Available online at www.sciencedirect.com

SCIENCE DIRECT®

Discrete Applied Mathematics 143 (2004) 307–311

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

Notes

A note on tolerance graph recognition☆

Ryan B. Hayward[a], Ron Shamir[b]

[a] *Department of Computing Science, University of Alberta, Edmonton, Alta., Canada T6G 2E8*
[b] *School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel*

Received 23 May 2002; received in revised form 23 May 2003; accepted 27 August 2003

Abstract

A graph $G=(V,E)$ is a *tolerance graph* if there is a set $I=\{I_v \mid v\in V\}$ of closed real intervals and a set $\tau=\{\tau_v \mid v\in V\}$ of positive real numbers such that $(x,y)\in E \Leftrightarrow |I_x\cap I_y| \geqslant \min\{\tau_x,\tau_y\}$. We show that every tolerance graph has a polynomial size integer representation. It follows that tolerance graph recognition is in NP.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Tolerance graph; Tolerance representation; Class NP

1. Overview

A graph $G=(V,E)$ is a *tolerance graph* if there is a set $I=\{I_v \mid v\in V\}$ of closed real intervals and a set $\tau=\{\tau_v \mid v\in V\}$ of positive real numbers called tolerances such that

$$(x,y)\in E \Leftrightarrow |I_x\cap I_y| \geqslant \min\{\tau_x,\tau_y\}. \tag{1}$$

Any such triple (V,I,τ) is a *tolerance representation* for G. A tolerance τ_v of a tolerance representation is *unbounded* if $\tau_v > |I_v|$. Notice that if a tolerance τ_v is unbounded, τ_v can have any value greater than $|I_v|$, since for any other vertex w, $|I_v\cap I_w| \geqslant \min\{\tau_v,\tau_w\}$ if and only if $|I_v\cap I_w| \geqslant \tau_w$. Each interval I_v of a tolerance representation is represented by its ordered pair of endpoints $[p_v^-, p_v^+]$, where p_v^- and p_v^+ denote the respective left and right endpoints; thus $p_v^- \leqslant p_v^+$ and $|I_v| = p_v^+ - p_v^-$. For each v in V with bounded tolerance, we refer to $p_v^{*-} = p_v^- + \tau_v$ and $p_v^{*+} = p_v^+ - \tau_v$ as the respective left and right *tolerance points* of v.

Tolerance graphs were introduced by Golumbic and Monma [8] as a generalization of interval graphs; see also [9]. While much is known about tolerance graphs (see the monograph by Golumbic and Trenk [10] and the bibliography by Golumbic [7]), the complexity of recognizing them is an open problem; in fact, even the complexity of verifying that a graph is a tolerance graph has been an open problem, as it was not known whether every tolerance graph has a polynomial size tolerance representation. In this paper we show that this is indeed the case, and so tolerance graph recognition is in NP.

Our method also applies to the related classes of *bounded tolerance graphs* [1,5], in which all tolerances are bounded, and (*bounded*) *bi-tolerance graphs* [2], in which left and right endpoint tolerances are specified separately, so the recognition of these graph classes is also in NP.

Roughly, our argument is that

- for any tolerance representation of a graph with n vertices, the adjacency conditions which define a tolerance representation are completely determined by two linear orders defined in terms of I and τ, namely the linear orders consisting

☆ The first author acknowledges the support of NSERC. The second author was supported in part by the Israeli Science Foundation founded by the Israeli Academy of Sciences and Humanities (Grant 309/02).

E-mail addresses: hayward@cs.ualberta.ca, rshamir@tau.ac.il (R. Shamir).

0166-218X/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2003.08.012

of (i) the $2n$ interval endpoints together with the $2n$ (or fewer) tolerant points, and (ii) the bounded tolerances and the lengths of all intervals with unbounded tolerance;
- given these linear orders, a system of inequalities (SI) can be formulated to which any feasible solution yields a tolerance representation of G;
- there is some feasible non-negative integer SI solution for which the size of the integers needed to represent the intervals and tolerances is polynomial in the size of G.

The details follow.

2. From tolerance representation to linear orders

We begin by identifying the essential order information encoded in a tolerance representation.

