Interval Graphs, Interval Orders, and Their Generalizations

W. T. TROTTER, JR.*

Abstract. We survey some recent results on interval graphs and interval orders concentrating on algorithmic questions and generalizations involving multiple intervals, intervals in higher dimensions, intervals with tolerances, and other geometric figures. We include some open problems and discuss directions for future research.

1. Introduction. Interval graphs and their generalizations have been investigated intensively by researchers in the mathematical, social and biological sciences for more than 25 years. Understandably, a good fraction of this interest stems from the wide range of applications of interval graphs. On the other hand, the study of interval graphs has yielded a substantial body of mathematical theory of independent interest. To gain some appreciation for both aspects of the subject and for details on its history, we encourage the reader to consult the books [16], [27] by Peter Fishburn and Martin Golumbic and the special volume [28] of Discrete Mathematics edited by Golumbic. These references also provide an extensive bibliography of papers in this area.

The purpose of this article is to survey recent research on interval graphs and interval orders concentrating on algorithmic questions and generalizations involving multiple intervals, higher dimensional analogues, and other geometric figures. Our goal is to demonstrate that interval graphs and their generalizations continue to provide interesting results even as new problems arise.

---

*Department of Mathematics, Arizona State University, Tempe, AZ 85287. Research supported in part by the National Science Foundation.

In view of our earlier remarks about the scope of the research conducted on interval graphs, we confess that this article cannot include a full discussion of some interesting and important work. To those authors whose work is not mentioned, we apologize in advance.

In the interests of brevity, we provide in the article only the basic definitions and terminology necessary to understand the results and problems discussed. Additional background material is given in the books [16], [27], [61], the papers [12], [14], [24], [25], [26], [34], [36] [37], [49] [78] [79] [81] and in some recent theses [6], [43], [47], [55], [58], [64], [73] and [75].

Let $F = \{I_x : x \in V\}$ be an indexed family of nondegenerate closed[1] intervals of the real line **R**. It is natural to associate with the family F a graph G(F) and a partially ordered set P(F) each having the index set V as point set. In the graph G(F), the edges are distinct pairs xy from V for which the intervals I_x and I_y have nonempty intersection. In the partially ordered set P(F), the ordering is defined by $x < y$ when I_x and I_y are disjoint, and the right end point of I_x is less than the left end point of I_y. The graph G(F) is called an <u>interval graph</u> and the partially ordered set P(F) is called an <u>interval order</u>.

2. Algorithmic Questions. Recall that the chromatic number of a graph is at least as large as the maximum clique size. Interval graphs enjoy the special property that these two parameters are always equal. This phenomenon can be easily established by noting that interval graphs are rigid circuit graphs [11], i.e.,they do not contain induced cycles on four or more vertices. In fact, interval graphs possess additional properties which are very useful in designing algorithms for coloring. To be more precise, suppose $G = (V,E)$ is an interval graph. Choose a representation $F = \{I_x : x \in V\}$ with all end points distinct. For each $x \in V$, let $I_x = [a(x), b(x)]$. Label the vertices $v_1, v_2, \ldots, v_n$ so that $a(v_1) < a(v_2) < \ldots < a(v_n)$. If the first-fit (or greedy) algorithm is then used to color the vertices of G using this ordering of the vertices, then an optimum coloring is acheived. By this we mean that if the maximum clique size of G is m, then exactly m different colors will be used on the vertices of G.

A graph G is <u>perfect</u> if every induced subgraph of G has chromatic number <u>equal to</u> its maximum clique size. Of course, interval graphs are perfect. The fact that the

[1]We may assume wihout loss of generality that the intervals are closed as long as the family F is finite. If desired, we may also assume that all end points are distinct.

ordering by left end points allows the first-fit algorithm to produce an optimum coloring is an even stronger statement. It shows that interval graphs are <u>perfectly orderable</u> as defined by Chvatal [5], i.e. there is an ordering of the vertices on which first-fit is optimal on the graph as well as each of its induced subgraphs. Three algorithmic questions are immediate.

