CS762: Graph-Theoretic Algorithms
Lecture 2: Interval Graphs
January 9, 2002

Scribe: Fred Comeau

Abstract

Interval graphs have many applications throughout society. In this lecture, we consider a number of these applications. We also consider what separates interval graphs from other graphs, and how different types of interval graphs are equivalent.

1 Introduction

In the study of graph theory, it is often useful to restrict consideration to certain classes of graph. One such class is that of interval graphs. We shall define these graphs rigorously later on, but for now consider an interval graph as a representation based on a set of intervals. Collections of intervals arise in many places of society, such as in scheduling problems, dating artifacts, placing personal traits into a developmental order, and the synchronization of traffic lights. This means that studying the class of interval graphs not only provides us with insight into their mathematical structure, but any discovery we make has implications for applications in society.

In considering interval graphs, as previously mentioned, we shall first provide a more rigorous definition. From here, we will move to consider some of the applications of interval graphs. After that, we consider how we can decide whether a given graph is an interval graph, and we will see some examples of graphs which aren't interval graphs. Finally, we will see that a number of different representations of interval graphs, based on the definition of the endpoints of the intervals, are all equivalent.

2 Definitions

An interval is a line segment with two definite end points. The ends of the interval can be either open or closed, as in Figure 1. Given some set of intervals such as in Figure 2, one can define an intersection graph.

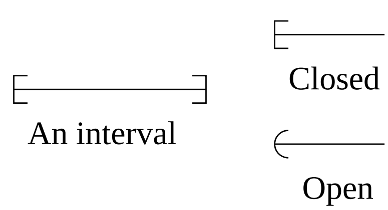

Figure 1: Interval examples

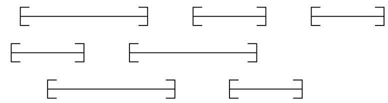

Figure 2: A set of intervals

Definition 1 *Given a set of intervals, an* intersection graph *has a vertex v for every interval I_v, and an edge (v, w) if and only if $I_v \cap I_w \neq \emptyset$.*

Definition 2 *A graph G is an interval graph if it is the intersection graph of some intervals.*

We conclude this section with a number of definitions and notation we will be using throughout the paper.

Definition 3 *An* induced subgraph *of a graph G is formed by considering some subset S of the vertices of G, along with all edges of G that have both endpoints in S.*

This means that in creating an induced subgraph, we can remove any vertices we like, as well as all edges incident on the removed vertices, but all other edges must remain. We cannot add additional edges. Closely related is the *subgraph.*

Definition 4 *A* subgraph *of some graph G is produced by taking some subset of the vertices and edges of G.*

Definition 5 *A graph G contains a* cycle *of size k, C_k, if there exists a path which starts at some vertex v, and ends at the same vertex, of length k, with no repeated vertices.*

3 Interval Graphs

Interval graphs have a number of applications. Some examples include:

archaeology Here the intervals are the time ranges of each artifact, and the problem is finding an order of the ages of the artifacts. However, there are some doubts as to whether the theoretical results are applicable to the practical setting [Spi97].

scheduling This is a common problem in computer science, in which we are given a number of tasks with different start and processing times, and we are interested in find orderings of these tasks which have certain properties. Closely related to this is activity selection, which amounts to finding a maximum independent set in an interval graph [CLRS00].

biology Here certain sequences of DNA are modeled as intervals, and the problem involves constructing maps of the DNA [WG86].

For a more comprehensive list, or more details on these, see [Rob76].

There are graphs that are not interval graphs:

1. $C_k : k \geq 4$. To see this, let us examine C_4. Label the nodes in C_4 as in Figure 3. Because a is connected to both b and c, their corresponding intervals must overlap. However, since b and c do not share an edge, their corresponding intervals are disjoint. An arrangement that depicts this can also be found in Figure 3. The problem occurs when we try and add an interval corresponding to d. Because d is connected to both b and c, it must overlap both of their corresponding intervals, however, d is not connected to a, and hence the corresponding intervals must *not* overlap. Placing an interval corrsponding to d is now impossible.

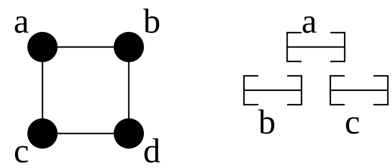

Figure 3: C_4

2. Any graph that contains an induced $C_k : k \geq 4$. This means that intervals graphs are closed under vertex deletion and edge contraction. To see this, note that if we delete a vertex, we can simply remove its interval from the underlying set, which results in an interval graph. In the case of edge contraction, we can simply merge two intervals in the underlying set, to produce another interval graph.

3. The tree of Figure 4. Similar to the C_4 case, we can add the first six intervals without a problem, as depicted in Figure 4. The problem occurs when trying to place an interval corresponding to g. g must intersect the interval of f, but must not intersect the interval of a. Since both ends of the interval of a are already taken, we have a problem.

