Available online at www.sciencedirect.com
ScienceDirect
Applied Mathematics Letters 21 (2008) 332–337
Applied Mathematics Letters
www.elsevier.com/locate/aml

A matrix characterization of interval and proper interval graphs

George B. Mertzios

Department of Computer Science, RWTH Aachen University, Germany

Received 24 June 2006; received in revised form 6 March 2007; accepted 5 April 2007

Abstract

In this work a matrix representation that characterizes the interval and proper interval graphs is presented, which is useful for the efficient formulation and solution of optimization problems, such as the k-cluster problem. For the construction of this matrix representation every such graph is associated with a node versus node zero–one matrix. In contrast to representations used in most of the previous work, the proposed matrix characterization does not make use of the maximal cliques in the graph investigated.

Keywords: Interval graph; Proper interval graph; Matrix representation; Matrix characterization

1. Introduction

A graph G is called an *interval graph* if its nodes can be assigned to intervals on the real line, so that two nodes are adjacent in G if and only if their assigned intervals intersect. The set of intervals assigned to the nodes of G is called a *realization* of G. A *proper interval graph* is an interval graph that has an intersection model, in which no interval contains another one properly. Both classes of graphs are important for their applications to scheduling problems, biology, VLSI circuit design, as well as to psychology and social sciences [1–3]. Several difficult optimization problems are solvable in polynomial time on interval graphs, while they are NP-hard in the general case. Some of these problems are the maximum clique, the maximum independent set [4,5], the Hamiltonian cycle and the Hamiltonian path problem [6]. A *unit interval graph* is an interval graph that has an intersection model, in which every interval has unit length. The classes of unit and proper interval graphs coincide [7–9].

There are several characterizations of interval graphs, as well as of the proper and unit interval graphs. An arbitrary graph G is an interval graph if and only if one of the following equivalent conditions hold: (a) it is chordal and its complement $\overline{G}$ is a comparability graph [8], (b) it contains no induced C_4 and $\overline{G}$ is transitively orientable, or (c) it is chordal and contains no asteroidal triple (AT) [7]. Alternatively, the interval graphs are characterized by the consecutive ones property [10], i.e., the maximal cliques can be linearly ordered such that, for every node v, the maximal cliques containing v occur consecutively [7,8]. Namely, in the clique versus node incidence matrix of any interval graph there is a permutation of its rows such that the ones in each column appear consecutively. On the other hand, the proper interval graphs are characterized as graphs containing no astral triples, as well as interval graphs without containing any induced claw $K_{1,3}$ [7,9]. Moreover, the interval order is a partial order, which is characterized

E-mail address: mertzios@cs.rwth-aachen.de.

0893-9659/$ - see front matter
doi:10.1016/j.aml.2007.04.001

by forbidding $\underline{2}+\underline{2}$ posets, while the proper interval order or, equivalently, the unit interval order or the semiorder, is an interval order without $\underline{1}+\underline{3}$ posets. Here, the posets $\underline{2}+\underline{2}$ and $\underline{1}+\underline{3}$ are the posets consisting of two disjoint chains of sizes 2, as well as 1 and 3 respectively [9,11].

In this work a matrix representation of the interval and proper interval graphs is proposed, which is based on a node versus node zero–one matrix of them. This characterization leads to a matrix representation of both classes of graphs, which is useful for the efficient formulation and solution of optimization problems. Specifically, it is used to solve efficiently the *k-cluster problem*, whose complexity on interval and proper interval graphs was an interesting open question [12]. Recently, it has been shown that the k-cluster problem on the interval graphs is solvable in polynomial time [13], while it is NP-hard for chordal graphs [12]. In the sequel, the proposed matrix characterization of the interval and its restriction on the proper interval graphs are presented in Sections 2 and 3 respectively.

2. The interval graphs in the general case

Without loss of generality, we may suppose that all intervals in a realization of an interval graph are closed, i.e. of the form $[a, b]$. However, this representation is too general. To this end, a more suitable interval representation form is proposed in Definition 1. Algorithm 1 describes how to transform an arbitrary given interval graph to this form. Recall that an interval graph can be recognized in linear time [14,15]. In the following, suppose we are given a realization of an interval graph G on n nodes.

