Some Applications of Graph Theory

Fred S. Roberts∗
Department of Mathematics and DIMACS, Rutgers University
draft of October 1, 2000

Abstract
We review three graph-theoretical concepts: graph coloring, intersection graph, and competition graph. For each, we mention a variety of applications, discuss generalizations of the concept arising from the applications, and describe some recent results and open questions.

∗Fred Roberts thanks the National Science Foundation for its support under grant NSF-SBR-9709134 to Rutgers University. He thanks the Combinatorial and Computational Mathematics Center at Pohang University of Science and Technology for sponsoring the conference that led to this paper and offers his congratulations and best wishes for Com2Mac’s success.
1 Introduction

Graph theory is an old subject, but one that has many fascinating modern applications. These applications in turn have offered important stimulus to the development of the field, leading to generalizations of important graph-theoretical concepts and challenging questions about them.

We will illustrate these points with three graph-theoretical concepts: graph coloring, intersection graph, and competition graph. For each, we will mention a variety of applications, concentrating on a few; discuss generalizations related to applications; and describe a few recent results and open questions.

We will use the terminology of graph theory from the book [124]. Formally, a graph consists of a set $V$ called the set of vertices or points and a set $E$ called the set of edges and consisting of unordered pairs of vertices.

2 Graph Coloring

A coloring of a graph is an assignment of a color to each vertex so that if two vertices are joined by an edge, they get different colors. We say that a graph is $k$-colorable if it can be colored in $k$ or fewer colors. The smallest $k$ so that graph $G$ is $k$-colorable is called the chromatic number of $G$ and is denoted by $\chi(G)$. For an introduction to graph coloring, see [124].

2.1 Applications of Graph Coloring

Graph coloring has many fascinating applications. See [127] for a survey of such applications. In channel assignment, we have a set of transmitters taken to be the vertices of a graph, an edge means that the corresponding transmitters interfere, and the color assigned to a vertex is its assigned channel. The idea is that if two transmitters interfere, they get different channels. The channel assignment problem was formulated graph-theoretically in [6, 60] and has been studied widely. See [166] for an example of a recent paper on the subject.

In traffic phasing, we have a set of individuals (or cars or ...) with requests to use a facility (room, tool, traffic intersection). These are taken to be the vertices of a graph and an edge means their requests interfere. The color assigned to a vertex is the time the facility is assigned to the individual. The idea is that if two individuals have interfering requests, the individuals will be given different times to use the facility. This problem arises in phasing traffic lights, but also in a variety of other scheduling problems. An early paper formulating a graph-theoretical approach to the traffic phasing problem is by Stoffers [144]. Stoffers’ method was discussed and generalized in [106, 107, 108, 118, 120, 123].

Graph coloring also arises in scheduling meetings of legislative committees. In work connected to the legislature in the State of New York in the USA, various legislative committees are taken to be vertices of a graph and an edge
means the committees have a member in common. The color assigned to a vertex is its assigned meeting time and, of course, if two committees have a common member, they must be assigned different meeting times. This problem was studied for the New York State legislature in [9] and an exposition of it can be found in [124]. A similar problem arises in scheduling final exams at a university or jobs in a factory.

Graph coloring has also arisen in assigning schedules to garbage trucks in the City of New York. Here, the vertices of a graph represent “tours” of garbage trucks, schedules of sites they visit on a given day, an edge means they visit a common site, and the color assigned to a vertex is the day of the week the “tour” is scheduled to run. If two tours visit a common site, they must get different days. This graph coloring problem is a subproblem that arises as part of a large and complicated problem in operations research and was discussed in [151]. See also [118, 124] for an exposition of the approach.

In the fleet maintenance problem, we have a set of vehicles (planes, cars, ships) scheduled for regular maintenance. These become the vertices of a graph and an edge means that the corresponding vehicles are in the maintenance facility at overlapping times. The color gives the space assigned to a vehicle and two vehicles scheduled into the facility at overlapping times must get different spaces. This problem first arose in shipyards with the “vehicles” being ships and was studied at IBM by Alan Hoffman and Ellis Johnson (see [50]). For more general discussion of the problem, see [106, 124].

In the task assignment problem, a variety of tasks need to be performed. The tasks become vertices of a graph and an edge means that the corresponding tasks use a common tool or space or common worker. Then the color is the time assigned to the task. If two tasks use a common tool or space or worker, then they must get different times for these tasks. See [106, 124] for a discussion.

The mobile radio frequency assignment problem is concerned with assigning frequencies to mobile phones in a region. The region is divided into zones, which are the vertices of a graph, and an edge between two zones means that the phones in the two zones interfere. The color assigned to a zone is the frequency to be used by the phones in that zone, and the restriction is that if two zones interfere, they must get different frequencies. This problem was first formulated graph-theoretically by Gilbert [48]. For discussion, see [106, 120, 123].

2.2 Variants of Ordinary Graph Coloring

The problems described in the previous subsection often have complications that make ordinary graph coloring an oversimplified model. These complications have given rise to fascinating new graph coloring concepts.
2.2.1 $T$-Coloring

One complication is that channels assigned to interfering transmitters might be subject to various distance constraints. We can formalize this by talking about a set $T$ of nonnegative integers. We can think of a channel as a positive integer. Then, the idea is that interfering transmitters cannot get channels which are separated by an integer in the set $T$, thought of as a disallowed separation.

Formally, we seek a function $f$ that assigns a positive integer to each vertex of a graph $(V,E)$ so that

$$\{x,y\} \in E \rightarrow |f(x) - f(y)| \notin T.$$ 

A function $f$ satisfying this condition is called a $T$-coloring. The special case of $T = \{0\}$ is ordinary graph coloring. Another simple case is $T = \{0,1\}$. Here, interfering transmitters get not only different channels, but non-adjacent channels.

Consider for example the case of a complete graph on three vertices and the set $T = \{0,1,4,5\}$. We can color the vertices “greedily,” using the lowest acceptable positive integer for each. In that case, we would color the first vertex 1, the second 3, and the third 9. Another $T$-coloring would use the integers 1, 4, 7. These two $T$-colorings are comparable in terms of the number of colors (channels) used. However, the second is better in terms of the separation between the largest and smallest colors used. This separation is called the span of the $T$-coloring and the minimum span over all $T$-colorings is denoted $sp_T(G)$.