Given a multiset $M=\{x_1,x_2,\ldots,x_t\}$ of real numbers, a *linear order* (or sorting) $\mathscr{L}(M)$ of M is an ordering $x_{\rho(1)}r_1x_{\rho(2)}r_2\ldots x_{\rho(t-1)}r_{t-1}x_{\rho(t)}$ of the elements of M, where for each j, the relation r_j is either equality $(=)$ or less-than $(<)$. Given a tolerance representation with vertex set V, define V_T and V_U as the vertex subsets whose tolerances are, respectively, bounded and unbounded.

Lemma 1. *For any tolerance representation $R=(V,I,\tau)$, the set of pairs of vertices x,y with $|I_x\cap I_y|\geqslant\min\{\tau_x,\tau_y\}$ is determined by the partition (V_T,V_U) of V together with the linear orders of the two multisets*

$$Q_1(R)=\{\,p_v^-,p_v^+\mid v\in V_U\}\cup\{\,p_v^-,p_v^+,p_v^{*-},p_v^{*+}\mid v\in V_T\}\quad \textit{and}$$

$$Q_2(R)=\{|I_v|\mid v\in V_U\}\cup\{\tau_v\mid v\in V_T\}.$$

Proof. There are six cases to consider, depending on the linear order of the four endpoints of I_x and I_y. By relabelling x and y if necessary, we may assume that $p_x^-\leqslant p_y^-$ and that $p_x^+\leqslant p_y^+$ if $p_x^-=p_y^-$, and so there are only three cases to consider. Whichever of these cases occurs can be determined from Q_1. In one case, $p_x^+\leqslant p_y^-$, so $|I_x\cap I_y|=0<\min\{\tau_x,\tau_y\}$. In another case, $p_x^-\leqslant p_y^-\leqslant p_y^+\leqslant p_x^+$, so $|I_x\cap I_y|=|I_y|$, and whether $|I_y|\geqslant\min\{\tau_x,\tau_y\}$ can be determined from Q_2, since the inequality holds if and only if $y\in V_T$, or $y\in V_U$ and $|I_y|\geqslant\tau_x$. In the last case, $p_x^-\leqslant p_y^-\leqslant p_x^+\leqslant p_y^+$, so $|I_x\cap I_y|=p_x^+-p_y^-$, and $p_x^+-p_y^-\geqslant\min\{\tau_x,\tau_y\}\Leftrightarrow p_y^-+\min\{\tau_x,\tau_y\}\leqslant p_x^+\Leftrightarrow p_y^{*-}\leqslant p_x^+$ or $p_y^-\leqslant p_x^{*+}$, and whether the last condition holds can be determined from Q_1. □

3. Expansive representations

We next present some simplifying assumptions which can be made about tolerance representations.

Call a tolerance representation *integer* (respectively, *rational*) if all endpoints and tolerances are integer (rational). Call a tolerance representation *expansive* if all endpoints and tolerance points are distinct and all bounded tolerances and unbounded interval lengths are distinct (namely, the multisets $Q_1(V,I,\tau)$ and $Q_2(V,I,\tau)$ are each sets). Call two tolerance representations R and R' *order-equivalent* if they preserve order relations among all endpoints and tolerance points and among all bounded tolerances and unbounded interval lengths, namely, if $\mathscr{L}(Q_1(R))=\mathscr{L}(Q_1(R'))$ and $\mathscr{L}(Q_2(R))=\mathscr{L}(Q_2(R'))$.

Lemma 2. *Every tolerance graph has an expansive integer representation.*

Proof. By making small perturbations to endpoints and tolerances, any tolerance representation can be transformed into a tolerance representation R of the same graph, such that in R all endpoints and tolerance points are distinct and all tolerances are distinct (see [10] or [12] for a detailed proof of this result). Let ε be the smallest difference between consecutive but non-equal terms in $\mathscr{L}(Q_1(R))$ or $\mathscr{L}(Q_2(R))$, and let t_x and t_y be any equal elements in $\mathscr{L}(Q_2(R))$ (so each of t_x,t_y is either a bounded tolerance or the length of an interval with unbounded tolerance). Let R' be the tolerance representation obtained from R as follows: if t_x is a tolerance τ_x, replace τ_x with $\tau_x-\varepsilon/2$; if t_x is the length $|I_x|$ of the interval I_x, replace this interval's right endpoint p_x^+ with $p_x^++\varepsilon/2$. It is easy to check that R' represents the same graph as R, that the elements of $Q_1(R')$ are all distinct (as are the elements of $Q_1(R)$), and that the number of distinct elements of $Q_2(R')$ is one more than the number of distinct elements of $Q_2(R)$, since t_x is now different from all other elements. Thus repeating this argument yields an expansive representation.