Definition 1. A representation of n intervals, having the following properties, is called a *Normal Interval Representation (NIR) form*:

1. all intervals are of the form $[i, j)$, where $0 \le i < j \le n$,
2. exactly one interval begins at i, for every $i \in \{0, 1, \ldots, n-1\}$.

Algorithm NIR(G):
Input: A realization with closed intervals of an interval graph G on n nodes.
Output: A NIR form of G.

1. Suppose that some intervals of the graph share exactly one common point $x \in$ R, as well as that the next greatest point at which some interval begins is $\xi_2 > x$ and the next smaller point at which some interval ends is $\xi_1 < x$ respectively. Replace any interval $[a, x]$ by the interval $[a, (x+\xi_2)/2]$ and any interval $[x, b]$ by the interval $[(x+\xi_1)/2, b]$. Repeat, until this step cannot be further applied.
2. Replace every interval $[a, b]$ by the interval $[a, b)$.
3. Suppose that exactly $l > 1$ intervals begin at the same point a and that the next greatest point at which some interval begins or at which some interval ends is $b > a$. Move the left end of the ith of these l intervals from a to $a + (i-1) \cdot (b-a)/l$. Repeat, until this step cannot be further applied.
4. Suppose that the left ends of the n intervals are $a_1 < a_2 < \cdots < a_n$. Replace any interval of the form $\left[a_j, b\right)$, where $j \le i$ and $a_i < b < a_{i+1}$, by the interval $\left[a_j, a_{i+1}\right)$. Also replace any interval of the form $[a_j, b)$ with $b \ge a_n$ by the interval $\left[a_j, a_n+1\right)$.
5. Move bijectively the point a_i to the point $i-1$, for $i = 1, 2, \ldots, n$, and the point a_n+1 to the point n.

Algorithm 1. The transformation of the given integral graph to the NIR form.

Theorem 1. *Algorithm 1 runs in time $O(n)$ and the resulting graph is isomorphic to the input G.*

Proof. Every step of Algorithm 1 operates on each interval at most twice. Therefore, since there are n intervals, its running time is $O(n)$. Additionally, two arbitrary intervals of the resulting NIR form intersect exactly if the corresponding intervals of G intersect. Consequently, no edge is added to or removed from G and thus the resulting graph is isomorphic to G. □

Note that the NIR form of an interval graph G is not unique. Indeed, suppose that in an NIR form of G no interval has the form $[a, i)$, for some $i \in \{1, 2, \ldots, n\}$. Then, by exchanging the left ends of the ith and the $(i+1)$th intervals, the qualitative relation between these intervals and the rest ones in G remains unchanged. Thus with such an operation on an NIR form, we may obtain another NIR form for the same graph. Here, by the term "qualitative relation" we mean whether two intervals intersect or not.

Lemma 1. *An arbitrary graph is an interval graph iff it can be represented by the NIR form.*

Proof. According to Theorem 1, any interval graph can be transformed to the NIR form. Conversely, the NIR form is clearly a set of intervals, i.e. it corresponds to an interval graph. □

Since no pair of graph intervals in the NIR form share a common left end, it is possible to define a perfect order over them. Using this order of the intervals we define in Definition 2 the proposed matrix representation of the interval graphs.

Definition 2. Consider the ith interval $[i-1, b)$ of the NIR form of the interval graph G, which corresponds to the node v_i of G. The *Normal Interval Representation (NIR) matrix* H_G of G is defined as the lower triangular portion of the adjacency matrix of G with zero diagonal, in which the ith row and column correspond to the node v_i.

The NIR matrix H_G of G can also be seen as looking at "half of the adjacency matrix", where in addition, requirements have been imposed on how to label the lines of the adjacency matrix. Using the above NIR matrix H_G of G, the k-cluster problem on an interval graph may be reformulated [13]. Indeed, the ith diagonal element of H_G has a chain of consecutive ones below it, while all remaining elements of this column are zero. The number x_i of the consecutive ones in the ith column equals the number of intervals within the $(i+1)$ th, ..., nth ones that intersect with the ith interval. Fig. 1(a) shows an example of the form of H_G.

