliques can be ordered so that those which contain vertex v occur consecutively for every $v \in V$. Booth and Leuker [1976] have shown that this step can also be executed in linear time. We shall look at the main ideas behind their algorithm and its implementation. The interested reader should consult their very readable paper for additional details. Subject to Corollary 8.8 and Exercise 3 we have the following.

Theorem 8.5 (Booth and Leuker [1976]). Interval graphs can be recognized in linear time.

The general consecutive arrangement problem is the following: *Given a finite set X and a collection $\mathscr{I}$ of subsets of X, does there exist a permutation π of X in which the members of each subset $I \in \mathscr{I}$ appear as a consecutive subsequence of π?* In the interval graph problem, X is the set of maximal cliques and $\mathscr{I} = \{I(v)\}_{v \in V}$, where $I(v)$ is the set of all maximal cliques containing v. The consecutive arrangement and consecutive 1's problems are equivalent: The rows of the matrix constitute X, and each column determines, or is determined by, a subset of X consisting of those rows containing a 1 in the specified column. Tucker [1972] has characterized the consecutive 1's problem in terms of forbidden configurations. Another characterization, due to Nakano [1973a], is stated as Exercise 12.

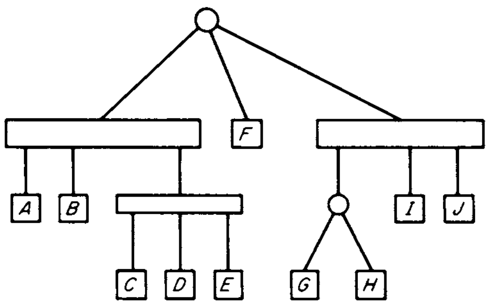

Figure 8.4. A PQ-tree.

Besides its use in recognizing interval graphs, the consecutive 1's problem has a number of other applications. These include a linear-time algorithm for recognizing planar graphs (see Booth and Leuker [1976]), and a storage allocation problem to be discussed in the next section (Application 8.4).

The data structure needed to solve the consecutive arrangement problem most efficiently is the PQ-tree. PQ-trees were invented by Leuker [1975] and Booth [1975] expressly for this purpose. They are used to represent *all* the permutations of X which are consistent with the constraints of consecutivity determined by $\mathscr{I}$. Most importantly, only a small amount of storage is required for this representation.

A *PQ-tree* T is a rooted tree whose internal nodes are of two types: P and Q. The children of a type P-node occur in no particular order, while those of a Q-node appear in an order which must be locally preserved. This will be explained in the next paragraph. We designate a P-node by a circle and a Q-node by a wide rectangle. The leaves of T are labeled bijectively by the elements of the set X (see Figure 8.4).

The *frontier* of a tree T is the permutation of X obtained by reading the labels of the leaves from left to right. In our example, the frontier is $[A\ B\ C\ D\ E\ F\ G\ H\ I\ J]$. Two PQ-trees T and T' are *equivalent*, denoted $T \equiv T'$, if one can be obtained from the other by applying a sequence of the following transformation rules.

1. Arbitrarily permute the children of P-node.
2. Reverse the children of a Q-node.

Figure 8.5 illustrates a PQ-tree which is equivalent to the tree in Figure 8.4. Its frontier is $[F\ J\ I\ G\ H\ A\ B\ E\ D\ C]$. Parenthetically, we obtain an equivalent tree by regarding T as a mobile and exposing it to a gentle summer breeze.

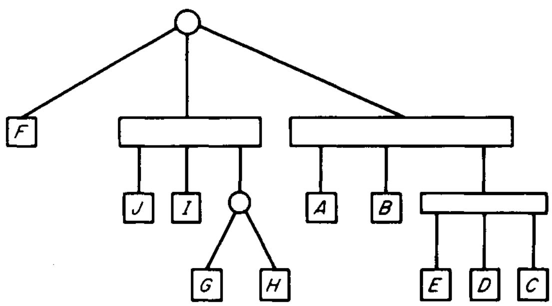

Figure 8.5. A PQ-tree equivalent to the tree in Figure 8.4.

Finally, any frontier obtainable from a tree equivalent with T is said to be *consistent* with T, and we define

$$\text{CONSISTENT}(T) = \{\text{FRONTIER}(T') \mid T' \equiv T\}.$$

It can be shown that the classes of consistent permutations of PQ-trees form a lattice. The null tree T_0 has no nodes and CONSISTENT$(T_0) = \varnothing$. The *universal tree* has one internal P-node, the *root*, and a leaf for every member of X (Figure 8.6).

Let us now relate PQ-trees to the consecutive arrangement problem. Let $\mathscr{I}$ be a collection of subsets of a set X, and let $\Pi(\mathscr{I})$ denote the collection of all permutations π of X such that the members of each subset $I \in \mathscr{I}$ occur consecutively in π. For example if $\mathscr{I} = \{\{A, B, C\}, \{A, D\}\}$, then $\Pi(\mathscr{I}) = \{[D\,A\,B\,C], [D\,A\,C\,B], [C\,B\,A\,D], [B\,C\,A\,D]\}$. We have the following important correspondence.

Theorem 8.6 (Booth and Leuker [1976]). (i) For every collection of subsets $\mathscr{I}$ of X there exists a PQ-tree T such that $\Pi(\mathscr{I}) = \text{CONSISTENT}(T)$.

(ii) For every PQ-tree T there exists a collection of subsets $\mathscr{I}$ such that $\Pi(\mathscr{I}) = \text{CONSISTENT}(T)$.

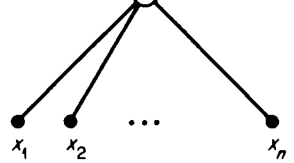

Figure 8.6. The universal tree T_u. CONSISTENT (T_u) includes all permutations of $X = \{x_1, x_2, \ldots, x_n\}$.

Note that the effect of Q-nodes is to restrict the number of permutations by making some of the brother relationships rigid. We leave it to the reader to verify that the tree in Figure 8.4 corresponds to the collection

$$\mathscr{I} = \{\{A, B\}, \{C, D\}, \{D, E\}, \{B, C, D, E\}, \{I, J\}, \{G, H\}, \{G, H, I\}\}.$$

The following procedure clearly calculates $\Pi(\mathscr{I})$.

procedure CONSECUTIVE(X, $\mathscr{I}$):
begin
  $\Pi \leftarrow \{\pi \mid \pi$ is a permutation of $X\}$;
  for each $I \in \mathscr{I}$ **do**
    $\Pi \leftarrow \Pi \cap \{\pi \mid$the members of I occur consecutively in $\pi\}$;;
  return Π;
end

Any naive implementation of this algorithm would be impractical because of the initially exponential size of Π. However, using PQ-trees we can represent Π with only $O(|X|)$ space. The equivalent program using PQ-trees is as follows.

procedure CONSECUTIVE(X, $\mathscr{I}$):
begin
  $T \leftarrow$ universal tree;
  for each $I \in \mathscr{I}$ **do**
    $T \leftarrow$ REDUCE(T, I);;
  return T;
end

This version makes use of a pattern matching routine REDUCE which attempts to apply a set of 11 *templates*. Each template consists of a *pattern* to be matched against the current PQ-tree and a *replacement* to be substituted for the pattern. The templates are applied from the bottom to the top of the tree. The null tree may be returned when no template applies. Two examples are illustrated in Figure 8.7. For details of the algorithm, the reader is directed to Booth and Leuker [1976]. There you will find the templates, a detailed version of the algorithm, a proof of correctness, and a proof of the following complexity theorem.

Theorem 8.7 (Booth and Leuker [1976]). The class of permutations $\Pi(\mathscr{I})$ can be computed in $O(|\mathscr{I}| + |X| + \text{SIZE}|\mathscr{I}|)$ steps where $\text{SIZE}(\mathscr{I}) = \sum_{I \in \mathscr{I}} |I|$.