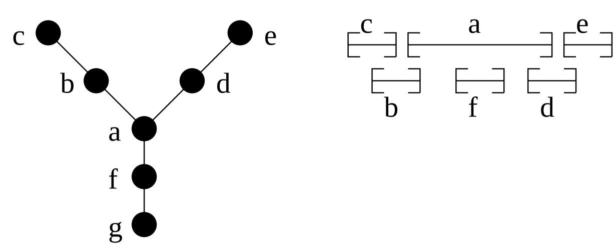

Figure 4: A tree

Until now, we have not differentiated between interval graphs with open intervals, closed intervals, or some mixture of the two. As we shall see, these all turn out to be equivalent.

Theorem 1 *Interval graphs with open, closed, or mixed intervals are all equivalent.*

Note this does not mean that we can simply make open intervals closed, or vice versa. It means that given an interval graph, we can always convert it to one with only closed, or only open intervals, which has the same structure as the original graph. To understand this, notice that if there are no 'ties' in the interval graph, that is, if there are no intervals which start or finish on exactly the same point, it makes no difference whether the intervals are open or closed. As such, the end points of these intervals can be changed as necessary. There can be three types of tie, as indicated in Figure 5. Mirror images of these three are also possible. In the case of a tie of two closed intervals, the two intervals intersect. In the case of a tie involving one or two open intervals, the two intervals do not intersect.

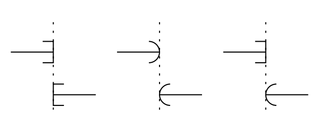

Figure 5: Interval Ties

4 Proof of Theorem 1

Proof: Here we only show the case of converting an arbitrary mixed set of intervals to one with all closed intervals. Similar constructions can be used in other cases.

We begin the proof by noticing several simple constructions that can be applied on interval 'ties'. In the case of a tie involving an open interval, we observe that the two intervals do not intersect, and hence can be replaced by two closed intervals, which are moved slightly apart, as in Figure 6. Figure 6 only depicts the case where both intervals in question are open, but the same construction applies if only one of the intervals is open. If the tie involves two closed intervals, we can move these so that they are overlapping slightly, as in Figure 7.

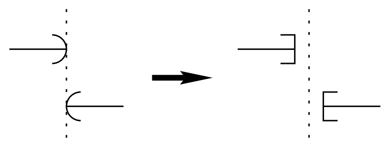

Figure 6: Open Tie Construction

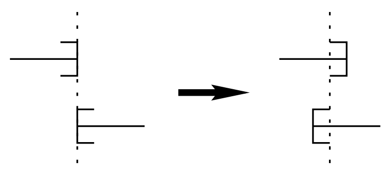

Figure 7: Closed Tie Construction

With these constructions in mind, we proceed with the proof. Given a set of intervals

$$\mathcal{I} = \{I_1, I_2, \ldots, I_n\}.$$

Here each interval $I_i = \{s_i, f_i\}$ is one of $[s_i, f_i], (s_i, f_i), [s_i, f_i),$ or $(s_i, f_i]$. That is, the endpoints of the interval are either open, closed or a combination of the two. We then define

$$\mathcal{I}' = \{I'_1, I'_2, \ldots, I'_n\}.$$

Now each interval is of the form $I'_i = [s'_i, f'_i]$, where s'_iandf'_i are defined according to Figure 8.

$$\begin{array}{ccc}
[s_i, f_i] & \rightarrow & [s_i - \frac{\epsilon}{2}, f_i + \frac{\epsilon}{2}] \\
[s_i, f_i) & \rightarrow & [s_i - \frac{\epsilon}{2}, f_i - \epsilon] \\
(s_i, f_i] & \rightarrow & [s_i + \epsilon, f_i + \frac{\epsilon}{2}] \\
(s_i, f_i) & \rightarrow & [s_i + \epsilon, f_i - \epsilon]
\end{array}$$

Figure 8: Interval Mapping

Here ϵ is a non-negative value which will be determined later. Notice Figure 8 formalizes the constructions we had previously introduced.