$T$-colorings were introduced by Hale [60] and formalized later by Cozzens and Roberts [29]. Since then, there have been dozens of papers written on $T$-colorings, and five Ph.D. theses [11, 92, 111, 146, 162]. For a survey of the literature on this topic, see [126].

There is much work to be done on $T$-colorings. For example, we don’t even know $sp_T(K_p)$ for every $T$-set and every complete graph $K_p$. ($K_p$ has $p$ vertices and all possible edges.) Cozzens and Roberts [29] showed that if $T = \{0,1,2,\ldots,r\}$, then

$$sp_T(G) = (r + 1)[\chi(G) - 1] = sp_T(K_{\chi(G)}) .$$

The result even holds if $T = \{0,1,2,\ldots,r\} \cup S$, where $S$ does not contain a multiple of $r + 1$. Such a set $T$ is called an $r$-initial set. For example, $T = \{0,1,2,5,7\}$ is 2-initial.

Raychaudhuri [111, 112] showed that if $T = \{0,s,2s,\ldots,ks\}$, then

$$sp_T(G) = st + skt - sk - 1 \text{ if } \chi(G) = st$$

and

$$sp_T(G) = st + skt + p - 1 \text{ if } \chi(G) = st + p, 0 < p < s .$$
In both cases, \( sp_T(G) = sp_T(K_{\chi(G)}) \).

Raychaudhuri showed that the same result holds if \( T = \{0, s, 2s, \ldots, ks\} \cup S \) where \( S \) is contained in \( \{s+1, s+2, \ldots, ks-1\} \). Such a set \( T \) is called a \( k \)-multiple of \( s \) set.

Another important open problem is to determine the sets \( T \) and graphs \( G \) for which an obvious greedy algorithm will obtain the optimum span. Cozzens and Roberts [29] showed that for unit interval graphs (to be defined below) and \( r \)-initial sets \( T \), the greedy algorithm obtains \( sp_T(G) \) when applied to a so-called compatible vertex ordering. Raychaudhuri [111, 112] showed that for triangulated graphs (graphs in which every cycle of length greater than 3 has a chord) and for either \( r \)-initial sets or \( k \)-multiple of \( s \) sets \( T \), the greedy algorithm obtains \( sp_T(G) \) when applied to the reverse of a so-called perfect elimination ordering.

2.2.2 \( L(2,1) \)-Coloring

Another complication in channel assignment is that transmitters might interfere at different levels. Then we might have a different distance separation requirement for each level of interference.

One formalization of this idea is to have two kinds of edges, e.g., a red edge if two transmitters are within 50 kilometers and a blue edge if they are within 100 kilometers. Then we would also have different sets \( T \) for the two colored edges. If \( \{x, y\} \) is a red edge, we would require \( |f(x) - f(y)| \notin T_1 \), and if \( \{x, y\} \) is a blue edge, we would require \( |f(x) - f(y)| \notin T_2 \).

Another formalization of this is the following: Given a graph \( G = (V,E) \), find a coloring of the vertices with positive integers so that if \( x \) and \( y \) are joined by an edge, they get colors that are not in the set \( T_1 = \{0, 1\} \), and if they have distance 2 in \( G \), then they are not in the set \( T = \{0\} \). Such an assignment of a coloring is called an \( L(2,1) \)-coloring of \( G \).

\( L(2,1) \)-colorings were introduced by Yeh [169] and Griggs and Yeh [56]. They studied the smallest \( k \) so that graph \( G \) has an \( L(2,1) \)-coloring using only integers in \( \{1, 2, \ldots, k+1\} \). This number is denoted \( \lambda(G) \). Griggs and Yeh showed that, in general, the problem of determining \( \lambda(G) \) is NP-complete. On the other hand, they showed that for trees \( T, \lambda(T) = \Delta(T) + 1 \) or \( \Delta(T) + 2 \), where \( \Delta(T) \) is the maximum degree of a vertex. Chang and Kuo [18] gave a polynomial algorithm for determining \( \lambda(T) \) for a tree \( T \). For arbitrary graphs, Chang and Kuo showed that \( \lambda(G) \leq \Delta^2(G) + \Delta(G) \). Tighter bounds on \( \lambda \) in terms of \( \Delta \) for special families of graphs are established in [18, 46, 47, 56, 69, 133, 167, 169]. Griggs and Yeh conjectured that \( \lambda(G) \leq \Delta^2(G) \). This conjecture remains open.

More recent work on \( L(2,1) \)-colorings is found in the papers [36, 37] and these papers provide a variety of references to the literature. They leave open in general the question of characterizing graphs \( G \) such that \( \lambda(H) < \lambda(G) \) for all proper subgraphs \( H \) of \( G \) and the problem of characterizing graphs \( G \) for which there is an \( L(2,1) \)-coloring using integers in \( \{1, 2, \ldots, \lambda(G) + 1\} \) such that
all these integers are used on some vertex.

2.2.3 Set Colorings; \(n\)-Tuple Colorings

In all of the problems we have mentioned, it makes sense to speak of assigning a set of colors to a vertex rather than a single color. If \(S(x)\) is the set of colors assigned to vertex \(x\), then it is natural to require that

\[
\{x, y\} \in E \rightarrow S(x) \cap S(y) = \emptyset.
\]

Such an assignment is called a set coloring. \(|S(x)| = 1\) for all \(x\) gives an ordinary graph coloring. If each set \(S(x)\) has exactly \(n\) colors, we call this an \(n\)-tuple coloring. The smallest number of colors required for an \(n\)-tuple coloring of \(G\) is called the \(n\)-tuple chromatic number of \(G\) and is denoted by \(\chi_n(G)\). Consider for example the 4-cycle. If we use the color sets \{red, blue\} and \{green, yellow\} alternating around the cycle, then we get a 2-tuple coloring. This shows that \(\chi_2(C_4) \leq 4\). Of course, it equals 4. \(n\)-tuple colorings were introduced by Gilbert [48] in connection with the mobile radio frequency assignment problem.