In the theorem the word *computed* means computed in its PQ-tree representation T. In the consecutive arrangements problem it is not necessary to calculate all of $\Pi(\mathscr{I})$. Rather, it is enough to produce one member of $\Pi(\mathscr{I})$ or to determine that $\Pi(\mathscr{I})$ is empty. This can be done very simply by calculating

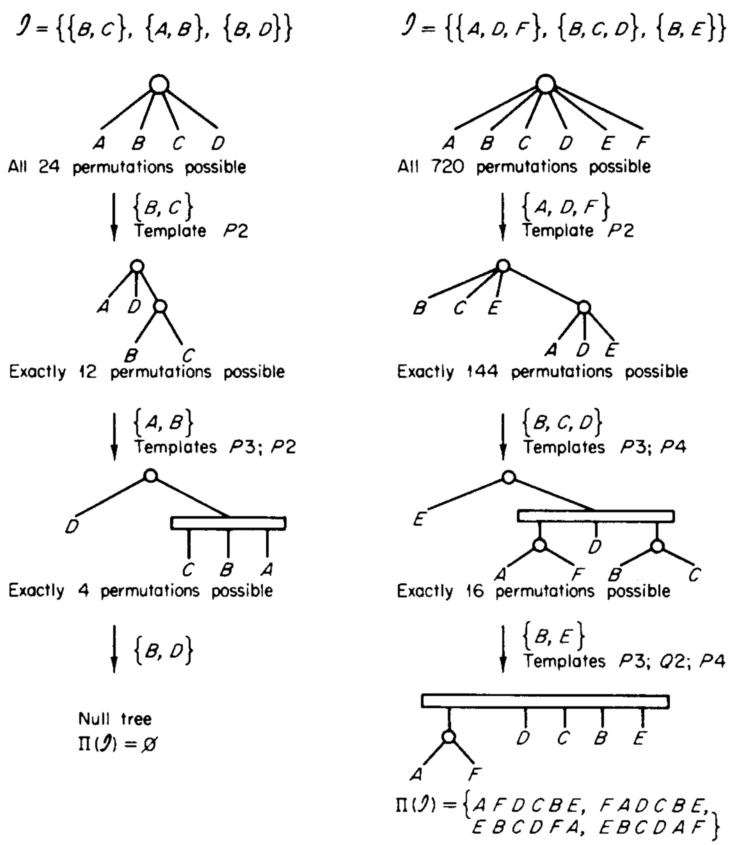

Figure 8.7. Two illustrations of the procedure CONSECUTIVE (PQ-tree version). The reductions make use of the templates in Booth and Leuker [1976].

FRONTIER(T). In the next section we suggest an application in which the permutations of $\Pi(\mathscr{I})$ may have to be compared according to secondary criteria.

Corollary 8.8. Let **M** be a (0, 1)-valued matrix with m rows, n columns, and f nonzero entries. Then, **M** can be tested for the consecutive 1's property in $O(m + n + f)$ steps.

Remark. If **M** is sparse ($f \ll mn$), then **M** would not be stored as an array. Rather, either a list of the nonzero entries or row lists of **M** would be used (Chapter 2).

$$\begin{pmatrix} 1&0&0&1&1&1\\ 0&1&1&1&1&1\\ 1&1&1&0&0&1\\ 1&0&1&0&0&1\\ 0&1&0&1&1&1\\ 0&1&1&0&1&0 \end{pmatrix} \rightarrow \begin{pmatrix} 1&0&1&0&0&1\\ 1&0&0&1&1&1\\ 0&1&0&1&1&1\\ 0&1&1&1&1&1\\ 0&1&1&0&1&0\\ 1&1&1&0&0&1 \end{pmatrix} \qquad \begin{pmatrix} 0&1&0&1&0&1\\ 1&1&0&0&1&0\\ 1&1&0&1&1&0\\ 1&1&1&1&1&0\\ 0&0&1&1&1&0\\ 0&0&0&1&1&0 \end{pmatrix}$$

$$\mathbf{M}_1 \qquad\qquad \mathbf{M}_2 \qquad\qquad \mathbf{M}_3$$

Figure 8.8. Matrix $\mathbf{M}_1$ has the circular 1's property for columns since its rows can be permuted to yield $\mathbf{M}_2$. Matrix $\mathbf{M}_3$ does not have this property.

A (0, 1)-valued matrix has the *circular 1's property* for columns if its rows can be permuted in such a way that the 1's in each column occur in a circular consecutive order; regard the matrix as wrapped around a cylinder. In Figure 8.8 the matrix $\mathbf{M}_1$ has the circular 1's property since its rows can be permuted to obtain $\mathbf{M}_2$. However, $\mathbf{M}_3$ does not have the circular 1's property.

Remark 1. $\mathbf{M}$ has circular 1's if and only if it has circular 0's.

Remark 2. The circular 1's property is preserved under *complementation* of any column, i.e., interchanging ones and zeros.

The circular 1's property was introduced by Ryser [1969].

Clearly, consecutive 1's implies circular 1's, but not conversely. Nonetheless, one can verify the latter property using a test for the former, as follows.

Let $\mathbf{M}$ be a (0, 1)-valued matrix, and let $\mathbf{M}'$ be obtained from $\mathbf{M}$ by complementing those columns with a 1 in the kth row (k chosen arbitrarily).

Theorem 8.9 (Tucker [1970, 1971]). Matrix $\mathbf{M}$ has the circular 1's property if and only if $\mathbf{M}'$ has the consecutive 1's property.

Proof. By Remark 2, if $\mathbf{M}$ has the circular 1's property, then so does $\mathbf{M}'$. By cyclically permuting the rows of $\mathbf{M}'$ so that the kth row (containing only zeros) is moved to the top, we shall obtain a matrix with consecutive 1's in each column. Conversely, if $\mathbf{M}'$ has the consecutive 1's property, then $\mathbf{M}'$ also has the circular 1's property. Hence, Remark 2 implies that $\mathbf{M}$ has the circular 1's property. ∎

The efficiency of testing for circular 1's and consecutive 1's depends partly upon the sparseness of $\mathbf{M}$. Thus, if $\mathbf{M}$ is sparse we shall want to choose k so that $\mathbf{M}'$ is also sparse. This can always be done provided $\mathbf{M}$ is stored as a list of its nonzero entries or by row lists.

Theorem 8.10 (Booth [1975]). An $m \times n$ (0, 1)-valued matrix $\mathbf{M}$ with f nonzero entries can be tested for the circular 1's property in $O(m + n + f)$ steps.

Proof. Let **M** be given as a list L of its nonzero entries. Testing for circular 1's can be accomplished as follows.

Step I. Scan L once, setting up row lists for **M** and counting the number c_i of ones in each row i: $O(m + f)$.

Step II. Choose a row k having minimum number of 1's: $O(m)$.

Step III. Form $\mathbf{M}'$ by complementing the appropriate columns. This may be carried out by scanning each row in parallel with row k, or by using an auxiliary Boolean n-vector, as illustrated in Appendix B: $O(\sum_i(c_i + c_k)) = O(m + f)$.

Remark. $\mathbf{M}'$ has at most $2f$ nonzeros since each row is at most doubled in its number of ones.

Step IV. Test $\mathbf{M}'$ for consecutive 1's: $O(m + n + 2f) = O(m + n + f)$. ∎

We have seen that testing a given matrix **M** for the consecutive or circular 1's properties can be executed efficiently. It is natural to ask, if **M** does not satisfy one or both of these properties, whether certain columns of **M** can be deleted in order that the remaining matrix satisfies the property. In general this problem is very difficult to answer.

Theorem 8.11 (Booth [1975]). Let **M** be an $r \times c$ (0, 1)-valued matrix, and let k be an integer ($k < c$). Deciding whether or not there exists an $r \times k$ submatrix of **M** satisfying the consecutive 1's property (or the circular 1's property) is NP-complete.

A proof follows from Exercise 15.

Kou [1977] presents two other extensions of the consecutive 1's property which are also NP-complete:

(1) minimizing the number of consecutive blocks of 1's appearing in the columns;

(2) minimizing the number of times a row must be split into two pieces to obtain consecutive 1's.

4. Applications of Interval Graphs

Interval graphs are among the most useful mathematical structures for modeling real world problems. The line on which the intervals rest may represent anything that is normally regarded as one dimensional. The linearity may be due to *physical restriction*, such as blemishes on a micro-organism, speed traps on a highway, or files in sequential storage in a computer. It may arise from *time dependencies* as in the case of the life span of persons or cars, or jobs on a fixed time schedule. A *cost function* may be the reason as with the approximate worth of some fine wines or the potential for growth of a portfolio of securities. And so the list goes on.

The task to be performed on an interval graph will vary from problem to problem. If what is required is to find a coloring or a maximum weighted stable set or a big clique, then fast algorithms are available. If a Hamiltonian circuit must be found, then there are no known efficient algorithms (unless the graph has more structure than just being an interval graph). Also, the speed with which such a problem can be solved will depend partially on whether we are given simply the interval graph G or, in addition, an interval representation of G.

Let us direct our attention to a few interesting applications of interval graphs.

Application 8.1. Suppose $c_1, c_2, \ldots, c_n$ are chemical compounds which must be refrigerated under closely monitored conditions. If compound c_i must be kept at a *constant* temperature between t_i and t'_i degrees, how many refrigerators will be needed to store all the compounds?

Let G be the interval graph with vertices $c_1, c_2, \ldots, c_n$ and connect two vertices whenever the temperature intervals of their corresponding compounds intersect. By the Helly property (Section 4.5), if $\{c_{i_1}, c_{i_2}, \ldots, c_{i_k}\}$ is a clique of G, then the intervals $\{[t_{i_j}, t'_{i_j}] \mid j = 1, 2, \ldots, k\}$ will have a common point of intersection, say t. A refrigerator set at a temperature of t will be suitable for storing all of them. Thus, a solution to the minimization problem will be obtained by finding a minimum clique cover of G.