Claim 1 *For any $i, j \in \{1, \ldots, n\}$, $I_i \cap I_j \neq \emptyset \iff I'_i \cap I'_j \neq \emptyset$.*

Proof:

$I_i \cap I_j \neq \emptyset \Rightarrow I_i' \cap I_j' \neq \emptyset$: Here we have two cases. The two intervals intersect either in a single point, as in a tie between two closed intervals, or they overlap for some non-empty open interval. In the first case, our construction will lengthen both intervals by $\frac{\epsilon}{2}$, thus they will now overlap by ϵ. In the second case, we have three possibilities. The first possibility is that both intervals are closed. Then we lengthen both by $\frac{\epsilon}{2}$, and thus their overlap is increased by ϵ. The next possibility is that both intervals are open. In this case, both are shortened by ϵ, and thus their intersection is shortened by 2ϵ. Finally, we can have one of each. One is lengthened by $\frac{\epsilon}{2}$, the other is shortened by ϵ, thus their intersection is shortened by $\frac{\epsilon}{2}$. In the worst case, we have shortened the overlap by 2ϵ, so setting $\epsilon \leq \frac{1}{3}$(The length of that interval) ensures that they will still overlap in all cases, and thus the claim holds for a nonempty interval, with appropriate choice of ϵ.

$I_i \cap I_j = \emptyset \Rightarrow I_i' \cap I_j' = \emptyset$: Here again we have two situations. The first involves a tie, in which an open interval participates. The open interval will be shortened by ϵ. The other interval will either be shortened by ϵ again, if it is open, or lengthened by $\frac{\epsilon}{2}$, if it is closed. In either case, the two intervals will be separated by at least $\frac{\epsilon}{2}$, as indicated in Figure 9. In the other case, the two intervals are separated by a length of space (greater than a point). In this case, if both are open, the space increases by 2ϵ. If there is one of each, the space increases by $\frac{\epsilon}{2}$. And if they are both closed, the space decreases by ϵ. So in the worst case, the space can decrease by ϵ. We can simply again choose $\epsilon \leq \frac{1}{3}$(The length of that interval), and this ensures the two intervals will remain separated. Thus the claim holds for an empty interval, again with appropriate choice of ϵ.

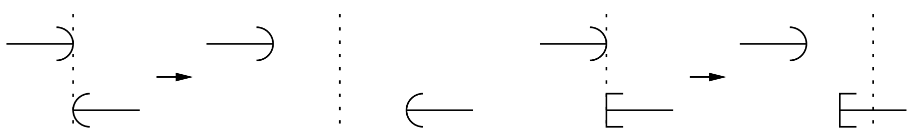

Figure 9: Closed Tie Construction

Finally, notice that setting

$$\epsilon \leq \frac{1}{3} \min\{|x_i - x_j|, x_i, x_j \in \{s_1, \ldots, s_n, f_1, \ldots, f_n\}, x_i \neq x_j\}$$

ensures that the claim holds for all intervals. In setting ϵ in this way, we also ensure no intervals are deleted, since the shortest interval must have length $> 3\epsilon$. $\square$

Since we have no restrictions on the choice of i and j in this claim, it holds for any pair of intervals. This means that the structure of the intervals remains intact, and Theorem 1 holds. $\square$

Because of the above theorem, we know that we can convert an arbitrary interval graph into one that has all closed intervals, and no ties. We can then number the vertices corresponding to the intervals according to the first occurrence of the interval, scanning from left to right. This gives us an ordering of the vertices. We can impose an edge direction on this ordering, such that each edge goes from left to right. This allows us to define the following sets.

Definition 6 *In a directed graph, the set of* predecessors *of a vertex v_i,* Pred(v_i)*, is the set of all vertices v_j for which there exists an edge from v_j to v_i. Similarly, the set of successors of a vertex v_i,* Succ(v_i)*, is the set of all vertices v_j for which there exists an edge from v_i to v_j.*

Definition 7 *The* indegree *of a vertex v_i,* indegree(v_i)*, is the number of incoming edges of v_i. Similarly, the* outdegree *of a vertex v_i,* outdegree(v_i)*, is the number of outgoing edges of vertex v_i.*

Notice that

$$\text{indegree}(v_i) = |\text{Pred}(v_i)|$$

and

$$\text{outdegree}(v_i) = |\text{Succ}(v_i)|.$$

In the context of interval graphs, we can order the vertices by numbering the intervals in order of their leftmost endpoint, reading from left to right, breaking ties arbitrarily. Given some vertex represented by v_i which has a starting endpoint s_i, $\text{Pred}(v_i)$ is the set of all vertices with intervals that start before s_i and do not end before s_i. With this in mind, also notice that all vertices in $\text{Pred}(v_i)$ also overlap with one another, and hence have edges between any pair. This means that $v_i \cup \text{Pred}(v_i)$ is a clique.

References

[CLRS00] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, 2nd edition.* MIT Press, McGraw-Hill Book Company, 2000.

[Rob76] F.S. Roberts. *Discrete Mathematical Models, with Applications to Social, Biological, and Environmental Problems.* Prentice-Hall, 1976.

[Spi97] J. Spinrad. Interval graphs. Available online at http://www.vuse.vanderbilt.edu/ spin/research.html, 1997.

[WG86] M. S. Waterman and J. R. Griggs. Interval graphs and maps of dna. *Bull. Math. Biol.*, 48(2):189–195, 1986.