Once we have an expansive representation, we can obtain an order-equivalent expansive rational representation by repeatedly defining ε as above, selecting any non-rational endpoint or tolerance, and replacing it with a rational number that differs from the non-rational number by less than $\varepsilon/2$. Once we have an expansive rational representation, we can obtain an order-equivalent integer expansive representation by mulitplying all endpoint and tolerance values by the least common multiple of all denominators. □

4. From linear orders to a system of inequalities

We now show how the partition (V_T, V_U) and the linear orders $Q_1(R), Q_2(R)$ which encode the essential information of a tolerance representation R can be described with a system of inequalities.

Given an expansive integer tolerance representation $R=(V,I,\tau)$, let n, n_b represent $|V|, |V_T|$, respectively, and let $S(R)$ be the system of equations (in fact, exactly one equation) and strict inequalities whose $2n+n_b$ variables correspond to the $2n$ interval endpoints and n_b bounded tolerances, and whose relations are

(i) the $2n+2n_b$ relations which establish the linear order of $Q_1(I,\tau)$ and set the minimum point in $Q_1(I,\tau)$ to 0,
(ii) the $n-1$ relations which establish the linear order of $Q_2(I,\tau)$.

Observe that since $p_v^{+*}, p_v^{-*}, |I_v|$ can be expressed in terms of p_v^+, p_v^-, τ_v, we do not need to introduce new variables for $p_v^{+*}, p_v^{-*}, |I_v|$. Also, observe that since R is a tolerance representation and expansive, the minimum value in $Q_2(R)$ is positive. Also, observe that $S(R)$ implicitly encodes the partition (V_T, V_U), since a vertex v is in V_T if and only if the τ_v is one of the elements of $Q_2(R)$.

Using $S(R)$, we can establish our main result.

Theorem 3. *Every tolerance graph with n vertices has a non-negative integer tolerance representation R in which the maximum integer is at most* $1+5n2^{5n}$.

Proof. Let G be a tolerance graph. By Lemma 2, G has an expansive integer representation R. Since shifting all endpoints by the same amount does not change any order information, we may assume that the smallest endpoint of R has value 0.

Let $S'(R)$ be the system of equalities and inequalities obtained from $S(R)$ by replacing each strict inequality $(<)$ with a non-strict inequality $(\leqslant)$ with a gap of at least 1 (namely, each relation $x_j < x_k$ is replaced with the relation $x_j + 1 \leqslant x_k$). Any integer solution to S is a solution to S', so S' has an integer solution (namely, R). Let R' be any integer solution to S'. Then R' is a tolerance representation of some graph H (this can be shown by letting the values of R' correspond to interval endpoints and tolerance lengths in exactly the same way that R was obtained); since R' is order-equivalent to R, $H=G$ by Lemma 1.

Following standard linear programming techniques, S' can be expressed as a matrix equation

$$Ax = b \quad x \geqslant 0, \tag{2}$$

where

- the first $2n$ variables $x_1,\ldots,x_{2n}$ correspond to the interval endpoints $p_1^-, p_1^+, \ldots, p_n^-, p_n^+$,
- the next n_b variables $x_{2n+1},\ldots,x_{2n+n_b}$ correspond to the bounded tolerances $\tau_1,\ldots,\tau_{n_b}$,
- the next $2n+2n_b-1$ (slack) variables $x_{2n+n_b+1},\ldots,x_{4n+3n_b-1}$ correspond to the $2n+2n_b-1$ gaps which determine $\mathscr{L}(Q_1(I,\tau))$,
- the next $n-1$ (slack) variables $x_{4n+3n_b},\ldots,x_{5n+3n_b-2}$ correspond to the $n-1$ gaps which determine $\mathscr{L}(Q_2(R))$,

and

- the first $2n+2n_b$ rows of A correspond to the relations which determine $\mathscr{L}(Q_1(R))$ and set the minimum value of $Q_1(R)$ to 0,
- the next $n-1$ rows of A correspond to the relations which determine $\mathscr{L}(Q_2(R))$.