Application 8.2. Benzer's problem, as stated in the introduction to this chapter, asks if the subelements inside the gene are linked together in a linear arrangement. To answer this question data were gathered on mutations of the gene. For certain microorganisms a mutant form may be assumed to arise from the standard form by alteration of some *connected* part of the internal structure. By experiment it can be determined whether or not the blemished part of two mutant genes intersect. (We would hope to show that the blemished parts are linear.)

From a large collection of mutants we obtain the pairwise intersection data of their blemishes and consider its intersection graph G. Are the intersection data compatible with the hypothesis of linearity of subelements in the gene? Equivalently, is G an interval graph? A positive answer does not confirm linearity! However, if the data are correct, a negative answer definitely refutes the hypothesis. Benzer experimented on the virus Phage T4; his findings were consistent wih linearity (see Benzer [1959, 1962] and Roberts [1976]).

Cohen, Komlós, and Mueller [1979] have shown that the asymptotic probability $P_{n,e}$ that a random graph with n vertices and e edges is an interval graph satisfies

$$P_{n,e} \sim \exp(-\lambda)$$

for large n and e and not too large e^6/n^5 where $\lambda \approx 32e^6/3n^5$. From this result and from some Monte Carlo estimates, they suggest, "it appears that the chance that Benzer observed an interval graph by chance alone is nearly zero." For related results see Cohen [1968, 1978] and Hanlon [1979a, 1979b].

The phenomenon of overlap in biology has been brought to light again recently. Kolata [1977] surveys some of these developments. She writes,

> Since the early days of molecular biology, genes have been pictured as nonoverlapping sequences of DNA [within the chromosome]. Detailed studies of a few bacterial and viral genes confirmed this view, and most investigators did not question it. [Furthermore,] the hypothesis of non-overlapping genes is a keystone for many genetic theories. [However, recent evidence seems to suggest that] viral genes and possibly bacterial genes may overlap. None of the studies with bacteria provide incontrovertible evidence that genes overlap, but all suggest that this phenomenon occurs. [If overlapping genes do exist,] current views of gene organization and the control of gene expression, as well as views of the information content of DNA molecules and the effects of mutations in DNA, may have to be substantially revised. [*Science* **176**, 1187–1188 (1977), copyright 1977 by the American Association for the Advancement of Science.]

Application 8.3. In archaeology *seriation* is the attempt to place a set of items in their proper chronological order. At the turn of the century, Flinders Petrie, a well-known archaeologist, formulated this problem, calling it "sequence dating," while studying 800 types of pottery found in 900 Egyptian graves. This problem has much in common with interval graphs and the consecutive 1's property. Let A be a set of artifacts (or aspects of artifacts) which have been discovered in various graves. To each artifact there ought to correspond a time *interval* (unknown to us) during which it was in use. To each grave there corresponds a *point* in time (also unknown) when its contents were interred. Our problem is to figure out these time relationships.

(a) Consider the incidence matrix **M** whose rows represent the graves and whose columns represent the artifacts which either are or are not present in a given grave. Under the assumption that a grave contains *every* member of A in use at the time of burial, the matrix $\mathbf{M}$ will have the consecutive 1's property for columns. Each permutation of the rows which gives consecutive 1's corresponds to an acceptable seriation of the graves and defines a possible interval assignment for A. Since there may be many of these, other methods will also have to be used to further limit the possibilities.

(b) Consider the graph G whose vertices represent the artifacts with two vertices being connected by an edge if their corresponding artifacts are found in *some* common grave. Under the assumption that every pair of artifacts whose usage intervals intersect are to be found together in some grave, we have the G is an interval graph and any interval assignment for G would be a candidate for the usage intervals of A. As before, additional techniques are required to choose the correct assignment. (See Kendall [1969a, 1969b], Hodson, Kendall, and Tăutu [1971], and Roberts [1976].) One further drawback to practical application is that there may be incomplete data so that the assumptions are not satisfied.

Application 8.4. Let X represent a set of distinct data items (*records*) and let $\mathscr{I}$ be a collection of subsets of X called *inquiries*. Can X be placed in linear sequential storage in such a way that the members of each $I \in \mathscr{I}$ are stored in consecutive locations? When this storage layout is possible, the records pertinent to any inquiry can be accessed with two parameters, a starting pointer and a length. Ghosh [1972, 1973] calls this the *consecutive retrieval* property; it is clearly a restatement of the consecutive arrangement property. Thus, the question can be answered efficiently using PQ-trees (Section 8.3). For related results see Nakano [1973a, 1973b], Ghosh [1974, 1975], Waksman and Green [1974], Patrinos and Hakimi [1976], L. T. Kou [1977], and Gupta [1979]. For an application of the circular 1's property to cyclic staffing problems, see Bartholdi *et al.* [1977].

Commentary

Application 8.5. At the Typical Institute of Mathematical Sciences (TIMS) each new faculty member visits the coffee lounge once during the first day of the semester and meets everyone who is there at the time. How can we assign the new faculty members to alcoves of the coffee lounge in such a way that no one ever meets a new person during the entire remainder of the semester? This is clearly a coloring problem on an interval graph. No specific algorithm is needed, however, since it usually happens naturally.

Additional applications of consecutive and circular 1's to such areas as file organization and cyclic staffing appear in the bibliography at the end of the chapter.

5. Preference and Indifference

Let V be a set. Let us assume that, for every pair of distinct members of V, a certain decision maker either clearly prefers one over the other or he feels indifferent about them. What is the nature of his preferences, and can they be quantified in an orderly manner? What does this imply about his decision processes?

We construct two graphs $H = (V, P)$ and $G = (V, E)$ as follows. For distinct $x, y \in V$,

$$xy \in P \Leftrightarrow x \text{ is preferred over } y,$$
$$xy \in E \Leftrightarrow \text{indifference is felt between } x \text{ and } y.$$

By definition, $H = (V, P)$ is an oriented graph, $G = (V, E)$ is an undirected graph, and $(V, P + P^{-1} + E)$ is complete.

What should we expect from the structure of H? If H has a cycle, then our decision maker is likely to be confused and is probably wasting time running around in circles.

Therefore, it is reasonable to require H to be acyclic. In fact, we would want H to be transitive. After all, if x is preferred over y and y is preferred over z, it is unlikely that a discriminating person would feel indifferent about x and z. Thus, we require that P be a partial order.*

Our example is not as whimsical as it may at first seem. The discussion above, and what will follow below, are important issues in decision theory and mathematical psychology. Analyzing how such preferences are made can enable us to understand and predict individual as well as group behavior. For example, how do we evaluate the decision making ability of a middle level corporate manager in order to determine if he is top management material?

The discipline of utility theory provides the mechanism for quantifying preference. One reasonable measure, due to Luce [1956], is the notion of a semiorder. We assign a real number $u(x)$ to each $x \in V$ so that for all x and y in V, x is preferred over y if and only if $u(x)$ is *sufficiently larger* than $u(y)$. Formally, letting $\delta > 0$, a real-valued function $u: V \to \mathbb{R}$ is called a *semiorder utility function* for a binary relation (V, P) if the following condition is satisfied:

$$xy \in P \Leftrightarrow u(x) \geq u(y) + \delta \qquad (x, y \in V). \tag{1}$$

* Krantz, Luce, Suppes, and Tversky [1971, p. 17] present an argument against transitivity of preference.

Clearly, a relation P satisfying (1) is a partial ordering of V. The quantity δ represents the amount of *fuzziness* that must be filtered out. This enables us to be indifferent about events that differ by a minuscule amount.

It is natural for us to ask the question, under what conditions does a preference relation (V, P) admit a semiorder utility function?

Theorem 8.12 (Scott and Suppes [1958]). There exists a semiorder utility function for a binary relation (V, P) if and only if the following conditions hold: For all $x, y, z, w \in V$,

(S1) P is irreflexive;
(S2) $xy \in P$ and $zw \in P$ imply $xw \in P$ or $zy \in P$.
(S3) $xy \in P$ and $yz \in P$ imply $xw \in P$ or $wz \in P$.

Such a relation P is called a *semiorder*. The conditions (S1)–(S3) constitute a set of axioms for a semiorder.* Proof of the necessity of these three conditions is straightforward and is given as Exercise 9. For the sufficiency half of the theorem, the reader is directed to the constructive proof of Rabinovitch [1977] or to the existence proofs of Scott [1964] and Suppes and Zinnes [1963].

