

A new representation of proper interval graphs with an application to clique-width¹

Pinar Heggernes² Daniel Meister² Charis Papadopoulos²

*Department of Informatics
University of Bergen
N-5020 Bergen, Norway*

Abstract

We introduce a new representation of proper interval graphs that can be computed in linear time and stored in $\mathcal{O}(n)$ space. This representation is a 2-dimensional vertex partition. It is particularly interesting with respect to clique-width. Based on this representation, we prove new upper bounds on the clique-width of proper interval graphs.

Keywords: proper interval graphs, representation model, clique-width

1 Introduction

Proper interval graphs are the intersection graphs of intervals of the real line where no interval properly contains another, and they are equivalent to unit interval graphs [18]. This graph class has many applications, such as physical mapping of DNA and genome reconstruction [9,19]. Proper interval graphs have been subject to extensive theoretical study, and there are several representations and many characterisations of them. In this paper we give a new representation of proper interval graphs that can be seen as a generalisation of previous representations.

An important characterisation of proper interval graphs is through proper interval orderings [15]. A proper interval ordering together with the leftmost

¹ This work is supported by the Research Council of Norway through grant 166429/V30.

² Emails: `pinar.heggernes@ii.uib.no`, `daniel.meister@ii.uib.no`,
`charis.papadopoulos@ii.uib.no`

neighbour of each vertex is an $\mathcal{O}(n)$ -space representation of proper interval graphs, that can be computed in linear time [3,7]. Other vertex ordering characterisations of proper interval graphs formulate conditions on neighbourhoods or maximal cliques [10]. During the last decade, many linear-time recognition algorithms for proper interval graphs have been developed. More recent ones generate vertex orderings that are of the above kinds if and only if the input graph is a proper interval graph [2,16,12,17]. Most of these algorithms are elegantly based on special breadth-first search (BFS) strategies. Other recognition algorithms also apply BFS strategies but with a different approach: for every connected component, find a vertex of special kind and run BFS starting with this vertex. A graph is then proper interval if and only if the BFS levels are cliques and the neighbourhoods between consecutive levels satisfy the so called chain property [3,14]. On the representation side, these latter algorithms compute an ordered vertex partition and verify neighbourhood properties. Similar to the linear orderings above, these partitions can be turned into graph representation by adding adjacency information.

For the representation of proper interval graphs that we introduce in this paper, we define a 2-dimensional structure similar to a matrix. The elements, called *bubbles*, are sets of vertices, and they define a partition of the vertex set of the graph. Two vertices are adjacent only if they belong to bubbles appearing in the same column or in consecutive columns. This new representation is called a *bubble model*, and the exact definition is given in Section 2. The two types of representations mentioned above, orderings and vertex partitions, are captured by the bubble model, which means that those representations can be “embedded” into our representation, hence it generalises previous representations.

We apply the bubble model to give two new upper bounds for the clique-width of proper interval graphs. Clique-width is a graph parameter that can measure the complexity of problems. In particular, all problems that can be expressed in a certain type of monadic second order logic can be solved in linear time on graph classes whose clique-width is bounded, provided a corresponding expression is given. Computing the clique-width is NP-hard, even when restricted to complements of bipartite graphs [8]. The clique-width of a graph can be bounded by a function of its treewidth or pathwidth [6,4,8]. This implies that the clique-width of a proper interval graph is at most the size of a maximum clique plus 1. By now, grids are the only known graph class of unbounded clique-width for which the clique-width can be computed in polynomial time [11]. Proper interval graphs have unbounded clique-width [11]. In Section 3, we give two new upper bounds on the clique-width of

proper interval graphs that are not based on pathwidth or maximum clique size. We show that there are graphs on which our results give a better upper bound than the size of a maximum clique. In particular, we show that the clique-width of a proper interval graph is bounded by the size of a maximum independent set, which also enables us to give a tight bound on the clique-width of co-chain graphs. Our new representation is of essential importance for proving this bound.