Denote further the desired k-subgraph of G with the maximum number of edges by C_k. Join the variable $z_i \in \{0, 1\}$ to the ith interval. The case $z_i = 1$ indicates that the ith node of G, i.e. the ith interval of its NIR form, is included in C_k. Let now $1 \le j < i \le n$. The jth and the ith intervals intersect in C_k if and only if the quantity $z_j \cdot z_i \cdot H_G(i, j) \in \{0, 1\}$ equals one. Indeed, in this case both intervals have been chosen in C_k, i.e. $z_i = z_j = 1$, and, simultaneously, the jth interval ends strictly further than $i-1$, where the ith one begins, i.e. $H_G(i, j) = 1$. Thus, the number of intersections among the k intervals of the realization of C_k equals

$$\sum_{i=2}^{n} \sum_{j=1}^{i-1} z_j \cdot z_i \cdot H_G(i, j) = z^{\mathrm{T}} \cdot H_G \cdot z \tag{1}$$

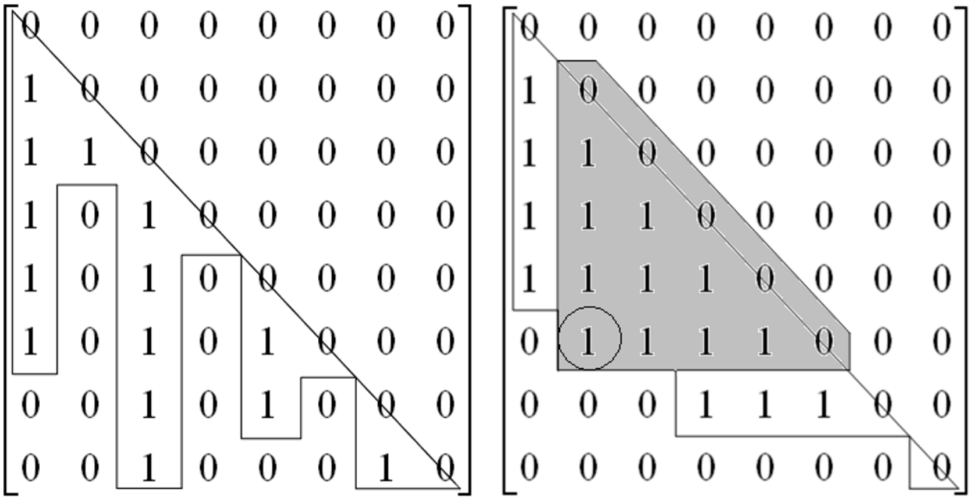
where $z = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix}^{\mathrm{T}}$ and H_G is the NIR matrix of G.

Since C_k has exactly k nodes, exactly k entries of the vector z are one. Thus, the k-cluster problem on G is equivalent to finding the appropriate subset $I \subseteq \{1, 2, \ldots, n\}$ of the satisfied entries of z, with $|I| = k$, so that the quantity $\sum_{\substack{i,j \in I \\ i>j}}, H_G(i, j)$ is maximized. If H_G were an arbitrary zero–one lower triangular matrix with zero diagonal, then the latter optimization problem would correspond to the k-cluster problem on an arbitrary graph, which is NP-hard. However, since H_G is an NIR matrix, then its special column structure allows us to solve it in polynomial time using a dynamic programming approach [13].

Lemma 2. *Any maximal clique of G corresponds bijectively to a row of its NIR matrix H_G, in which at least one of its unit elements or its zero diagonal element does not have any chain of 1's below it.*

Proof. Consider an arbitrary row of H_G, let it be the ith one, in which exactly the i_1th, i_2th, ..., i_rth elements equal one. Clearly, the ith and the jth intervals intersect for every $j \in \{i_1, i_2, \ldots, i_r\}$, since $H_G(i, j) = 1$. The i_1th, i_2th, ..., i_rth intervals of G intersect each other also, due to the NIR form of H_G. Thus, the i_1th, i_2th, ..., i_rth, ith intervals build a clique Q in G. Consider now the case where in this row at least one of its i_1th, i_2th, ..., i_rth, ith elements, say the jth one, does not have any chain of 1's below it. Suppose also that there exists another clique Q' in G, which strictly includes Q. Since $H_G(l_1, j) = H_G(i, l_2) = 0$ for every $l_1 > i$ and $l_2 \in \{1, 2, \ldots, i\} \backslash \{i_1, i_2, \ldots, i_r\}$, the l_1th and the jth, as well as the ith and the l_2th intervals, do not intersect. Therefore, Q' cannot be a clique, which is a contradiction. Thus, Q is a maximal clique.