Here is a useful early result about \(n\)-tuple colorings. Given graphs \(G\) and \(H\), the lexicographic product \(G[H]\) is defined as follows: \(V(G[H]) = V(G) \times V(H)\) and \((a, b)\) is joined to \((c, d)\) by an edge if and only if \(\{a, c\} \in E(G)\) or \(a = c\) and \(\{b, d\} \in E(H)\). Stahl [142] proved that \(\chi_n(G) = \chi(G[K_n])\). This result is often more useful theoretically than practically – just try to compute \(\chi_2(C_4)\) this way! Theoretically, for example, the result is used to prove that if \(G\) is a weakly \(\gamma\)-perfect graph, then \(\chi_n(G) = n\chi(G)\). (\(G\) is weakly \(\gamma\)-perfect if its chromatic number is equal to the size of its largest clique, i.e., the largest complete subgraph.) This holds for \(G = C_p\) for \(p\) even. For example, we know that \(\chi_2(C_4) = 4 = 2\chi(C_4)\).

To show how difficult \(n\)-tuple coloring can be, we consider the Kneser graph \(G(m, p)\): The vertices are the \(p\)-element subsets of an \(m\)-element set. There is an edge between two subsets if and only if they are disjoint. A sticky open question in the theory of \(n\)-tuple colorings is to find \(\chi_n(G(m, p))\). This is important because ordinary colorings are “homomorphisms” into complete graphs while \(n\)-tuple colorings are homomorphisms into Kneser graphs. Lovász [93] calculated \(\chi(G(m, p)) = \chi_1(G(m, p))\) in 1978 in the process of settling an important open problem in graph theory by proving Kneser’s conjecture. (Kneser’s conjecture says: If we split the \(p\)-element subsets of a \((2p + k)\)-element set into \(k + 1\) classes, one of the classes will contain two disjoint \(p\)-element subsets.)

A formula for \(\chi_n(G(m, p))\) was conjectured by Stahl [142]: If \(n - 1 = qp + r, q \geq 0, 0 \leq r < p\), then

\[
\chi_n(G(m, p)) = (q + 1)m - 2(p - r - 1).
\]

This formula is known to hold for a variety of values of \(n, p, m\). See for example [40, 44, 142]. However, its truth or falsity remains open in general.
Many sets other than sets of \( n \) integers have been studied in connection with set colorings. Among the most important are sets consisting of real intervals or unions of real intervals. These are discussed briefly below, in Sec. 3.5. See [107, 127] for a variety of concrete cases and references.

### 2.2.4 List Coloring

In many applications, there is an extra complication. There are some acceptable colors for each vertex and the color assigned to a vertex must be chosen from the set of acceptable colors. For instance, in channel assignment, we specify a set of acceptable channels and in traffic phasing a set of acceptable times.

Given a graph \( G \), let \( L(x) \) be a non-empty set of integers assigned to vertex \( x \). We call \( L \) a list assignment for \( G \). We seek a graph coloring \( f \) so that for every vertex \( x \), \( f(x) \) is in \( L(x) \). Such a coloring is called a list coloring for \( (G, L) \).

If a list coloring exists, we say that \( (G, L) \) is list colorable. List colorings were introduced in the 1970's, independently by Vizing [156] and by Erdős, Rubin, and Taylor [32], and there have been a very large number of papers about this subject in the past decade. Some recent survey articles are [3, 88, 155].

A great deal of emphasis has been placed on the case where each set \( L(x) \) has the same fixed number of elements, \( k \). If \( (G, L) \) can be list colored for every possible list assignment \( L \) in which all \( |L(x)| = k \), we say that \( G \) is \( k \)-choosable.

To give an example, we note that \( C_5 \) is not 2-choosable. Use \( L(x) = \{1, 2\} \) on each vertex. If there is a list coloring for this list assignment, then \( C_5 \) has an ordinary coloring using two colors. On the other hand, it is easy to show that \( C_4 \) is 2-choosable.

Erdős, Rubin, and Taylor [32] characterized 2-choosable graphs. However, the characterization of 3-choosable graphs remains a wide open problem.

There have been major results on choosability in the past ten years. For example, Erdős, Rubin, and Taylor [32] conjectured that every planar graph is 5-choosable but that there are planar graphs that are not 4-choosable. Alon and Tarsi [4] proved that bipartite planar graphs are 3-choosable. Voigt [157] showed that not all planar graphs are 4-choosable. Thomassen [147] proved that every planar graph is 5-choosable. In spite of this progress, many open questions about list coloring and other generalizations of ordinary graph coloring remain. See [67] for many such problems.

### 3 The Second Concept: Intersection Graph

Let \( F = \{S_2, S_3, \ldots, S_n\} \) be a family of sets. We can build a graph corresponding to \( F \) by taking the vertex set to be \( F \) and including an edge between \( S_i \) and \( S_j \) if and only if \( S_i \cap S_j \neq \emptyset \). This graph is called the intersection graph of the family of sets. For instance, consider the sets...
Then in the intersection graph, the edges join $S_2$ to the other vertices.

It is natural to ask: What graphs arise this way? Formally put: Given a graph $G = (V, E)$, can we assign a set $S(x)$ to each vertex $x$ of $V$ so that for all $x \neq y$,

\[ \{x, y\} \in E \iff S(x) \cap S(y) \neq \emptyset \]

Marczewski [102] proved that very graph is the intersection graph of some family of sets.

3.1 Interval Graphs

The question of what graphs arise as intersection graphs is much more interesting and leads to really important ideas if we restrict the families of sets to sets of certain kinds. We shall concentrate on one important special case. If $F$ is a family of intervals on the real line, then its intersection graph is called an interval graph. We shall return to the question of what graphs are interval graphs. First, we note that this one very simple concept has a large number of fascinating applications. For general references on interval graphs and their applications, see [34, 50, 118, 120, 148].

3.2 Applications of Interval Graphs

A variant of the following problem was one of the motivating problems for the notion of interval graph (Hajós [59]). A group of students goes to the library and later a book is missing. To find out who might have taken it, we try to reconstruct when people were there. We assume that each student stays in the library for a certain interval of time, then leaves. For each pair of students $x$ and $y$, we know whether or not $x$ and $y$ saw each other. Can we construct time intervals so that if $x$ and $y$ saw each other, their time intervals in the library overlap, and if they didn’t see each other, then their time intervals in the library do not overlap? We define a graph $G$ by letting the vertices be the individuals and an edge between two individuals mean that they saw each other. Then our question is equivalent to the question: Is $G$ an interval graph? Even if the answer is yes, we need a way to find the intervals. There are good algorithms for doing so, as we shall note below. Also, even finding the intervals is just a first step in solving the mystery. However, it is an important first step.