Since S' has a feasible (rational) solution, (2) has a feasible solution, and so by a fundamental theorem of linear inequalities (see Theorem 3.4 in [3] or Theorem 7.1 in [14]), S' has a basic feasible solution (that is, a feasible solution in which at most $3n+2n_b-1$ of the variables are non-zero). Also, A has rank $3n+2n_b-1$, since the submatrix induced by the columns indexing the $3n+2n_b-2$ slack variables and the variable corresponding to the minimum value of Q_1 clearly has full rank. Thus any submatrix B of A corresponding to such a basis is invertible, and $x_B = B^{-1}b$ is a basic feasible solution.

By Cramer's rule [4], each entry of B^{-1} can be written as a ratio whose absolute value is of the form $|\det B_{jk}|/|\det B|$, where B_{jk} is the submatrix of B obtained by deleting column j and row k.

Now let x'_B and b' be the vectors obtained by multiplying each entry of x_B and b, respectively, by the scaling factor $s = |\det B|$. Notice that x'_B is a basic feasible solution to

$$Ax = b' \quad x \geqslant 0 \tag{3}$$

and that all entries of x'_B are integers. Also, the tolerance representations given by x_B and x'_B are clearly order-equivalent, since one is obtained from the other by multiplicative scaling of endpoints, tolerance points, and tolerances (this can also be seen by noting that the only difference between (2) and (3) is to change the minimum gap size from 1 to s).

The maximum value of a variable x'_j of x'_B is at most $(3n + 2n_b - 1)\max_{j,k}\{|\det B_{jk}|\}$, since each entry of b is in $\{0, \pm 1\}$. Since each relation of Q_1 is a comparison between two expressions of the form $p_j^{\pm}(\pm\tau_j)$, and each relation of Q_2 is a comparison between two expressions of the form τ_j or $p_k^+ - p_k^-$, each row of B has at most four non-zero entries, and those entries are each $+1$ or -1. Hadamard's inequality [11] states that for a $0, \pm 1$ $m \times m$ matrix Z

$$\det Z \leqslant \prod_{j=1}^{m} \left(\sum_{k=1}^{m} z_{jk}^2 \right)^{1/2}.$$

It follows that

$$|\det B_{jk}| \leqslant \prod_{p=1, p\neq j}^{3n+2n_b-1} \left(\sum_{q=1, q\neq k}^{3n+2n_b-1} b_{pq}^2 \right)^{1/2} \leqslant \prod_{j=1}^{3n+2n_b-2} 4^{1/2} = 2^{3n+2n_b-2}$$

and so the maximum value of a variable of x'_B is at most $(3n + 2n_b - 2)2^{3n+2n_b-2} \leqslant 5n2^{5n}$.

Since the tolerance representation under construction is based on S', to this point values have been assigned for all interval endpoints and all bounded tolerances. Finally, for every vertex which has a tolerance which did not appear in $Q_2(R)$, assign the value of $1+M$ to this vertex's tolerance, where M is the maximum integer used so far. This assignment completes the construction of the desired tolerance representation by ensuring that every vertex which is supposed to has an unbounded tolerance. Thus the theorem holds. □

Corollary 4. *Any n-vertex tolerance graph has an* $O(n^2)$*-bit tolerance representation, and verifying that such a representation represents the given graph takes only* $O(n^3)$ *time. Thus the tolerance graph recognition problem* ("*given a graph, is it a tolerance graph*?") *is in* NP.

Proof. The tolerance representation constructed in Theorem 3 consists of $3n$ non-negative integers representing $2n$ endpoints and n tolerances; since the maximum integer requires $O(\lg(1 + 5n2^{5n})) = O(n)$ bits, the tolerance representation requires a total of $O(n^2)$ bits.