Conversely, let Q be a maximal clique in G, which contains the i_1th, i_2th, ..., $i_{|Q|}$th intervals of its NIR form, where $i_1 < i_2 < \cdots < i_{|Q|}$. Consider now the $i_{|Q|}$th row of H_G. Since Q is a clique, the i_1th, i_2th, ..., $i_{|Q|-1}$th intervals intersect with the $i_{|Q|}$th one and therefore $H_G\left(i_{|Q|}, j\right) = 1$ for every $j \in \left\{i_1, i_2, \ldots, i_{|Q|-1}\right\}$. Suppose $i_{|Q|} < n$. Then, if $H_G\left(i_{|Q|}+1, j\right) = 1$ for every $j \in \left\{i_1, i_2, \ldots, i_{|Q|}\right\}$, the $i_{|Q|+1}$th row corresponds to another clique Q' that includes Q strictly, which is a contradiction. Thus, at least one of the i_1th, i_2th, ..., $i_{|Q|}$th elements of the



(a) (b)

Fig. 1. (a) The NIR matrix H_G of an interval graph G. (b) The SNIR matrix $H_{G'}$ of a proper interval graph G'.

$i_{|Q|}$th row does not have any chain of 1's below it. Finally, in the case where $i_{|Q|} = n$, obviously none of the i_1th, i_2th, ..., $i_{|Q|}$th elements of the $i_{|Q|}$th has any chain of 1's below it. □

3. The proper interval graph case

Consider now the case where G is a proper interval graph. Since G is also an interval graph, Algorithm 1 can be applied to it, leading thus to a special type of the NIR form, as is described in Definition 3.

Definition 3. An NIR form of n intervals is called a *Stair Normal Interval Representation (SNIR) form*, iff it has the following additional property:

If for the intervals $[a, b)$ and $[c, d)$, $a < c$ holds, then $b \leq d$ also holds.

Lemma 3. *Every proper interval graph is transformed to the SNIR form, after applying Algorithm 1 on it.*

Proof. Suppose that the left end of a vector $v_1 = [a, b]$ is initially strictly less than the left end of another vector $v_2 = [c, d]$, i.e. $a < c$. Then the same also holds for their right ends, respectively, i.e. $b < d$, since otherwise v_2 would strictly include v_1, which is a contradiction. Suppose that v_1 and v_2 are transformed by the algorithm to the vectors $v_1' = [a', b')$ and $v_2' = [c', d')$ respectively. Then, it holds that $a' < c'$ and $b' \leq d'$, since v_1 and v_2 may be "aligned" by their right ends (Step 4 of Algorithm 1), while the relative order of their left ends remains unchanged, due to the structure of the algorithm. Thus, the NIR form obtained satisfies the condition of Definition 3, i.e. it is a SNIR form. In the special case of two identical intervals, i.e. $a = c$ and $b = d$, Algorithm 1 returns the same right end $b' = d'$ for them, while their left ends are ordered by increasing order, i.e. in this case the obtained NIR form is also a SNIR form. □

Definition 4. The NIR matrix H_G that corresponds to the SNIR form of a proper interval graph G is called the *Stair Normal Interval Representation (SNIR) matrix* of G.

Definition 5. Consider the SNIR matrix H_G of the proper interval graph G. The matrix element $H_G(i, j)$ is called a *pick* of H_G iff:

1. $i \geq j$,
2. if $i > j$ then $H_G(i, j) = 1$,
3. $H_G(i, k) = 0$ for every $k \in \{1, 2, \ldots, j-1\}$ and
4. $H_G(l, j) = 0$ for every $l \in \{i+1, i+2, \ldots, n\}$.

Given the pick $H_G(i, j)$ of H_G, the set

$$S := \{H_G(k, l) : j \leq l \leq k \leq i\}$$

of matrix entries is called the *stair* of H_G, which corresponds to this pick.