Interval graphs arose independently from Hajós’ problem and from a problem in molecular genetics known as Benzer’s Problem [7, 8]. Seymour Benzer was a Nobel Prize-winning geneticist. In 1959, he was studying the “fine structure” of bacterial genes; it was not known whether or not the collection of DNA
composing a bacterial gene was linear. He asked: Is the fine structure of the gene linear or is it circular or does it have a different topology? How do we tell if we can’t see it? At that time, there were no sophisticated methods (such as gel electrophoresis) for answering questions like this. Benzer observed that we can determine whether or not two connected substructures inside the gene overlap. We do it by gathering mutation data. He then asked: Is the overlap information consistent with a linear structure? Equivalently, the question can be stated this way: Can we assign an “interval” to each substructure so that two substructures overlap if and only if their intervals do? Or: Is the graph of overlaps among substructures an interval graph? There is no longer active interest in Benzer’s problem, but there is a great deal of modern work in molecular biology that involves interval graphs. For example, interval graphs arise in connection with research on restriction maps which show the location of certain sites (short specific sequences) on a specific DNA ([125, 163, 165]). More generally, interval graphs also play an important role in the whole study of physical mapping in molecular biology. See [57, 136, 164] for a general introduction to this problem and [1, 2, 10, 68, 71, 138, 139, 170] for examples of research papers involving interval graphs in physical mapping.

Interval graphs also arise in the study of preference and indifference in economics and psychology. Suppose we consider some alternative purchases about which we do not have a good estimate of the value, such as antiques for example. We can associate with each alternative \( u \) a range of possible values \( J(u) \). We would be comfortable in saying that we prefer \( u \) to \( v \) if the interval \( J(u) \) is strictly to the right of the interval \( J(v) \), and otherwise not. Suppose that we are indifferent between \( u \) and \( v \) if and only if we neither prefer \( u \) to \( v \) nor \( v \) to \( u \). Then we are indifferent between \( u \) and \( v \) if and only if \( J(u) \) and \( J(v) \) overlap. Let \( V \) be the set of alternatives under consideration and define a graph on \( V \) by letting edges correspond to indifference. If observed indifference judgments fit our assumptions, then this graph is an interval graph. For further discussion of interval graphs as models of indifference, see [116, 118, 122].

Seriation or sequence dating is an important area in in archaeology. Here, we study a variety of types of artifacts (e.g., pottery) dug up in archaeological digs. We would like to place them in chronological order, assuming each type of artifact was in use over an interval of time. Let us build a graph with the types of artifacts as vertices. Suppose that an edge means that they were found in common in some dig. If they were found in common in some dig, then it is reasonable to conclude that the time intervals during which the two artifacts were in use overlapped. Assuming enough digs, then we can assume that we have two types of artifacts appearing in common in some dig if and only if their time intervals overlapped. Thus, the graph is an interval graph and we can use the corresponding assignment of intervals to find a preliminary chronological order. For more on seriation in archaeology, see [64, 73, 74, 75, 76, 77, 78, 118, 120, 121, 140].

An analogous problem arises in seriation in developmental psychology. In
studying the development of children, psychologists have noted that traits such as crawling, sitting up, etc. each develop over a certain interval of time over the course of development of the child. There is a stage when a child crawls, a stage that overlaps when the child is starting to sit up, a stage when the child begins to pull itself up to a standing position, etc. It is hypothesized that the pattern of development is common to all children. In studying this hypothesis, we can build a graph whose vertices are the traits and which has an edge between two traits if they are found in common in some child. Assuming we have studied enough children, we expect that two traits were found in common in some child if and only if they developed in overlapping intervals of time. Thus, the graph is an interval graph and we can use the corresponding assignment of intervals to find a preliminary developmental order. For more on seriation in developmental psychology, see [26, 118, 120].

3.3 What Graphs are Interval Graphs?

There are various theorems that allow us to tell if a graph is an interval graph. Early results were obtained by Lekkerkerker and Boland [91], Gilmore and Hoffman [49], and Fulkerson and Gross [42], and a polynomial algorithm was developed by Booth and Lueker [12].

Here is one sample result. A clique in a graph G is a complete subgraph and we say it is maximal if it is not contained in any larger clique. We define the maximal clique-vertex incidence matrix as follows. The rows correspond to the maximal cliques $K_1, K_2, \ldots, K_p$ and the columns to the vertices $x_1, x_2, \ldots, x_n$, with the $i, j$ entry equal to 1 if $K_i$ contains $x_j$ and equal to 0 otherwise.

Consider for example the graph $H = (V, E)$ where

$$V = \{a, b, c, d, e, f, g\}$$
$$E = \{\{a, b\}, \{b, c\}, \{a, c\}, \{d, e\}, \{f, g\}, \{e, f\}\}.$$

Then the maximal cliques are $K_1 = \{a, b, c\}, K_2 = \{d, e\}, K_3 = \{f, g\}, K_4 = \{e, f\}$ and the maximal-clique vertex incidence matrix is the following matrix:

$$\begin{pmatrix}
a & b & c & d & e & f & g \\
K_1 = abc & 1 & 1 & 1 & 0 & 0 & 0 \\
K_2 = de & 0 & 0 & 0 & 1 & 1 & 0 \\
K_3 = fg & 0 & 0 & 0 & 0 & 1 & 1 \\
K_4 = ef & 0 & 0 & 0 & 1 & 1 & 0
\end{pmatrix}$$

In the 4-cycle $C_4$ with vertices $a, b, c, d$ in order around the cycle, we have the maximal-clique vertex incidence matrix
Fulkerson and Gross [42] showed that a graph is an interval graph if and only if the rows of the maximal clique-vertex incidence matrix can be reordered so that the 1’s in each column appear consecutively. We see that this can be done with the graph $H$ defined above. If we reorder the maximal cliques, we get the following matrix:

$$
\begin{pmatrix}
 a & b & c & d \\
 ab & 1 & 1 & 0 & 0 \\
 bc & 0 & 1 & 1 & 0 \\
 cd & 0 & 0 & 1 & 1 \\
 ad & 1 & 0 & 0 & 1 \\
\end{pmatrix}
$$

However, there is no way to reorder the rows of the maximal-clique vertex incidence matrix for $C_4$ in order to accomplish the same goal. That shows that $C_4$ is not an interval graph, as is easy to demonstrate directly. A matrix whose rows can be permuted so that the 1’s in each column appear consecutively is said to have the consecutive 1’s property (for columns). This property is very useful in a variety of applications. See [28] for a discussion.