2 A representation of proper interval graphs – bubble models

A graph G is called *proper interval graph* if every vertex of G can be assigned an interval of the real line such that no interval is properly contained in another, and two vertices are adjacent if and only if their corresponding intervals have a non-empty intersection. A *vertex ordering* σ for G is a permutation of $V(G)$. We write $u \prec_\sigma v$ if u appears before v in σ . Ordering σ is called a *proper interval ordering* if for every triple u, v, w of vertices of G where $u \prec_\sigma v \prec_\sigma w$ and $uw \in E(G)$, $uv, vw \in E(G)$. We call this condition the *umbrella property*.

Theorem 2.1 ([15]) *A graph G is a proper interval graph if and only if G has a proper interval ordering.*

Let A be a finite and non-empty set. A *2-dimensional bubbles structure* \mathcal{B} for A is a 2-dimensional arrangement of *bubbles*, $\langle B_{i,j} \rangle_{1 \leq j \leq k, 1 \leq i \leq r_j}$, and every bubble $B_{i,j}$ contains a subset of A where every object of A appears in exactly one bubble. Some bubbles may be empty. To give an intuition, bubbles are put into a matrix-like setting, and bubble $B_{i,j}$ appears in row i and column j . For every $j \in \{1, \dots, k\}$, bubbles $B_{1,j}, \dots, B_{r_j,j}$ are grouped to the j -th column of \mathcal{B} . Column j starts with bubble $B_{1,j}$ and ends with bubble $B_{r_j,j}$.

Definition 2.2 Let A be a finite non-empty set. Let $\mathcal{B} = \langle B_{i,j} \rangle_{1 \leq j \leq k, 1 \leq i \leq r_j}$ be a 2-dimensional bubbles structure for A . The *graph defined by \mathcal{B}* , denoted as $G(\mathcal{B})$, is defined as follows:

- (1) $G(\mathcal{B})$ has a vertex for every element in A , and
- (2) uv is an edge of $G(\mathcal{B})$ if and only if there are indices i, j, i', j' such that $a_u \in B_{i,j}$ and $a_v \in B_{i',j'}$, where a_u and a_v are the elements of A corresponding to u and v , respectively, and $|j - j'| \leq 1$ and one of the two conditions holds: either $j = j'$ or $(i - i') \cdot (j - j') < 0$.

Adjacent vertices of the graph defined by a 2-dimensional bubbles struc-

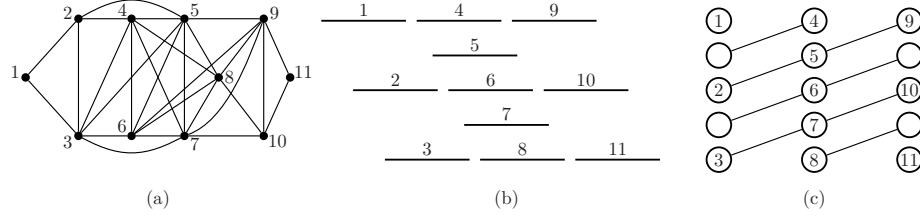


Fig. 1. (a) A proper interval graph G , (b) a proper interval model for G and (c) a bubble model for G .

ture are contained in the same column or in neighbouring columns. It follows directly that vertices that appear in the same column form a clique. Furthermore, vertices in the same bubble are true twins, since they are adjacent and they have the same neighbours in the neighbouring columns. An alternative definition for adjacency of vertices in neighbouring columns is: $u \in B_{i,j}$ and $v \in B_{i',j+1}$ are adjacent if and only if $i > i'$.

Definition 2.3 A *bubble model* for a graph G is a 2-dimensional bubbles structure \mathcal{B} for $V(G)$ such that $G = G(\mathcal{B})$.

Figure 1 gives an example of a proper interval graph, that is represented by a proper interval model and a bubble model. The line segments between bubbles in neighbouring columns emphasise the neighbourhood property. For example, vertex 10 is adjacent to 7 and 8 but not to 4, 5, 6.

Theorem 2.4 A graph is a proper interval graph if and only if it has a bubble model.