Dean and Keller [1968] prove that the number of nonisomorphic semiorders on an n-set is $\binom{2n}{n}/(n + 1)$. In particular, they show that each isomorphism class has a unique representative, called a *normal natural* partial order (NNPO), and they then demonstrate a one-to-one correspondence between (a) the NNPOs, (b) the normal subgroups of the upper triangular group of $n \times n$ matrices, and (c) the set of nondecreasing paths from $(0, 0)$ to (n, n) on a Cartesian grid which never rises above the line $x = y$. Rabinovitch [1978] shows that every semiorder may be expressed as the intersection of at most three linear orders. No similar result holds for orders satisfying only (S1) and (S2), i.e., *interval orders*. Jamison and Lau [1973] characterize the choice functions of semiorders. They also have a good table of references. For further investigation see the works of Fishburn [1970a–1970d, 1971, 1973, 1975] and the excellent book by Roberts [1979c].

Our attention has thus far been focused on semiorders from the standpoint of the preference relation (V, P). We now investigate the indifference relation $G = (V, E)$ of our semiorder (V, P). A number of characterizations are known for these undirected graphs. First, G is a special type of interval graph (Exercise 7). Second, a necessary condition easily follows from a semiorder utility function, namely, the existence of a real-valued function u on V satisfying

$$xy \in E \Leftrightarrow |u(x) - u(y)| < \delta \qquad (x \neq y).$$

* Unfortunately, the term semiorder was used in Ghouila-Houri [1962] and later in Berge [1973] in a different context.

We will see in the next theorem that this latter condition is also sufficient. The theorem provides a number of equivalent characterizations of *indifference graphs*, which are, simply stated, the class of cocomparability graphs of semiorders. Additional characterizations appear as Exercises 10 and 11.

Theorem 8.13 (Roberts [1969]). Let $G = (V, E)$ be an undirected graph. The following conditions are equivalent.

(i) There exists a real-valued function $u: V \to \mathbb{R}$ satisfying, for all distinct vertices $x,y \in V$,

$$xy \in E \Leftrightarrow |u(x) - u(y)| < 1.$$

(ii) There exists a semiorder (V, P) such that $\bar{E} = P + P^{-1}$.

(iii) $\bar{G}$ is a comparability graph and every transitive orientation of $\bar{G} = (V, \bar{E})$ is a semiorder.

(iv) G is an interval graph containing no induced copy of $K_{1,3}$.

(v) G is a proper interval graph.

(vi) G is a unit interval graph.

Proof. (i) $\Rightarrow$ (vi) Let u be a real-valued function satisfying

$$xy \in E \Leftrightarrow |u(x) - u(y)| < 1 \qquad (x \neq y).$$

To each vertex $x \in V$ we associate the open interval $I_x = (u(x) - \frac{1}{2}, u(x) + \frac{1}{2})$. Clearly,

$$I_x \cap I_y \neq \varnothing \Leftrightarrow |u(x) - u(y)| < 1 \qquad (x \neq y).$$

Therefore, the collection $\{I_x\}_{x \in V}$ is a unit interval representation for the graph G.

(vi) $\Rightarrow$ (v) Since no unit interval can properly contain another unit interval, a unit interval representation for G will be proper.

(v) $\Rightarrow$ (iv) Let $\{I_x\}_{x \in V}$ be a proper interval representation of G. Suppose G contains an induced subgraph $G_{\{y, z_1, z_2, z_3\}}$ isomorphic to $K_{1,3}$ where $\{z_1, z_2, z_3\}$ is a stable set and y is adjacent to each z_i $(i = 1, 2, 3)$. If I_{z_j} is that interval among the intervals $I_{z_1}, I_{z_2}, I_{z_3}$ which lies entirely between the other two, then I_y must properly contain I_{z_j}, a contradiction. Thus, G can have no induced copy of $K_{1,3}$.

(iv) $\Rightarrow$ (iii) (A. A. J. Marley [unpublished].) Since G is an interval graph, its complement $\bar{G} = (V, \bar{E})$ is a comparability graph. Let F be a transitive orientation of $\bar{G}$. Using transitivity and Theorem 8.1 it is straightforward to show that F satisfies the axioms (S1) and (S2) of a semiorder (Exercise 7). We will show that (S3) also holds provided that G contains no induced copy of $K_{1,3}$. Suppose $xy \in F$ and $yz \in F$, while $xw \notin F$ and $wz \notin F$. By transitivity

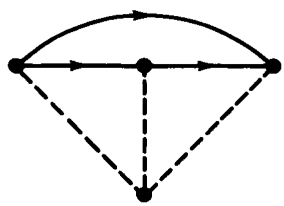

Figure 8.9. Solid edges are in the transitive orientation F of $\bar{G} = (V, \bar{E})$. Broken edges are in $G = (V, E)$.

of F, $wx \notin F$ and $zw \notin F$, and $wy \notin F$, $yw \notin F$, but $xz \in F$ (see Figure 8.9). Therefore, $G_{\{x,y,z,w\}}$ is isomorphic to $K_{1,3}$, a contradiction.

(iii) $\Rightarrow$ (ii) Immediate.

(ii) $\Rightarrow$ (i) If (V, P) is a semiorder, then there exists a real-valued function $u'\colon V \to \mathbb{R}$ and a number $\delta > 0$ such that

$$xy \in P \Leftrightarrow u'(x) - u'(y) \geq \delta.$$

Define $u(x) = u'(x)/\delta$. Since $P + P^{-1} = \bar{E}$, clearly

$$xy \in E \neq |u(x) - u(y)| < 1 \qquad (x \neq y).$$

∎

6. Circular-Arc Graphs

The intersection graphs obtained from collections of arcs on a circle are called *circular-arc graphs.* A circular-arc representation of an undirected graph G which fails to cover some point p on the circle will be topologically the same as an interval representation of G. Specifically, we can cut the circle at p and straighten it out to a line, the arcs becoming intervals. It is easy to see, therefore, that every interval graph is a circular-arc graph. The converse, however, is false. In fact, circular-arc graphs are, in general, not perfect graphs. For example, the chordless cycles C_5, C_7, C_9, ... are circular-arc graphs (see Figure 8.10).

As with interval graphs, it is immaterial whether we choose open arcs or closed arcs. The same class of intersection graphs will arise in either case (Exercise 13). We shall adopt the convention of open arcs. We call G a *proper* circular-arc graph if there exists a circular-arc representation for G in which no arc properly contains another.

In Section 1.2 we discussed an application of circular-arc graphs to the traffic light phasing problem due to Stoffers [1968]. The astute reader may well be able to adapt some of the applications of interval graphs given in Section 8.4 to the more general class of circular-arc graphs. Stahl [1967]

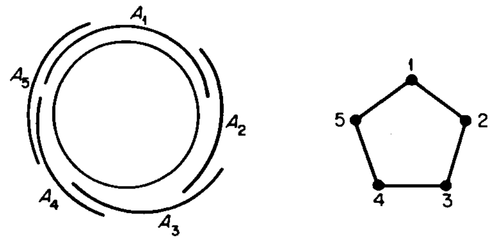

Figure 8.10. A circular-arc representation of the nonperfect graph C_5.

suggests such a problem in genetics. Other relevant papers on applications of circular-arc graphs include Luce [1971], Hubert [1974], Tucker [1978], and Trotter and Moore [1979].

A characterization of circular-arc graphs due to Tucker, originally formulated in terms of the augmented adjacency matrix of a graph, is equivalent to the following.

Theorem 8.14 (Tucker [1970b, 1971]). An undirected graph $G = (V, E)$ is a circular-arc graph if and only if its vertices can be (circularly) indexed $v_1, v_2, \ldots, v_n$ so that for all i and j

$$v_i v_j \in E \Rightarrow \begin{cases} \text{either} & v_{i+1}, \ldots, v_j \in \mathrm{Adj}(v_i) \\ \text{or} & v_{j+1}, \ldots, v_i \in \mathrm{Adj}(v_j). \end{cases} \tag{2}$$

(If $i < j$, then $v_{j+1}, \ldots, v_i$ means $v_{j+1}, \ldots, v_n, v_1, \ldots, v_i$.)

Proof. Let G have a circular-arc representation (open arcs). We may assume, without loss of generality, that no pair of arcs share a common endpoint (Exercise 14). Moving clockwise once around the circle from an arbitrary starting point, index the vertices according to the order in which the counterclockwise endpoints of their corresponding arcs occur. Let A_i denote arc corresponding to v_i. Clearly, v_i is adjacent to v_j if and only if the counterclockwise endpoint of A_j lies within A_i or vice versa. In the former case, each of $A_{i+1}, \ldots, A_j$ intersects A_i, and in the latter case each of $A_{j+1}, \ldots, A_i$ intersects A_j. Thus (2) is satisfied.