An ordering of maximal cliques for which the maximal clique-vertex incidence matrix has 1’s appearing consecutively in each column is called a consecutive ordering. We shall see that consecutive orderings are very important in applications.

### 3.4 Circular-Arc Graphs

Other families of intersection graphs are also of interest. We say that $G$ is a circular-arc graph if it is the intersection graph of a family of arcs on a circle. While $C_4$ is not an interval graph, it is easy to see that it is a circular-arc graph.

It seems natural to conjecture the following: A graph is a circular-arc graph if and only if the rows of the maximal clique-vertex incidence matrix can be reordered so that the 1’s in each column appear consecutively in a circular sense (i.e., continuing from bottom to top). However, this is not the case. The graph $\Gamma$ consisting of a triangle with a pendant edge added to each vertex is a circular-arc graph. However, there is no way to reorder rows of the maximal clique-vertex incidence matrix so that the 1’s in each column appear consecutively in a circular sense. ($\Gamma$ has vertices $a, b, c, x, y, z$ with $a, b, c$ forming a triangle and edges from $x$ to $a$, $y$ to $b$, and $z$ to $c$.) Technically speaking, one needs to add a condition that the family of circular arcs has the Helly property. Any subfamily of arcs
that pairwise overlap has a common point. Then, we can conclude that a graph is the intersection graph of a family of circular arcs with the Helly property if and only if we can reorder the rows of the maximal clique-vertex incidence matrix so that the 1’s in each column appear consecutively in a circular sense. Any family of circular arcs whose intersection graph is the graph \( \Gamma \) defined above would have circular arcs corresponding to \( a, b, c \) pairwise overlapping, but no one point belonging to all three such arcs. Thus, \( \Gamma \) is not the intersection graph of a family of circular arcs with the Helly property.

The problem of recognizing or characterizing circular-arc graphs is difficult. There is a long literature on this problem. Finally, good algorithms were described for testing for whether not a graph is a circular-arc graph. However, structural understanding of these graphs remains difficult to obtain. For work on the circular-arc graph recognition problem, see for example [33, 87, 141, 149, 150, 152, 153, 154]

### 3.5 Connection to Traffic Phasing

The notions of circular-arc graph and interval graph have applications in traffic phasing. We have different streams of traffic approaching an intersection. We would like to put in a traffic light. The light will give a green signal to each traffic stream over an interval of time. (Similar problems arise in scheduling other facilities such as computers, rooms, etc.)

We can think of a large circular clock and the traffic streams each receiving an arc of time along that clock for their green time. Some traffic streams are compatible and some are not. We build an incompatibility graph by taking the vertices to be the traffic streams and an edge between two traffic streams if they are incompatible. A green light assignment is called feasible if incompatible traffic streams get non-overlapping circular arcs. If \((V, E)\) is the incompatibility graph, then if \(J(x)\) is a circular arc assigned to traffic stream \(x\) in a feasible green light assignment,

\[
\{x, y\} \in E \rightarrow J(x) \cap J(y) = \emptyset.
\]

If we think of each \(J(x)\) as a set of colors or times assigned to a vertex \(x\), we have a set coloring problem, more specifically, what is sometimes called an interval coloring problem.

It is useful to consider the compatibility graph instead, the graph \((V, E')\) whose edges join compatible traffic streams. For this graph, we have:

\[
\{x, y\} \in E' \leftarrow J(x) \cap J(y) \neq \emptyset.
\]

Note that we do not necessarily have \(\leftarrow\). The set of all edges in \(E'\) corresponding to \(x, y\) so that \(J(x) \cap J(y) \neq \emptyset\) is a subset of \(E'\). Thus, a feasible assignment defines a so-called spanning subgraph \(H\) of the compatibility graph (spanning meaning that the subgraph has the same vertex set and a subset of the edges).
Moreover, since the $J(x)$’s are circular arcs, $H$ is a circular-arc graph. Thus, feasible traffic light assignments correspond to spanning subgraphs of the compatibility graph that are circular-arc graphs. Unfortunately, circular-arc graphs are not easy to identify.

Observe that if the last green light ends before the first starts, then $H$ is an interval graph. Real traffic lights have this property. Thus, feasible traffic light assignments correspond to spanning subgraphs of the compatibility graph that are interval graphs. Fortunately, interval graphs are easy to identify.

Consider for example a traffic intersection with eastbound traffic stream $a$, westbound traffic stream $b$, northbound traffic stream $c$, southbound traffic stream $d$, left-turning traffic stream $e$ going from south to west, and left-turning traffic stream $f$ going from north to east. (This intersection is studied in [118].) A compatibility graph is given by taking

$$V = \{a, b, c, d, e, f\}$$

$$E = \{\{a, b\}, \{c, d\}, \{d, f\}, \{e, f\}, \{c, e\}\}.$$  

Note that this is not a circular-arc graph or an interval graph. However, omitting edge $\{c, e\}$ gives us an interval graph spanning subgraph $H$ and this corresponds to a feasible green light assignment, which can be found by an assignment of intervals (or circular arcs) whose intersection graph is this spanning subgraph. One such assignment is the following:

$$J(a) = (0, 1), J(b) = (0, 1), J(c) = (1, 3), J(d) = (1, 6), J(e) = (6, 10), J(f) = (3, 10).$$

Of all feasible green light assignments, which ones are the best or most efficient? We can make each green light interval very short, thus almost surely obtaining a feasible assignment. However, we usually specify that each green light interval have a certain minimum length. If the traffic light is at an isolated intersection with no other traffic lights nearby, then it is reasonable to try to find an assignment that minimizes the sum total of waiting times. If the intersection is not isolated, vehicles can be expected to arrive at the intersection at given times. Then, we try to minimize the (weighted) time lags between ideal and realized starting times.