Q_1: Suppose the vertices of an interval graph are labelled in some arbitrary fashion (not necessarily by left end points) as $v_1, v_2, \dots, v_n$. The first-fit algorithm is then used to color the vertices in this order. If the maximum clique size of G is m, how many colors are used (The answer does not depend on n)?

Q_2: How may colors are required to color an interval graph with maximum clique size m in an on-line fashion? By "on-line," we mean that the vertices are received one at a time. When the new vertex arrives, all adjacencies to previous vertices are given. Based on this information, a color must then be assigned to this vertex and this assignment is permanent. Only then is the next vertex given.

Q_3: For which graphs is there a bounded number of permutations of the vertices so that for every induced subgraph, the first fit algorithm provides an optimum coloring when applied to one of the permutations? Of course, this question also makes sense if we replace "bounded" by "small" and replace "optimum" by "nearly optimum.

The second question was answered completely by Kierstead and Trotter [45]. They showed that there exists an on-line algorithm which will color an interval graph with maximum clique size m using 3m-2 colors. They also showed that this result was sharp. The optimum algorithm they produced is polynomial in the number of vertices; however, this algorithm is not first-fit.

It is surprising the Q_1 remains open. A. Gyarfas and J. Lehel [35] have shown that the first-fit algorithm will not use more than 3mlogm colors, but no one has been able to show that the correct answer is not linear in m. Perhaps this is another relatively innocent looking problem with solution mf(m) where f(m) is a slow growing function in the spirit of the iterated log or inverse Ackerman functions. Question 3 is wide open, although the merit of the answer will no doubt depend on what implications are obtained.

The concept of a coloring can be generalized in many ways. In [56], Opsut and Roberts discuss the reduction of some of these problems to linear programming problems. Of course, this reduction to a polynomial algorithm is only valid when the underlying graph satisfies some special properties, for instance, being an interval graph.

3. Multiple Intervals. Define the <u>interval number</u> of a graph G, denoted i(G), as the least t for wich G is the intersection graph of a family of sets with each set the union of t intervals of the real line. An interval graph has interval number 1, while the interval number of a cycle of 4 or more vertices is 2.

In [80] Trotter and Harary showed that $i(K_{m,n}) = \lceil (mn+1)/(m+n) \rceil$ and conjectured that if G is any graph on n vertices, then $i(G) \leq \lceil (n+1)/4 \rceil$. This conjecture was settled in the affirmative by J. Griggs [32] (see also [1]). The inequality is best possible as is evidenced by the complete bipartite graph with balanced sides. For multipartite graphs, Hopkins and Trotter [40] showed that if $n_1 \geq n_2 \geq \ldots \geq n_p$ then $i(K_{n_1,n_2}) \leq i(K_{n_1,n_2,\ldots,n_p}) \leq i(K_{n_1,n_2}) + 1$. Subsequently, Hopkins, Trotter and West [41] showed that $i(K_{n_1,n_2}) = i(K_{n_1,n_2,\ldots,n_p})$ unless $(n_1,n_2) = (7,5)$ or $n_1 = n_2^2 - n_2 - 1$. In these two cases, the upper bound may be obtained.

In [33], Griggs and West showed that if the maximum degree in G is d, then $i(G) \leq \lceil (d+1)/2 \rceil$. They also showed that this inequality is tight if G is regular and triangle-free. They also showed that there exists an absolute constant c so the $i(G) \leq c\sqrt{e}$ where e is the number of edges in G. The best possible value of c is not known.

In [69], Scheinerman and West showed that the interval number of a planar graph is at most 3. This result is easily seen to be best possible. Scheinerman [67] has shown that there exists an absolute constant c for which any graph of genus g has interval number at most $c\sqrt{g}$. The complete bipartite graphs show that this result is best possible up to the value of c. Perhaps it is possible to obtain an exact answer in the spirit of the Heawood map coloring formula.