Conversely, let the vertices be indexed as required in (2). We will construct a circular-arc representation for G. Let p_k be the kth hour marker on an n-hour clock. For each vertex v_i, let v_{m_i} be the first vertex in the cyclic sequence $v_{i+1}, v_{i+2}, \ldots, v_i$ which is *not* adjacent to v_i. Draw an open arc A_i clockwise from p_i to p_{m_i}. By construction, A_i intersects A_j $(i \neq j)$ if and only if either $p_j \in A_i$ or $p_i \in A_j$. But also

$$p_j \in A_i \Leftrightarrow v_{i+1}, \ldots, v_j \in \mathrm{Adj}(v_i).$$

Therefore, by (2), $v_i v_j \in E$ if and only if $A_i \cap A_j \neq \emptyset$. ∎

Theorem 8.14 gives us a method for recognizing circular-arc graphs and constructing a circular-arc representation. However, since the characterization is quantified over all permutations of the vertices, this method will be impractical for all but very small graphs. Tucker [1978] approaches the problem of trying to find a more efficient recognition algorithm. Details of a polynomial time algorithm will appear in Tucker [1979].

In view of Theorem 8.3 it is tempting to guess that a circular-arc graph is characterized by the circular 1's property of its clique matrix or some other matrix. Unfortunately, this is not the case. Three related theorems, however, are stated here without proof.

We call G a *Helly circular-arc graph* if there exists a circular-arc representation for G which satisfies the Helly property.

Theorem 8.15 (Gavril [1974]). An undirected graph G is a Helly circular-arc graph if and only if its clique matrix has the circular 1's property for columns.

The *augmented adjacency matrix* of G is obtained from the adjacency matrix by adding 1's along the main diagonal.

Theorem 8.16 (Tucker [1970b, 1971]). An undirected graph G is a circular-arc graph if its augmented adjacency matrix has the circular 1's property for columns.

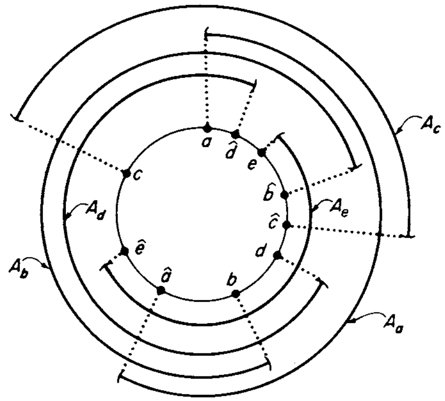

Figure 8.11. A collection of arcs of a circle with representing sequence of endpoints $\sigma = [a, \hat{d}, e, \hat{b}, \hat{c}, d, b, \hat{a}, \hat{e}, c]$.

Theorem 8.17 (Tucker [1970b, 1971]). An undirected graph G is a proper circular-arc graph if and only if its augmented adjacency matrix has the circular 1's property for columns and, for every permutation of the rows and columns that is a cyclic shift or inversion of their circular 1's order, the last 1 in the first column does not occur after the last 1 of the second column, excluding columns which are either all zeros or all ones.

It is useful to regard a collection of arcs as a sequence σ of its endpoints listed clockwise. Without loss of generality, we shall assume that no two arcs share a common endpoint (Exercise 14). In σ the symbol x denotes the counterclockwise endpoint of arc A_x and $\hat{x}$ denotes its clockwise endpoint. For example, $\sigma = [a, \hat{d}, e, \hat{b}, \hat{c}, d, b, \hat{a}, \hat{e}, c]$ represents the collection of arcs in Figure 8.11. Any cyclic permutation of σ would be an equally valid representation. The manner in which two arcs A_x and A_y intersect is uniquely determined by the pattern of the subsequence of σ involving $\{x, \hat{x}, y, \hat{y}\}$. Some examples are shown in Table 8.1. We shall utilize this model in proving the next theorem.

Theorem 8.18. If G is a proper circular-arc graph, then G has a proper circular-arc representation in which no two arcs share a common endpoint and no two arcs together cover the entire circle (i.e., they do not intersect at both ends).

Proof. The proof will be induction on the number of "circle covering" pairs of arcs. Let $\mathscr{A} = \{A_x\}_{x \in V}$ be a proper circular-arc representation of $G = (V, E)$. We may assume that no two arcs share a common endpoint. Suppose A_a and A_b cover the entire circle, that is, they intersect in two

Table 8.1

Coding a family of arcs as a sequence of letters*

Pattern of subsequence	Interpretation
$[x, \hat{x}, y, \hat{y}]$	$A_x \cap A_y = \varnothing$
$[\hat{y}, y, x, \hat{x}]$	$A_x \subset A_y$
$[x, y, \hat{x}, \hat{y}]$	A_x and A_y overlap at one end
$[x, \hat{y}, y, \hat{x}]$	A_x and A_y overlap at both ends

* Some examples of how the pattern of the subsequence of σ involving $\{x, \hat{x}, y, \hat{y}\}$ determines the manner of intersection of arcs A_x and A_y. Any cyclic permutation of a pattern leaves the interpretation unchanged.

disjoint subarcs. Let σ be the sequence of endpoints of the arcs going clockwise from the counterclockwise endpoint of A_a. Thus, $[a, \hat{b}, b, \hat{a}]$ is the subsequence of σ involving these letters, and σ may be expressed as the concatenation $\sigma = \tau\rho$, where

$$\sigma = [\underbrace{a, \ldots, \hat{b},}_{\tau} \underbrace{\ldots, b, \ldots, \hat{a}, \ldots}_{\rho}].$$

For any $x \in V$, it is impossible for x and $\hat{x}$ to appear in τ in the order $[x, \hat{x}]$ since such an appearance would imply $A_x \subseteq A_a$, contradicting the supposition that $\mathscr{A}$ is proper.

Consider the new sequence $\sigma' = \tau'\rho$, where τ' is obtained from τ by listing those entries of τ with hats followed by those without hats but preserving the relative order of each type. For any $x, y \in V$, this unshuffling operation will leave unchanged the subsequence of σ involving $\{x, \hat{x}, y, \hat{y}\}$ *unless* either $[x, \hat{y}]$ or $[y, \hat{x}]$ is a subsequence of τ. Since the cases are analogous, we assume that $[x, \hat{y}]$ is in τ. We allow the possibility that x equals a or that y equals b. Now $\hat{x}$ may either precede x or follow $\hat{a}$ in σ, and y must fall between $\hat{y}$ and b, for otherwise $\mathscr{A}$ would not be proper. This situation is depicted in Figure 8.12. Clearly, either $[\hat{x}, x, \hat{y}, y]$ or $[x, \hat{y}, y, \hat{x}]$ is a subsequence of σ, indicating that A_x and A_y overlap at both ends. After the transformation from σ or σ' occurs, these become, respectively, $[\hat{x}, \hat{y}, x, y]$ or $[\hat{y}, x, y, \hat{x}]$, which correspond to arcs which properly overlap at only one end.

Let $\mathscr{A}'$ be a set of arcs corresponding to σ'. We have just shown that (i) some doubly intersecting arcs in $\mathscr{A}$ are transformed into singly intersecting arcs in $\mathscr{A}'$, and (ii) all other pairs in $\mathscr{A}$, including nonintersecting, singly intersecting, the remaining doubly intersecting, and properly contained (of which there were none) arcs, were left unchanged. Thus, $\mathscr{A}'$ is a proper circular-arc representation of G with fewer circle covering pairs, and the theorem follows by induction. ∎

This theorem was used in Section 1.2 to show that every proper circular-arc graph is also the graph of intersecting chords of a circle. (See also Chapter 11.)

We conclude this section by presenting a polynomial-time algorithm which finds a maximum clique of a circular-arc graph. The algorithm appears in

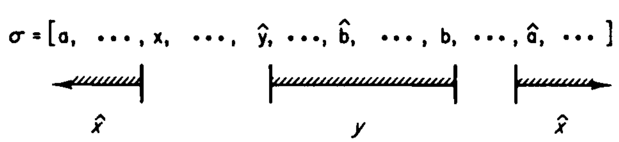

Figure 8.12. A view of where in σ the letters $\hat{x}$ and y could be hiding.

Gavril [1974] along with efficient solutions of the stable set problem and the clique cover problem for circular-arc graphs. The complexity of the coloring problem is unknown for these graphs.

Let $\{A_x\}_{x \in V}$ be a circular-arc representation of $G = (V, E)$. Assume that no two arcs share an endpoint. Denote the counterclockwise and clockwise endpoints of A_x by $\check{x}$ and $\hat{x}$, respectively. For $x \in V$, we define

$$Y_x = \{v \in V \mid \check{x} \in A_v\} + \{x\},$$
$$Z_x = \{v \in V - Y_x \mid \hat{x} \in A_v\}.$$

Each of Y_x and Z_x are complete sets, so the induced subgraph $G_{Y_x+Z_x}$ is the complement of a bipartite graph. Thus, finding a maximum clique of $G_{Y_x+Z_x}$ can be done in polynomial time.