We concentrate on the isolated traffic intersection. A good procedure was developed by traffic scientist Karl Stoffers [144]. See [118] for an exposition of this procedure. Stoffers’ idea was:

- 1. Find all spanning subgraphs of the compatibility graph that are interval graphs.
- 2. For each such spanning subgraph, find the most efficient feasible green light assignment.
• 3. Find the overall most efficient assignment by comparing the solutions in step 2.

How do we find the solution in step 2? We use the Fulkerson-Gross characterization of interval graphs given in Sec. 3.3.

• 2a. Find a consecutive ordering of maximal cliques, i.e., an ordering so that the maximal clique-vertex incidence matrix has consecutive 1’s: \(K_1, K_2, \ldots, K_p\). Each clique corresponds to a “phase” during which all of the traffic in that clique gets a green light.

• 2b. For each vertex \(u\), let it get green light in the first \(K_i\) that contains it and continue until the last \(K_i\) that contains it. Since the matrix has consecutive 1’s, each vertex is in a consecutive set of maximal cliques and so we end up with an interval of green light time.

In the specific example given above, the maximal clique-vertex incidence matrix with rows permuted to give consecutive 1’s is the following:

\[
\begin{pmatrix}
 a & b & c & d & e & f \\
 K_1 = ab & 1 & 1 & 0 & 0 & 0 \\
 K_2 = cd & 0 & 0 & 1 & 1 & 0 \\
 K_3 = df & 0 & 0 & 0 & 1 & 1 \\
 K_4 = ef & 0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

\(K_1, K_2, K_3, K_4\) is a consecutive ordering and corresponds to one phasing. In this phasing, we start with a green light for streams \(a\) and \(b\), the eastbound and westbound traffic. We then have a green light for the northbound and southbound traffic. In the third phase, we turn off the green light for the northbound traffic and turn on a green light for left-turning traffic from north to east. Finally, in the fourth phase, we have both left turn arrows on.

Another consecutive ordering is \(K_1, K_4, K_3, K_2\). Here, we start with the east-west traffic, then turn on both left-turn arrows, then replace the left-turning traffic from south to west with southbound traffic, and finish with north-south traffic.

A different interval graph spanning subgraph would give some entirely different phasings. The reader might wish to experiment to see what phasings can be found if instead of omitting edge \(\{c, e\}\) from the compatibility graph, we omit instead edge \(\{c, d\}\).

The rest of the solution requires us to find the lengths of the intervals (circular arcs).

• 2c. Assign maximal clique \(K_i\) a duration \(d_i\).

• 2d. How do we find the values \(d_i\) that minimize the total waiting time? This is a linear programming problem. It was formulated in some generality by Opsut and Roberts [107, 108].
Note that the method is not at all efficient in a technical sense. Even the problem of finding all maximal cliques of a graph is an NP-complete problem. However, the graphs to which the method is applied are relatively small, so the method can be applied in practice in a reasonably efficient way.

For more general discussion of the above procedure, see [106, 107, 108, 123].

### 3.6 Unit Interval Graphs

The special case of interval graphs where every interval has the same unit length is an important case. Such graphs, called unit interval graphs, were characterized in [115, 116] and have found numerous applications in philosophy of science, preference and utility measurement, molecular biology, etc. See [50, 120, 122] for examples of some of these applications.

### 3.7 Unions of Two Intervals

A variant of the traffic phasing problem is to allow green light assignments where each “green light set” is a union of two intervals. This problem is difficult in part because it is even an open problem to characterize the double interval graphs, intersection graphs of families of sets each of which is a union of at most two intervals. For some results about double interval graphs, see [55, 61, 111, 113].

### 3.8 Intersection Graphs of “Boxes” in Euclidean Space

Suppose we have boxes in Euclidean $k$-space, generalized rectangles with sides parallel to the coordinate axes. In 2-space, these are rectangles. What graphs arise as the intersection graphs of boxes in $k$-space? This is in general an unsolved problem, even for $k = 2$, and is known to be NP-complete for $k > 2$ [27, 168]. For $k = 1$, boxes reduce to intervals, and the intersection graphs of boxes are exactly the interval graphs.

It is easy to see that every graph is the intersection graph of boxes in $k$-space for some sufficiently large $k$. The smallest such $k$ is called the boxicity of the graph. This concept was introduced in [115, 117]. It is still fundamentally an open problem to understand the boxicity of large classes of graphs, even though this concept goes back to 1967.

Intersection graphs of boxes have a variety of important applications. One of the major applications is in ecology and we turn to this as we discuss the third basic concept of emphasis in this paper, that of competition graph.

### 4 Third Concept: Competition Graph

Suppose $D = (V, A)$ is a digraph. The competition graph $C(D)$ has vertex set $V$ and an edge between $x$ and $y$ if there is $u$ in $V$ so that $(x, u)$ and $(y, u)$ are arcs
of $A$. This notion was introduced by Joel Cohen in 1968 [19]. We say that $G$ is a competition graph if it is $C(D)$ for some $D$. In some cases, we require that $D$ be acyclic, but we shall not make that assumption here except in certain cases. Since Cohen’s work, a large literature on the subject of competition graphs has arisen. This literature is surveyed in [80, 95, 125, 128].

4.1 Applications of Competition Graphs

Competition graphs arose in the work of Cohen [19] in connection with an application in ecology. Here, $D = (V,A)$ represents a food web for an ecosystem. $V$ is the set of species in the ecosystem and $(x,y) \in A$ if $x$ preys on $y$. Then $x,y$ is an edge of $C(D)$ iff $x$ and $y$ compete for a common prey $u$. Consider for example the following food web $F$ (part of a food web in [16]). The species are Bear (B), Deer (D), Fox (F), Grass (G), Hare (H), Owl (O), Raccoon (R), Shrew (S), and Wildcat (W) and the predator-prey relations are given by the ordered pairs $(B,D)$, $(B,G)$, $(B,S)$, $(O,G)$, $(D,G)$, $(F,O)$, $(F,H)$, $(F,S)$, $(H,G)$, $(R,O)$, $(W,O)$, $(W,S)$. The Foxes and Wildcats compete because they both eat owls, and so there is an edge $\{F,W\}$ in the competition graph. The ecological application of competition graphs was a primary motivation for the paper [20], the book [21] and a large number of papers on the topic.