To the best of my knowledge, no one has investigated algorithmic questions for graphs with bounded interval number. Even the special case $i(G) \leq 2$ would be challenging since these graphs need not be perfect. Since forests have interval number at most 2, it will however be necessary to provide additional restictions on the class if we seek an algorithm for coloring graphs with a number of colors bounded by a function of the maximum clique size. Also no one has investigated generalizations of interval orders involving multiple intervals. Here it is not clear what the appropriate definition should be.

4. Higher Dimensional Analogues. F. Roberts [59] defined the <u>boxicity</u> of a graph G, denoted Box (G), as the least t for which G is the intersection of the boxes in $\mathbf{R}^t$ (A box is the cartesian product of t nondegenerate closed intervals in R). Roberts showed that the boxicity of a graph on n vertices does not exceed $\lfloor n/2 \rfloor$ and Trotter [77] and Witsenhausen [90] completely characterized those graphs for which the inequality is tight. One such example is the complement of a matching. Recently, C. Thomassen [76] has shown that the boxicity of a planar graph does not exceed 3. In fact, he showed that intersecting boxes may be required to intersect only along a face. Cozzens and Roberts [8] obtained bounds on boxicity by studying edge coverings of $\bar{G}$ by complements of interval graphs. Cozzens and Roberts [9] obtained such bounds by relating the boxicity to the notion of a set of linear orders which forms a k-suitable of arrangements, an idea introduced by Joel Spencer [74].

Feinberg [15] defined the <u>circular dimension</u> of a graph G, here denoted cd(G), as the least t for which there exists a mapping which assigns to each vertex $x \in V(G)$ a sequence $A_x(1), A_x(2), \ldots, A_x(t)$ of arcs of the unit circle so that $xy \in E(G)$ exactly when $A_x(i) \cap A_y(i) \neq \phi$ for $i=1,2,\ldots,t$. Feinberg gave a formula for the circular dimension of a complete multipartite graph. Evidently, $cd(G) \leq Box(G)$ although if G is the complement of a matching of m edges [15], then Box(G) = m and cd(G) = 1. It is not known whether the maximum value of the circular dimension of a graph on n vertices is o(n). Other recent results about circular dimension are developed as part of a general theory of dimensional properties of graphs by Cozzens and Roberts [10]

It is a relatively straightforward calculation (see [13], [71] and [77] for example) to show that there exists an absolute constant c so that almost all labelled graphs on n vertices have interval number, boxicity, and cicular dimension all exceeding cn/logn. More sophisticated counting arguments are used by Scheinerman [65], [66] in his development of a theory of random interval graphs.

The concept of boxicity extends to partially ordered sets. Bogart and Trotter [3] defined the <u>interval dimension</u> of a partially ordered set P, denoted Idim(P), as the least t for which there exists a mapping assigning to each point $x \in P$ a sequence $[a_x(j), b_x(j)]$, $j=1,2,\ldots, t$ of non-degenerate closed intervals of R so that $x < y$ in P if and only if $b_x(j) < a_y(j)$ for $j=1,2,\ldots, t$. Bogart and Trotter [2], [3] also produced a number of inequalities for the interval dimension of a partially ordered set in terms of the cardinality and the width of certain subposets.

Scheinerman introduced a generalization of interval graphs involving intervals in posets of higher dimension. Recall that the <u>dimension</u> of a poset P is the least t for which P is the intersection of t linear extensions. When P is a poset and $a < b$ in P, define the <u>interval</u> [a,b] by $[a,b] = \{x \varepsilon P: a \leq x \leq b \text{ in } P\}$. Then define the <u>poset boxicity</u> of G, here denoted pBox(G), as the least t so that G is the intersection graph of intervals in a poset of dimension t. A graph with poset boxicity 1 is an interval graph.

The concept of poset boxicity is not very well understood at this time. Trotter and West [82] have shown that there exists an absolute constant c so that the poset boxicity of a graph on n vertices is at most cloglogn. This result is probably far from best possible. In fact,