Let K be a maximum clique of G. Choose a vertex $x \in K$ such that A_x does not properly contain any arc A_w $(w \in K)$. Hence, for every $w \in K$, $x \neq w$, either $\check{x} \in A_w$ or $\hat{x} \in A_w$. Therefore, K is a clique of $G_{Y_x+Z_x}$.

A maximum clique of G can be obtained as follows: For each $x \in V$, construct and find a maximum clique $K(x)$ of $G_{Y_x+Z_x}$; then select the largest among the $K(x)$.

EXERCISES*

1. Discuss how interval graphs and the consecutive 1's property could be applied to the following problem. Several psychological traits are to be examined in children. Assign an age period to each trait representing the natural order in the development process during which the trait is present. What traits would be appropriate for such a study?

2. Let **M** be a symmetric (0, 1)-valued matrix. Prove that either **M** has the consecutive 1's property for columns and rows or **M** has neither property. Prove the same result for circular 1's.

3. Prove that the clique matrix of an interval graph $G = (V, E)$ has at most $O(|V| + |E|)$ nonzero entries. Is this equally true for triangulated graphs (Fulkerson and Gross [1965])?

4. Let $\mathscr{I}$ be a family of intervals on a line, and let k be the maximum possible number of pairwise disjoint intervals in $\mathscr{I}$. Prove that there exist k points on the line such that each interval contains at least one of these points (T. Gallai).

5. Let **A** and **B** be (0, 1)-valued matrices having the same shape. Prove that if $\mathbf{A}^T\mathbf{A} = \mathbf{B}^T\mathbf{B}$, then either both **A** and **B** have the consecutive 1's property for columns or neither has it (Fulkerson and Gross [1965, Theorem 2.1]). A

* Also review the exercises from Chapter 1.

stronger version of this is the following: If $\mathbf{A}^T\mathbf{A} = \mathbf{B}^T\mathbf{B}$ and $\mathbf{A}$ has no subconfiguration of either of the forms below, then $\mathbf{A} = \mathbf{PB}$ for some permutation matrix $\mathbf{P}$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

(Ryser [1969, Theorem 4.1]).

6. Let P be a binary relation on a set V. A real-valued function $u: V \to \mathbb{R}$ is called an *ordinal utility function* for (V, P) if

$$xy \in P \Leftrightarrow u(x) > u(y).$$

(a) Show that (V, P) admits an ordinal utility function if and only if P is irreflexive, antisymmetric, and satisfies the negative transitivity condition (transitive indifference),

$$xy \notin P, \, yz \notin P \Rightarrow xz \notin P.$$

A (preference) relation satisfying the conditions in (a) is called a *weak order* in decision theory and a *preorder* in some mathematics literature. An ordinal utility function is like scores on an exam; this makes a weak ordering almost a total ordering (ties being allowed). Armstrong [1950] first observed that transitive indifference has important empirical shortcomings in a preference model.* To resolve these shortcomings, Luce [1956] introduced semiorders.

(b) What is the structure of the indifference graphs of weak orders?

7. Let $G = (V, E)$ be an undirected graph. Prove that the following conditions are equivalent:

(i) G is an interval graph.

(ii) $\bar{G}$ has a transitive orientation P satisfying axioms (S1) and (S2) of a semiorder.

(iii) Every transitive orientation P of $\bar{G}$ satisfies (S1) and (S2).

8. Consider the lexicographic ordering of the plane: A point (x, y) is strictly less than a point (x', y') if either $x < x'$ or both $x = x'$ and $y < y'$. Clearly for every pair of distinct points, one of them is strictly less than the other. Prove that there cannot exist a real-valued function f defined on the points of the plane which preserves the lexicographic ordering (i.e., $f(x, y) < f(x', y') \Leftrightarrow (x, y) < (x', y')$) (Debreu [1954]).

* He wrote,
The nontransitiveness of indifference must be recognized and explained on [sic] any theory of choice, and the only explanation that seems to work is based on the imperfect powers of discrimination of the human mind whereby inequalities become recognizable only when of sufficient magnitude [1950, p. 122].

9. Prove the necessity half of the Scott–Suppes theorem.

10. Prove that G is a proper interval graph if and only if its augmented adjacency matrix satisfies the consecutive 1's property for columns (Roberts [1968]). (The *augmented adjacency matrix* of G is obtained from the adjacency matrix by adding 1's along the main diagonal.)

11. We say that vertices a and b are *equivalent*, denoted $a \approx b$, if their neighborhoods $N(a)$ and $N(b)$ are equal. A vertex x is called *extreme* if $N(x)$ is complete (i.e., x is a simplicial vertex) and $[a, b \in N(x),\ a \not\approx x,\ b \not\approx x]$ implies $[\exists z \in N(a) \cap N(b),\ z \notin N(x)]$ (see Figure 8.13). Finally, let G^* be the quotient graph obtained from G by coalescing the vertices of each $\approx$-equivalence class and preserving the adjacencies between classes.

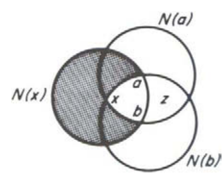

Figure 8.13. The shaded area is empty.

Prove that the following conditions are equivalent to those in Theorem 8.13 for an indifference graph $G = (V, E)$.

(vii) For every connected, induced subgraph H of G, either H^* has exactly one vertex (i.e., H is complete) or H^* has exactly two (nonadjacent) extreme vertices (Roberts [1969]).

(viii) G is triangulated and contains none of the graphs in Figure 8.14 as induced subgraph (Wegner [1967]).

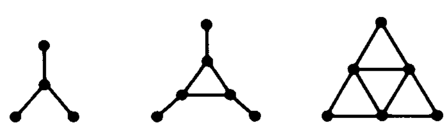

Figure 8.14. Forbidden subgraphs.

12. Let $\mathbf{M} = [m_{ij}]$ be an incidence matrix and define the row sets A_i and column sets B_j as follows:

$$A_i = \{j \mid m_{ij} = 1\}, \qquad B_j = \{i \mid m_{ij} = 1\}.$$

(a) Show that the following are equivalent.

(i) *Row intersection property*: For every i, j, k,

$$A_i \cap A_j \subseteq A_k \quad \text{or} \quad A_j \cap A_k \subseteq A_i \quad \text{or} \quad A_k \cap A_i \subseteq A_j.$$

(ii) *Column intersection property:* For every i, j, k,

$$B_i \cap B_j \subseteq B_k \quad \text{or} \quad B_j \cap B_k \subseteq B_i \quad \text{or} \quad B_k \cap B_i \subseteq B_j$$

(Nakano [1973b]).

(b) The matrix $\mathbf{M}$ is *closed* if $B_i \cap B_j \neq \varnothing$ implies $B_i \cup B_j = B_k$ for some column set B_k of $\mathbf{M}$. The closure cl($\mathbf{M}$) of $\mathbf{M}$ is defined by adding columns to $\mathbf{M}$ inductively: $\mathbf{M}^{(0)} = \mathbf{M}$; $\mathbf{M}^{(n)}$ has column sets $B_i^{(n-1)} \cup B_j^{(n-1)}$ for all $B_i^{(n-1)}$ and $B_j^{(n-1)}$ of $\mathbf{M}^{(n-1)}$ satisfying $B_i^{(n-1)} \cap B_j^{(n-1)} \neq \varnothing$.

Prove that $\mathbf{M}$ has the consecutive 1's property for columns if and only if cl($\mathbf{M}$) has the column intersection property (Nakano [1973a]).

13. Let $\mathscr{A} = \{A_x\}_{x \in V}$ be a finite collection of closed arcs of a circle. Show that there exists another collection $\mathscr{A}' = \{A'_x\}_{x \in V}$ of open arcs such that

(i) $A_x \cap A_y = \varnothing \Leftrightarrow A'_x \cap A'_y = \varnothing$,
(ii) $A_x \subseteq A_y \Leftrightarrow A'_x \subseteq A'_y$ $\quad (x, y \in V)$.

14. Let $\mathscr{A} = \{A_x\}_{x \in V}$ be a finite collection of open arcs of a circle. Show that there exists another collection $\mathscr{A}' = \{A'_x\}_{x \in V}$ of open arcs of a circle satisfying the following for all $x, y \in V$:

(i) $A_x \cap A_y = \varnothing \Leftrightarrow A'_x \cap A'_y = \varnothing$;
(ii) $A_x \subseteq A_y \Leftrightarrow A'_x \subseteq A'_y$;
(iii) no two arcs of $\mathscr{A}'$ have a common endpoint.

15. Let $\mathbf{A}$ be the $n \times m$ incidence matrix (vertices-versus-edges) of an undirected graph G.

(a) Prove that G has a Hamiltonian circuit if and only if $\mathbf{A}$ has an $n \times n$ submatrix satisfying the circular 1's property.