A second application of competition graphs is in communication over a noisy channel. Here, $V$ is the set of letters in an alphabet and $(x,y) \in A$ if when $x$ is sent, $y$ could be received. Then $x,y$ is an edge of $C(D)$ iff $x$ and $y$ could be received as the same letter. $C(D)$ is the confusion graph. This graph arises in Shannon’s theory of noisy channels and in his definition of the capacity of such a channel (see for example [58, 94, 104, 120, 137]). These concepts give rise to challenging open questions in graph theory. For example, it took a long time to calculate the so-called Shannon capacity of the simple confusion graph $C_5$ (see [94]), but we still don’t know the capacity of all cycles.

Competition graphs also arise in channel assignment. Here, $V$ is a set of transmitting/receiving stations and $(x,y) \in A$ if a signal sent at $x$ can be received at $y$. Then $x,y$ is an edge of $C(D)$ iff messages sent by $x$ and $y$ can be received at the same place. $C(D)$ is the conflict graph. The channel assignment problem we have discussed before in Sec. 2.1 is concerned with coloring this graph. The channel assignment application has given rise to a lot of work on competition graphs of undirected graphs. See [98, 99, 101, 114].

A fourth application of competition graphs is in modeling of complex systems. In large-scale computer models, for instance those dealing with energy or economic systems, we often deal with a matrix $M$ representing the constraints of an LP or a similar problem. Here, we can take $V$ to be the set of rows of $M$ and $(x,y) \in A$ if $M(x,y) \neq 0$. Then $x,y$ is an edge of $C(D)$ iff rows $x$ and $y$ have a nonzero entry in the same column. They “link” a common variable. $C(D)$ is known as the row graph of $M$. The row graph has many uses, including applications to the structure of linear programs. The properties of row graphs are
studied in [51, 52, 53, 54] and applications to the structure of linear programs are discussed in [51, 109, 110].

The problem of phylogenetic tree reconstruction is the following: Given a set of species and some information about the similarities between pairs of species, reconstruct an evolutionary (phylogenetic) tree, an oriented tree whose vertices are species and which has an arc from vertex $u$ to vertex $v$ if $v$ is a direct ancestor of $u$, and do so in such a way that two species in $S$ are closer together in the tree iff they are more similar. Usually, we want species in $S$ to be leaves (vertices of indegree 0), but suppose we don’t make this assumption. Suppose we also consider the very special case where all similarities are 0 or 1 and build a graph $G$ on the vertex set $V$ by taking an edge between $x$ and $y$ if their similarity is 0. Roberts and Sheng [130] showed that under a particular definition of distance, the problem of reconstructing the phylogenetic tree reduces to the problem of finding an oriented tree $T$ which, after addition of loops at each vertex, has a competition graph that contains $G$ as an induced subgraph. Related papers on competition graphs and phylogenetic tree reconstruction are [129, 131]. See [128] for a survey paper on this topic.

4.2 What Graphs are Competition Graphs?

A widely studied problem is to identify what graphs are competition graphs. In particular, there has been a great deal of work on the special case of competition graphs of acyclic digraphs. Acyclicity is a natural assumption in the ecological application that motivated the subject. Note that every acyclic digraph has a vertex with no outgoing arcs and therefore every competition graph of such a digraph has an isolated vertex. This is the case in the food web $F$ defined in Sec. 4.1. Vertices G and S are isolated in the competition graph. We start by discussing competition graphs of acyclic digraphs.

It was observed in [119] that given any graph $G$, $G$ together with sufficiently many isolated vertices is a competition graph of some acyclic digraph. (The proof is straightforward: Add a new “prey” vertex $p(x, y)$ for each edge $x, y$ of $G$ and let $x$ and $y$ prey on $p(x, y)$.) We use the notation $G \cup I_r$ for $G$ together with $r$ isolated vertices. The smallest number $r$ so that $G \cup I_r$ is the competition graph of an acyclic digraph is called the competition number of $G$ and is denoted by $k(G)$. This concept was introduced in [119], where it was observed that characterization of competition graphs of acyclic digraphs is equivalent to the calculation of competition numbers. Opsut [105] proved that computation of $k(G)$ is an NP-complete problem.

While the problem of characterizing competition graphs is NP-complete, there are some interesting graph-theoretical approaches to it. Competition graphs of acyclic digraphs were characterized by Dutton and Brigham and Lundgren and Maybee [31, 96] in terms of edge coverings by cliques. This idea was generalized to characterizations of graphs of competition number at most $m$ by Lundgren and Maybee [96] (corrected by Kim [79]), and of competition graphs.
of arbitrary digraphs with loops allowed by Dutton and Brigham [31] and to competition graphs of arbitrary digraphs without loops by Roberts and Steif [132].

An interesting special situation is to characterize competition graphs of special types of digraphs. For instance, motivated by the channel assignment problem, Raychaudhuri and Roberts [114] characterized competition graphs of symmetric digraphs that are unit interval graphs. Other work on competition graphs of symmetric digraphs is found in the papers [98, 99] and elsewhere – see survey [128]. Competition graphs of strongly connected and of hamiltonian digraphs are studied in [41, 100] and of tournaments in [38, 39].

There has been a great deal of research on competition number. We mention here two long-term open problems. Opsut [105] calculated the competition number of a line graph and conjectured that a similar result held for all graphs with the property that two cliques could cover all of the vertices of the open neighborhood of any vertex. In spite of almost 20 years of work, Opsut’s conjecture remains open. See [83, 158, 159, 161] for partial results.

The second open problem concerns an elimination procedure for computing the competition number. Such a procedure was introduced by Roberts [119], who conjectured that it always led to the competition number. Opsut [105] showed that this was false. A modified elimination procedure due to Kim and Roberts [85] is known to calculate the competition number for a large class of graphs, and it remains open to determine whether or not it always works.

4.3 Ecological Niches

The competition between species is a major motivating application for the study of competition graphs. In an ecosystem, each species has a “normal, healthy” environment. This environment can be characterized as follows: There is a range (interval) of acceptable values of certain parameters, for instance temperature, humidity, acidity, etc. Suppose there are \( k \) such parameters. The set of all points in Euclidean \( k \)-space so that on each parameter, the point is within the given range, is called the species’ ecological niche. The ecological niche is a “box” in \( k \)-space with sides parallel to the coordinate axes. It is an old ecological principle that two species compete if and only if their ecological niches overlap, i.e., if their boxes overlap. Joel Cohen [19] suggested we try to find the smallest number of parameters \( k \) that are needed to account for observed competitions.