The somewhat technical proof of Theorem 2.4 is constructive and results in an $\mathcal{O}(n)$ -time algorithm for computing a bubble model, given the graph as a proper interval ordering with stored leftmost neighbour of each vertex [13]. Note that the algorithm does not output a real bubble model but a space efficient version, that represents consecutive empty bubbles in a column by a single bubble and an occurrence number.

3 Upper bounds on the clique-width of proper interval graphs

In this section we show two different approaches for computing a clique-width expression for a proper interval graph. Clique-width was introduced by Courcelle, Engelfriet, Rozenberg [5]. The *clique-width* of a graph G , denoted by $\text{cwd}(G)$, is defined as the smallest number of labels needed to construct G ,

using the following operations:

- creation of a new vertex v with label i , denoted by $i(v)$
- disjoint union, denoted by \oplus
- changing all labels i to j , denoted by $\rho_{i \rightarrow j}$
- adding edges between all vertices with label i and all vertices with label j , $i \neq j$, denoted by $\eta_{i,j} = \eta_{j,i}$.

An expression built by using the four operations is called *clique-width expression*. If k labels are used in a clique-width expression then it is called a k -expression.

First, we give an upper bound on the clique-width in the size of a maximum independent set. By $\alpha(G)$, we denote the size of a maximum independent set of a graph G .

Theorem 3.1 *Let G be a proper interval graph. Then $\text{cwd}(G) \leq \alpha(G) + 1$. Given G by a proper interval ordering together with the leftmost neighbour of each vertex, an $(\alpha(G) + 1)$ -expression for G can be constructed in $\mathcal{O}(n)$ time.*

Proof. We only prove correctness. Let $\mathcal{B} = \langle B_{i,j} \rangle_{1 \leq j \leq k, 1 \leq i \leq r_j}$ be a bubble model for G . We give an algorithm for constructing a $(k + 1)$ -expression. We use labels $1, \dots, k + 1$, where the first k labels are assigned to the columns of \mathcal{B} , and label $k + 1$ is used to add a new vertex. Our algorithm visits the bubbles of \mathcal{B} row by row in a top-down manner and within a row from left to right. Vertices in the same bubble are treated sequentially. This defines an ordering on the vertices of G . Let x be the first vertex. The expression for $G[\{x\}]$ is $1(x)$. Now, let y be a vertex of G , and assume that the expression t for the graph on the vertices preceding y has already been defined. Let y be in a bubble in column j . The expression for the graph induced by the vertices not succeeding y then is $\rho_{k+1 \rightarrow j}(\eta_{k+1,j+1}(\eta_{k+1,j}((k+1)(y) \oplus t)))$. Correctness follows directly from the properties of bubble models, since the neighbours of y are only in the subgraph in columns j and $j + 1$. Note that, if $j = k$, the subexpression involving column $j + 1$ is obsolete. It can be shown that there is a bubble model for G with (exactly) $\alpha(G)$ columns. This implies $\text{cwd}(G) \leq \alpha(G) + 1$. ■

As an interesting consequence, we mention that Theorem 3.1 gives a bound on the clique-width of co-chain graphs, which form a subclass of proper interval graphs³. Combining the already known results on the clique-width of chain graphs [11] and complements of graphs [6], co-chain graphs are graphs of

³ We refer to [1,10] for the definitions of graph classes mentioned in this paragraph.

clique-width at most 6. Hence, by the following corollary, we are able to give a better bound. Furthermore, the bound given below is tight, since there are co-chain graphs that are not cographs. Observe that every co-chain graph G has $\alpha(G) \leq 2$, since G is the complement of a bipartite graph.