(b) Prove that G has a Hamiltonian path if and only if $\mathbf{A}$ has an $n \times (n - 1)$ submatrix satisfying the consecutive 1's property (Booth [1975]).

16. Give an example of a circular-arc graph whose augmented adjacency matrix does not have the circular 1's property.

17. A graph G is a *unit circular-arc graph* if there exists a circular-arc representation for G in which every arc is of unit length. (The diameter of the circle is variable.) Verify that the graph in Figure 8.15 is a proper circular-arc graph but is not a unit circular-arc graph (Tucker [1974]). (Here we assume that all arcs are open or all arcs are closed.)

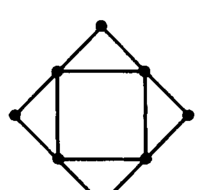

Figure 8.15.

18. Let G be a circular-arc graph. Show that $\chi(G) \leq 2\omega(G)$ (Tucker [1974]).
19. A matrix $\mathbf{M}$ is said to have the *unimodular* property if every square submatrix of $\mathbf{M}$ has determinant equal to 0, +1, or −1. (Every entry of such a matrix is necessarily 0, +1, or −1.) Show that any (0, 1)-valued matrix satisfying the consecutive 1's property is unimodular.

Research Problem. Let $\mathscr{I}$ be a collection of intervals whose intersection graph is G, and let $\mathrm{IC}(\mathscr{I})$ denote the number of different sized intervals in $\mathscr{I}$. Define the *interval count* of G, denoted $\mathrm{IC}(G)$, to be $\min\{\mathrm{IC}(\mathscr{I}) \mid \mathscr{I}$ is an interval representation of $G\}$. Clearly, $\mathrm{IC}(G) = 1$ iff G is a unit interval graph. Also, $\mathrm{IC}(K_{1,3}) = 2$.

(i) For any $k \geq 2$, characterize all graphs G with $\mathrm{IC}(G) = k$.
(ii) Find good upper and lower bounds for $\mathrm{IC}(G)$.

Leibowitz [1978] has constructed graphs of interval count k for all integers k. She has also found three classes of graphs with interval count 2, namely, trees that are interval graphs, interval graphs with a vertex whose removal leaves a unit interval graph, and threshold graphs.

Research Problem. Define the *interval number* of $G = (V, E)$, denoted $\mathrm{IN}(G)$, to be the minimum number t for which there exists a collection of sets $\mathscr{U} = \{U_v\}_{v \in V}$, where U_v is the union of t (not necessarily disjoint) intervals on the real line, such that G is the intersection graph of $\mathscr{U}$, i.e., $xy \in E$ iff $U_x \cap U_y \neq \emptyset$. Clearly, $\mathrm{IN}(G) = 1$ iff G is an interval graph. Also, any circular-arc graph has interval number at most 2.

(i) For any $k \geq 2$, characterize the graphs G with $\mathrm{IN}(G) = 2$.
(ii) Calculate the interval numbers of some special classes of graphs.
(iii) What are the best bounds for $\mathrm{IN}(G)$?

Trotter and Harary [1979] and Griggs and West [1979] have shown that the interval number of a tree is at most 2 and that

$$\mathrm{IN}(K_{m,n}) = \lceil (mn + 1)/(m + n) \rceil.$$

Griggs and West have also shown that $\mathrm{IN}(G) \leq \lceil (\delta + 1)/2 \rceil$, where δ is the maximum degree of a vertex, with equality holding for triangle-free regular graphs. Griggs [1979] has proven that $\mathrm{IN}(G) \leq \lceil (n + 1)/4 \rceil$ for all n-vertex graphs.

Bibliography

Abbott, Harvey, and Katchalski, Meir
[1979] A Turán type problem for interval graphs, *Discrete Math.* **25**, 85–88.
Armstrong, W. E.
[1950] A note on the theory of consumer's behavior, *Oxford Econ. Papers* **2**, 119–122.

Bartholdi, John J., III, Orlin, James B., and Ratliff, H. Donald

[1977] Circular ones and cyclic staffing, Tech. Report No. 21, Dept. of Oper. Res., Stanford Univ.

Berge, Claude

[1973] "Graphs and Hypergraphs," Chapter 16. North-Holland, Amsterdam. MR50 #9640.

Benzer, S.

[1959] On the topology of the genetic fine structure, *Proc. Nat. Acad. Sci. U.S.A.* **45**, 1607–1620.

[1962] The fine structure of the gene, *Sci. Amer.* **206**, 70–84.

Booth, Kellogg S.

[1975] PQ-tree algorithms, Ph.D. thesis, Univ. of California. (Also available as UCRL-51953, Lawrence Livermore Lab., Livermore, California, 1975.)

Booth, Kellogg, S., and Leuker, George S.

[1975] Linear algorithms to recognize interval graphs and test for the consecutive ones property, *Proc. 7th ACM Symp. Theory of Computing*, 255–265.

[1976] Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms, *J. Comput. Syst. Sci.* **13**, 335–379.

Cohen, Joel E.

[1968] Interval graphs and food webs: a finding and a problem, RAND Document 17696-PR.

[1978] "Food Webs and Niche Space." Princeton Univ. Press, Princeton, New Jersey.

Cohen, Joel E., Komlós, János, and Mueller, Thomas

[1979] The probability of an interval graph, and why it matters, "Proc. Symp. on Relations Between Combinatorics and Other Parts of Mathematics" (D. K. Ray-Chaudhuri, ed.). Amer. Math. Soc., Providence, Rhode Island.

Dean, Richard A., and Keller, Gordon

[1968] Natural partial orders, *Canad. J. Math.* **20**, 535–554. MR37 #1279.

Debreu, Gerard

[1954] Representation of a preference ordering by a numerical function, *in* "Decision Processes" (R. M. Thrall, C. H. Coombs, and R. L. Davis, eds.), pp. 159–165. Wiley, New York. MR16, p. 606.

Proves that there does *not* always exist a real-value order preserving function for an uncountable totally ordered set.

Eswaran, Kapali P.

[1975] Faithful representation of a family of sets by a set of intervals, *SIAM J. Comput.* **4**, 56–68. MR51 #14677.

Representation is not the intersection model. It is another use of the consecutive 1's property.

Fishburn, Peter C.

[1970a] An interval graph is not a comparability graph, *J. Combin. Theory* **8**, 442–443. MR42 #7541.

[1970b] Intransitive indifference with equal indifference intervals, *J. Math. Psych.* **7**, 144–149. MR40 #7155.

[1970c] "Utility Theory for Decision Making." Wiley, New York. MR41 #9401.

[1970d] Intransitive indifference in preference theory: A survey, *Oper. Res.* **18**, 207–228. MR41 #8050.

[1971] Betweenness, orders and interval graphs, *J. Pure Appl. Algebra* **1**, 159–178. MR47 #1689.

[1973] Interval representations for interval orders and semi-orders, *J. Math. Psych.* **10**, 91–105. MR47 #4870.

[1975] Semiorders and choice functions, *Econometrica* **43**, 975–977. MR55 #14082.

Fulkerson, D. R., and Gross, O. A.
[1964] Incidence matrices with the consecutive 1's property, *Bull. Amer. Math. Soc.* **70**, 681–684. MR32 #7444.
[1965] Incidence matrices and interval graphs, *Pacific J. Math.* **15**, 835–855. MR32 #3881.
Gavril, Fanica
[1974] Algorithms on circular-arc graphs, *Networks* **4**, 357–369. MR51 #12614.
Ghosh, Sakti P.
[1972] File organization: the consecutive retrieval property, *Comm. Assoc. Comput. Mach.* **15**, 802–808. Zbl246 #68004.
[1973] On the theory of consecutive storage of relevant records, *J. Inform. Sci.* **6**, 1–9. MR48 #1551.
[1974] File organization: consecutive storage of relevant records on a drum-type storage, *Inform. Control* **25**, 145–165. MR49 #6706.
[1975] Consecutive storage of relevant records with redundancy, *Comm. Assoc. Comput. Mach.* **18**, 464–471. MR52 #4743.
Ghouilà- Houri, Alain
[1962] Caractérisation des graphes non orientés dont on peut orienter les arrêtes de manière à obtenir le graphe d'une relation d'ordre, *C.R. Acad. Sci. Paris* **254**, 1370–1371. MR30 #2495.
Gilmore, Paul C., and Hoffman, Alan J.
[1964] A characterization of comparability graphs and of interval graphs, *Canad. J. Math.* **16**, 539–548; abstract in *Int. Congr. Math.* (Stockholm) (1962), 29 (A) MR31 #87.
Griggs, Jerrold R.
[1979] Extremal values of the interval number of a graph, II, *Discrete Math.* **28**, 37–47.
Griggs, Jerrold R., and West, Douglas B.
[1979] Extremal values of the interval number of a graph, *SIAM J. Algebraic Discrete Meth.*, to be published.
Gupta, U.
[1979] Bounds on storage for consecutive retrieval, *J. Assoc. Comput. Mach.* **26**, 28–36.
Hadwiger, H., Debrunner, H., and Klee, V.
[1964] "Combinatorial Geometry in the Plane," p. 54. Holt, New York. M22 #11310; MR29 #1577.
The problem of characterizing circular-arc graphs was posed here.
Hajös, G.
[1957] Über eine Art von Graphen, *Intern. Math. Nachr.* **11**, Problem 65.
First posed the problem of characterizing interval graphs.
Hanlon, Phil
[1979a] Counting interval graphs, submitted for publication.
[1979b] The asymptotic number of unit interval graphs, submitted for publication.
Hodson, F. R., Kendall, D. G., and Tăutu, P., eds.
[1971] "Mathematics in the Archaeological and Historical Sciences." Edinburgh Univ. Press, Edinburgh.
Hubert, Lawrence
[1974] Some applications of graph theory and related non-metric techniques to problems of approximate seriation: The case of symmetric proximity measures, *British J. Math. Statist. Psych.* **27**, 133–153.
Jamison, Dean T., and Lau, Lawrence J.
[1973] Semiorders and the theory of choice, *Econometrica* **41**, 901–912; corrections **43** (1975), 979–980. MR55 #14081.
See also Fishburn [1975].