To show that a decision problem is in the class NP, it is sufficient to show that there is a polynomial time algorithm which verifies all yes-instances; see for example Chapter 2 in [6], Proposition 9.1 in [13], or Definition 7.16 in [15].

Obviously, any tolerance representation (V, I, τ) of a tolerance graph $G = (V, E)$ with n vertices is a certificate that G is a tolerance graph, since to verify that G is a tolerance graph it suffices to check that (1) holds for each pair of vertices. Moreover, this verification takes only $O(n^2)$ integer addition, subtraction, and comparison operations, and these operations can be performed in time linear in the number of bits needed to represent the integers. It follows from Theorem 3 that every tolerance graph has an integer tolerance representation in which each number in the representation is represented as a string of at most $1 + \lg(1 + 5n2^{5n}) \in O(n)$ bits. Thus the arithmetic operations needed to verify that G is a tolerance graph using this particular tolerance representation take $O(n^3)$ time.

A reasonable encoding of an n-vertex graph requires $\Omega(n)$ space, and so the time required for this verification is polynomial in the size of the encoding of G. It follows that recognizing tolerance graphs is in NP. □

An interesting open problem is whether every tolerance graph has a representation that is substantially more compact than the one proved here. For example, every n-vertex interval graph has an expansive integer representation in which each integer is in $O(n)$ and thus requires merely $O(\log n)$ bits. We conjecture that each tolerance graph has a representation in which each interval and tolerance has only a poly-logarithmic number of bits.

Acknowledgements

These results were obtained at the Workshop on Structured Families of Graphs, Fields Institute, May 2000; the authors thank organizers Derek Corneil, Lorna Stewart, and Jerry Spinrad for their hospitality. We thank an anonymous referee for making insightful comments that improved the clarity of the paper. The first author also thanks Marty Golumbic, Wolfgang Holzmann, Garth Isaak, Mark Kayll and Jerry Spinrad for helpful comments, and Mike Saks for several crucial linear programming observations.

References

[1] K.P. Bogart, P.C. Fishburn, G. Isaak, L. Langley, Proper and unit tolerance graphs, Discrete Appl. Math. 60 (1995) 99–117.
[2] K.P. Bogart, A.N. Trenk, Bipartite tolerance orders, Discrete Math. 132 (1994) 11–22.
[3] V. Chvátal, Linear Programming, W.H. Freeman, New York, 1983.
[4] G. Cramer, Introduction à l'analyse des lignes courbes algébriques, Les frères Cramer et C. Philibert, Genève, 1750.
[5] S. Felsner, Tolerance graphs and orders, J. Graph Theory 28 (1998) 129–140.
[6] M.R. Garey, D.S. Johnson, Computers and Intractibility, W.H. Freeman, San Francisco, 1979.
[7] M.C. Golumbic, A Bibliography on Tolerance Graphs, http://www.cs.biu.ac.il/∼research/tolerance.
[8] M.C. Golumbic, C.L. Monma, A generalization of interval graphs with tolerances, Proceedings of the 13th Southeastern Conference on Combinatorics, Graph Theory and Computing, Cong. Numer. 35 (Utilitas Math., Winnipeg) (1982) 321–331.
[9] M.C. Golumbic, C.L. Monma, W.T. Trotter, Tolerance graphs, Discrete Appl. Math. 9 (1984) 157–170.
[10] M.C. Golumbic, A.N. Trenk, Tolerance graphs, Cambridge University Press, Cambridge, 2003.
[11] J. Hadamard, Resolution d'une question relativ aux determinants, Bull. Sci. Math. 28 (1893) 240–246.
[12] G. Isaak, K.L. Nyman, A.N. Trenk, A hierarchy of classes of bounded bitolerance orders, Ars Combinatoria 69 (2003) 33–53
[13] C. Papadimitriou, Computational Complexity, Addison-Wesley-Longman, Reading, 1994.
[14] A. Schrijver, Theory of Linear and Integer Programming, Wiley, Chichester, 1986.
[15] M. Sipser, Introduction to the Theory of Computation, PWS Publishing, Boston, 1997.