His approach was to define competition in a different way, specifically by using a food web and letting its competition graph define competition. The question is: How many parameters are needed so that each species corresponds to an ecological niche in \( k \)-space and so that there is an edge between \( x \) and \( y \) in the competition graph if and only if their ecological niches overlap? In other words, what is the smallest \( k \) so that \( G \) is the intersection graph of boxes in \( k \)-space?

In the case of the food web \( F \) defined in Sec. 4.1, there are five maximal cliques: \( K_1 = \{B,D,H,O\} \), \( K_2 = \{B,F,W\} \), \( K_3 = \{F,R,W\} \), \( K_4 = \{G\} \), \( K_5 = \)
It is easy to see that $K_1, K_2, K_3, K_4, K_5$ is a consecutive ordering of maximal cliques and hence that the competition graph is an interval graph. So, one dimension suffices. What is surprising is that, in practice, one dimension usually suffices. What does that mean? It means that the competition graph of most real world food webs is an interval graph!

This was first observed in 1968 by Joel Cohen [19]. For more than thirty years people have been trying to explain why, without complete success. The approaches to this problem have been both mathematical and ecological. This has led to a large scientific literature. We discuss this problem in the next subsection.

4.4 What Food Webs Have Competition Graphs that are Interval Graphs?

A fundamental open problem in applied graph theory is to characterize the acyclic digraphs (food webs) whose competition graphs are interval graphs. There are many approaches to this problem. We give some sample results. From a purely computational point of view, the fundamental open problem is easy to solve: Given a digraph, compute its competition graph (easy to do) and use one of the standard linear-time algorithms for checking if this is an interval graph. (A fundamental algorithm is the Booth-Lueker algorithm that uses PQ-trees. See [12].) However, we seek a solution that highlights the structure and properties of the food web that lead to an interval graph competition graph.

Cohen [21] took a statistical approach to this problem. He generated food webs randomly and determined whether or not their competition graphs were interval graphs. By varying the distributions from which the food webs were generated according to some assumptions corresponding to structure, he tried to find models which gave rise to interval graph competition graphs with high probability. (Other statistical/probabilistic approaches to the structure of competition graphs can be found in the papers [15, 22, 23, 24, 25, 103].)

Lundgren and Maybee [97] found a characterization of acyclic digraphs whose competition graph is an interval graph that makes use of the Fulkerson-Gross characterization of interval graphs discussed in Sec. 3.3 and essentially reduces the problem to determining directly if a competition graph is an interval graph. It only begins to shed light on the structural properties of a digraph required for it to have a competition graph that is an interval graph. Steif [143] approached the problem from just such a structural point of view. However, he obtained a negative result: There is no forbidden subdigraph characterization of acyclic digraphs with interval graph competition graphs. Sugihara [145] showed statistically that the frequency with which food webs have interval graph competition graphs could be accounted for by requiring that the competition graph be triangulated, i.e., have no cycles of length greater than 3 as generated subgraphs.

Hefner, Jones, Kim, Lundgren, and Roberts [63] approached the problem by studying digraphs with limited indegrees and outdegrees. They characterized
the acyclic digraphs with indegree and outdegree at most 2 at each vertex and which have interval graph competition graphs. This is the most elementary case, but there is evidence that real world food webs tend to have very low indegrees and outdegrees (see [22]), so small cases do give useful insights. The problem remains open for higher bounds on indegree and outdegree. Hefner, et al. also studied the special case where every vertex has the same indegree and every vertex has the same outdegree. This turns out to be closely related to combinatorial designs, in particular block designs. Hefner, et al. showed, for instance, that under certain circumstances, an acyclic digraph where every vertex has indegree 0 or \( k \) and outdegree 0 or \( r \) has an interval graph competition graph if there is a \((b, v, r, k, \lambda)\)-design.

\( D \) is an interval digraph if two real intervals \( J(x) \) and \( K(x) \) can be assigned to each vertex \( x \) so that \((x, y)\) is an arc of \( D \) iff \( J(x) \cap K(y) \neq \emptyset \). (Interval digraphs are not necessarily acyclic.) Langley, Lundgren, and Merz [89] showed that interval digraphs have interval graph competition graphs and that every interval graph is the competition graph of some interval digraph.

### 4.5 Variants of Competition Graphs

As with the notions of graph coloring and intersection graph, there are many interesting variations of the notion of competition graph. For instance, if we take an edge between \( x \) and \( y \) iff there is a vertex \( u \) with arcs \((u, x), (u, y)\) in \( D \), then we have the common enemy graph of \( D \); if there is an edge between \( x \) and \( y \) iff there are \( u \) and \( v \) with arcs \((x, u), (y, u)\) and \((v, x), (v, y)\) in \( D \), then we have the competition-common enemy graph of \( D \); and if there is an edge between \( x \) and \( y \) iff there is \( u \) with arcs \((x, u), (y, u)\) in \( D \) or there is \( v \) with arcs \((v, x), (v, y)\) in \( D \), then we have the niche graph of \( D \). In these graphs, in the ecological interpretation, we have an edge between two species iff they have a common enemy, or both a common prey and a common enemy, or either a common prey or a common enemy. Common enemy graphs are studied in [90, 145, 160], competition-common enemy graphs in [43, 70, 86, 134, 135], and niche graphs in [5, 13, 14, 17, 35, 45] – to give some references.

Still another variant arises if we take an edge between \( x \) and \( y \) iff there are \( u_1, u_2, \ldots, u_p \) with arcs \((x, u_1), (y, u_1), (x, u_2), (y, u_2), \ldots, (x, u_p), (y, u_p)\) in \( D \). In this case, we speak of the \( p \)-competition graph of \( D \). See [65, 66, 81, 82].

For each of these types of graphs, there are characterization problems, notions analogous to competition number, and many open questions.

### 5 Closing Comment

Graph theory has found widespread application in numerous fields. In turn, these fields have stimulated the development of many new graph-theoretical concepts and led to many challenging graph theory problems. We can anticipate
that the continued interplay between graph theory and many areas of application will lead to important new developments as we begin a new century.

References


22


[91] Lekkerkerker, C.B., and Boland, J. Ch., Representation of a finite graph by a set of intervals on the real line, *Fund. Math.*, 51 (1962), 45-64.


30