Jean, Michel

[1969] An interval graph is a comparability graph, *J. Combin. Theory*, 189–190. MR39 #4036.
Result is false; see Fishburn [1970a].

Kendall, D. G.

[1969a] Incidence matrices, interval graphs, and seriation in archaeology, *Pacific J. Math.* **28**, 565–570. MR39 #1344.

[1969b] Some problems and methods in statistical archaeology, *World Archaeol.* **1**, 68–76.
The consecutive 1's property is applied to sequence dating.

Klee, Victor

[1969] What are the intersection graphs of arcs in a circle? *Amer. Math. Monthly* **76**, 810–813.

Kolata, Gina Bari

[1977] Overlapping genes: more than anomalies? *Science* **176**, 1187–1188.

Kotzig, Anton

[1963] Paare Hajóssche Graphen, Časopis Pěst. Mat. **88**, 236–241. MR27 #2971.
Studies bipartite interval graphs.

Kou, L. T.

[1977] Polynomial complete consecutive information retrieval problems, *SIAM J. Comput.* **6**, 67–75. MR55 #7006.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A.

[1971] "Foundation of Measurement," Vol. I. Academic Press, New York.

Leibowitz, Rochelle

[1978] Interval counts and threshold graphs, Ph.D. thesis, Rutgers Univ.

Lekkerkerker, C. G., and Boland, J. Ch.

[1962] Representation of a finite graph by a set of intervals on the real line, *Fund. Math.* **51**, 45–64. MR25 #2596.

Leuker, G. S.

[1975] Interval graph algorithms, Ph.D. thesis, Princeton Univ.

Lipski, W., Jr., and Nakano, T.

[1976/1977] A note on the consecutive 1's property (infinite case), *Comment. Math. Univ. St. Paul.* **25**, 149–152. MR55 #12531.

Luce, R. Duncan

[1956] Semiorders and a theory of utility discrimination, *Econometrica* **24**, 178–191. MR17, p. 1222.

[1971] Periodic extensive measurement, *Compositio Math.* **23**, 189–198.

Mirkin, B. G.

[1972] Description of some relations on the set of real-line intervals, *J. Math. Psych.* **9**, 243–252. MR47 #4893.

Nakano, Takeo

[1973a] A characterization of intervals; the consecutive (one's or retrieval) property, *Comment. Math. Univ. St. Paul.* **22**, 49–59. MR48 #5871.

[1973b] A remark on the consecutivity of incidence matrices, *Comment. Math. Univ. St. Paul.* **22**, 61–62. MR49 #86.

Orlin, James B.

[1979a] Circular ones and cyclic capacity scheduling, Res. Report, Dept. of Operations Research, Stanford Univ.

[1979b] Coloring periodic interval graphs, Res. Report, Dept. of Operations Research, Stanford Univ.

Patrinos, A. N., and Hakimi, S. L.

[1976] File organization with consecutive retrieval and related properties, *in* "Large Scale Dynamical Systems" (R. Sacks, ed.). Point Lobos, North Hollywood, California.

Propp, James

[1978] A greedy solution for linear programs with circular ones, I.B.M. Res. Report RC 7421.

Rabinovitch, Issie

[1977] The Scott–Suppes theorem on semiorders, *J. Math. Psych.* **15**, 209–212. MR55 #10334.

[1978] The dimension of semiorders, *J. Combin. Theory A* **25**, 50–61.

Renz, Peter L.

[1970] Intersection representations of graphs by arcs, *Pacific J. Math.* **34**, 501–510. MR42 #5839.

Roberts, Fred S.

[1968] Representations of indifference relations, Ph. D. thesis, Stanford Univ.

[1969] Indifference graphs, *in* "Proof Techniques in Graph Theory" (F. Harary, ed.), pp. 139–146. Academic Press, New York. MR40 #5488.

[1971] On the compatibility between a graph and a simple order. *J. Combin. Theory* **11**, 28–38. MR43 #7362.

[1976] "Discrete Mathematical Models, with Applications to Social, Biological and Environmental Problems." Prentice-Hall, Englewood Cliffs, New Jersey.

[1978] "Graph Theory and Its Applications to Problems of Society," NFS-CBMS Monograph No. 29. SIAM Publications, Philadelphia, Pennsylvania.

[1979a] Indifference and seriation, *Ann. N.Y. Acad. Sci.* **328**, 173–182.

[1979b] On the mobile radio frequency assignment problem and the traffic light phasing problem, *Ann. N.Y. Acad. Sci.* **319**, 466–483.

[1979c] "Measurement Theory, with Applications to Decision-Making, Utility, and the Social Sciences." Addison-Wesley, Reading, Massachusetts.

Ryser, H. J.

[1969] Combinatorial configurations, *SIAM J. Appl. Math.* **17**, 593–602. MR41 #1559.

Scott, Dana S.

[1964] Measurement structures and linear inequalities, *J. Math. Psych.* **1**, 233–247.

Scott, Dana S., and Suppes, Patrick

[1958] Foundation aspects of theories of measurement, *J. Symbolic Logic* **23**, 113–128. MR22 #6716.

Stahl, F. W.

[1967] Circular genetic maps, *J. Cell. Physiol.* Suppl. **70**, 1–12.

Stoffers, K. E.

[1968] Scheduling of traffic lights—a new approach, *Transport. Res.* **2**, 199–234.

Suppes, Patrick, and Zinnes, J.

[1963] Basic measurement theory, *in* "Handbook of Mathematical Psychology," Vol. I (R. D. Luce, R. R. Bush, and E. Galanter, eds.), pp. 1–76. Wiley, New York.

Trotter, William T., Jr., and Harary, Frank

[1979] On double and multiple interval graphs, *J. Graph Theory* **3**, 205–211.

Trotter, William T., Jr. and Moore, John, I., Jr.

[1979] Characterization problems for graph partially ordered sets, lattices and families of sets (to be published).

Tucker, Alan C.

[1970a] Characterizing the consecutive 1's property, *Proc. 2nd Chapel Hill Conf. on Combinatorial Mathematics and its Applications*, Univ. North Carolina, Chapel Hill. 472–477. MR42 #1681.

[1970b] Characterizing circular-arc graphs, *Bull. Amer. Math. Soc.* **76**, 1257–1260. MR43 #1877.
Superseded by Tucker [1971].

[1971] Matrix characterizations of circular-arc graphs, *Pacific J. Math.* **39**, 535–545. MR46 #8915.

[1972] A structure theorem for the consecutive 1's property, *J. Combin. Theory* **12**, 153–162. MR45 #4999.

[1974] Structure theorems for some circular-arc graphs, *Discrete Math.* **7**, 167–195. MR52 #203.

[1975] Coloring a family of circular-arc graphs, *SIAM J. Appl. Math.* **29**, 493–502. MR55 #10309.

[1978] Circular arc graphs: new uses and a new algorithm, *in* "Theory and Application of Graphs," Lecture Notes in Math., Vol. 642, pp. 580–589. Springer-Verlag, Berlin.

[1979] An efficient test for circular arc graphs, *SIAM J. Comput.*, to be published.

Waksman, Abraham, and Green, Milton W.

[1974] On the consecutive retrieval property in file organization, *IEEE Trans. Comput.* **C-23**, 173–174. MR50 #11886.

Wegner, G.

[1967] Eigenschaften der Nerven Homologische-einfacher Familien in R^n, Ph.D. thesis, Göttingen.