Recall that the left and the right ends of the ith interval in the SNIR form of G correspond to the ith and the $(x_i + i)$th elements of the ith column of H_G respectively, where x_i equals the number of consecutive ones below the ith diagonal element. Therefore, due to Definition 3, we have that $x_i + i \geq x_j + j$ for $i > j$. Consequently, any stair of H_G consists of unit matrix elements, except the diagonal elements of H_G, while the corresponding pick is the lowermost left matrix entry of this stair. As is seen in Fig. 1(b), the SNIR matrix H_G has a stair shape and equals the union of all its stairs. In this figure a stair of H_G can be recognized, where the corresponding pick is marked with a circle.

Lemma 4. *An arbitrary graph is a proper interval graph iff it can be represented by the SNIR form.*

Proof. Due to Lemma 3, any proper interval graph can be written in the SNIR form. Conversely, the SNIR form is clearly a set of intervals, where none of them strictly includes another one, i.e. it is a realization of a proper interval graph. □

Lemma 5. *Any stair of the SNIR matrix H_G corresponds bijectively to a maximal clique in G.*

Proof. Due to Lemma 2, every maximal clique of G corresponds bijectively to a row of H_G, in which at least one of its unit elements or its zero diagonal element does not have any chain of 1's below it. However, since G is a proper interval graph and due to Definition 5, it is concluded that such a row corresponds bijectively to a pick of H_G and therefore to a stair of it, as is shown in Fig. 1(b). □

4. Conclusions

An efficient matrix representation characterizing the interval and proper interval graphs has been presented, which is useful for the efficient formulation and solution of optimization problems, such as the k-cluster problem. For the construction of this matrix representation every such graph is associated with a node versus node zero–one matrix. In contrast to representations used in most of the previous work, the matrix characterization presented does not make use of the maximal cliques in the graph investigated.

Acknowledgement

I wish to thank the anonymous reviewer for useful suggestions, which improved the presentation of the work.

References

[1] M.C. Golumbic, A.N. Trenk, Tolerance Graphs, Cambridge University Press, Cambridge, 2004.
[2] A.V. Carrano, Establishing the order to human chromosome-specific DNA fragments, in: A.D. Woodhead, B.J. Barnhart (Eds.), Biotechnology and the Human Genome, Plenum Press, New York, 1988, pp. 37–50.
[3] S.B. Sadjad, H.Z. Zadeh, Unit interval graphs, properties and algorithms, School of Computer Science, University of Waterloo, February 2004.
[4] U.I. Gupta, D.T. Lee, J.Y.T. Leung, Efficient algorithms for interval graphs and circular-arc graphs, Networks (1982) 459–467.
[5] Ju Yuan Hsiao, Chuan Yi Tang, An efficient algorithm for finding a maximum weight 2-independent set on interval graphs, Inform. Process. Lett. 43 (5) (1992) 229–235.
[6] M.S. Chang, S.L. Peng, J.L. Liaw, Deferred-query — An efficient approach for problems on interval and circular-arc graphs (extended abstract), in: WADS, 1993, pp. 222–233.
[7] A. Brandstädt, V.B. Le, J.P. Spinrad, Graph Classes: A Survey, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1999.
[8] M.C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, vol. 57. Annals of Discrete Mathematics, Amsterdam, The Netherlands, 2004.
[9] K.P. Bogart, D.B. West, A short proof that "proper = unit", Discrete Math. 201 (1–3) (1999) 21–23.
[10] D.R. Fulkerson, O.A. Gross, Incidence matrices and interval graphs, Pacific J. Math. 15 (1965) 835–855.
[11] R.P. Stanley, An introduction to hyperplane arrangements, in: Lecture Notes Based on a Lecture Series at the Park City Mathematics Institute, 2006.

[12] D.G. Corneil, Y. Perl, Clustering and domination in perfect graphs, in: Discrete Applied Mathematics, vol. 9, 1984, pp. 27–39.
[13] G.B. Mertzios, A polynomial algorithm for the k-cluster problem on interval graphs, in: Proceedings of Combinatorics 2006, in: Electronic Notes in Discrete Mathematics, vol. 26, Ischia, Naples, Italy, June 25–July 1, 2006, pp. 111–118.
[14] W.L. Hsu, A simple test for interval graphs, in: WG '92: Proceedings of the 18th International Workshop on Graph-Theoretic Concepts in Computer Science, Springer-Verlag, London, 1993, pp. 11–16.
[15] D.G. Corneil, S. Olariu, L. Stewart, The ultimate interval graph recognition algorithm? in: SODA '98: Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998, pp. 175–180.