Corollary 3.2 *For every co-chain graph G , $\text{cwd}(G) \leq 3$, and a 3-expression can be constructed in $\mathcal{O}(n + m)$ time.*

Our second bound in the clique-width uses the notion of groups. Let G be a connected proper interval graph with bubble model \mathcal{B} . A set of non-empty bubbles of a given column j is called *group* if the vertices of the bubbles have the same neighbourhood in G with respect to the vertices in column $j + 1$. By the definition of the bubble models, we know that every vertex of G belongs to exactly one group. Moreover observe that only consecutive non-empty bubbles may belong to the same group. That is, if $B_{i,j}$ and $B_{i',j}$ belong to the same group such that $i < i'$ then every non-empty bubble $B_{i'',j}$ for which $i < i'' < i'$ also belongs to the group. For instance, in Figure 1 every non-empty bubble of the first column defines a group by itself, whereas in the second column there are three groups defined as $\{\{4\}, \{5, 6\}, \{7, 8\}\}$, and all vertices in the third column define a single group. Let v be a vertex of G and let $L(\mathcal{B}, v)$ be the set of vertices that are in a column to the left of v , in a bubble above v or in the same bubble as v (excluding v) in \mathcal{B} . Let $n_{\mathcal{B}}(v)$ be the number of groups consisting of vertices from $L(\mathcal{B}, v)$ containing at least one neighbour of v . The *group number* of G with respect to \mathcal{B} , denoted by $\varphi(G, \mathcal{B})$, is defined as $\varphi(G, \mathcal{B}) =_{\text{def}} \max_{v \in V} n_{\mathcal{B}}(v)$. By $\omega(G)$, we denote the size of a maximum clique of a graph G .

Theorem 3.3 *Let G be a proper interval graph with bubble model \mathcal{B} . Then $\text{cwd}(G) \leq \varphi(G, \mathcal{B}) + 2 \leq \omega(G) + 1$. A $(\varphi(G, \mathcal{B}) + 2)$ -expression for G can be computed in $\mathcal{O}(n + m)$ time.*

Proof. We only prove correctness here. We prove $\varphi(G, \mathcal{B}) + 1 \leq \omega(G)$. Let \mathcal{B} be a bubble model for G . Consider $n_{\mathcal{B}}(v)$ for a vertex v . We call *active groups of v* those groups consisting of vertices from $L(\mathcal{B}, v)$ and that contain at least a neighbour of v . We show that the vertices of the active groups of v together with v form a clique in G for each vertex v . Observe that every vertex of an active group is adjacent to v and form a clique in G . Consider the active group F of v that is furthest away from v in \mathcal{B} . Then, for any other group F' which is between F and v we know that every vertex of F is adjacent to every vertex of F' and every vertex of F' is adjacent to v . This means that every group between F and v is active, and by induction, it follows that the

vertices of the active groups of v together with v form a clique in G . This implies $\varphi(G, \mathcal{B}) + 1 \leq \omega(G)$.

Now, we show that there exists a $(\varphi(G, \mathcal{B}) + 2)$ -expression for G . We iteratively construct an expression by adding vertices in their order in \mathcal{B} : from left to right column by column, and within a column from top row to bottom row. Let v be a vertex and assume an expression t for the graph induced by $L(\mathcal{B}, v)$ has been defined. Let $q = n_{\mathcal{B}}(v)$. Assume that a label is assigned to every group and all the vertices of the same group have the same label. We change the labels of all the vertices in non-active groups to label 1 and then assign a distinct label from $\{2, \dots, q+1\}$ for every active group of v . This can be done by using the appropriate ρ operation at most $q+1$ times. Next we use label $q+2$ to add v and join the vertices of label $i \in \{2, \dots, q+1\}$ with the vertex v ; $(q+2)v$ and $\eta_{i,q+2}$ define the appropriate operations. Finally, we put v into the group that v belongs to by relabelling (if v an existing group). This shows $\text{cwd}(G) \leq \varphi(G, \mathcal{B}) + 2$. ■

By Theorems 3.1 and 3.3 we obtain the following result.

Corollary 3.4 *For a proper interval graph G ,*
 $\text{cwd}(G) \leq \min\{\alpha(G) + 1, \varphi(G) + 2\} \leq \omega(G) + 1$.

We point out that there are proper interval graphs G for which $\varphi(G)$ or $\alpha(G)$ is significantly smaller than $\omega(G)$. An easy example can be derived from the graph G shown in Figure 1 by extending the three columns of the bubble model so that $2\varphi(G) = \omega(G)$. Hence our bounds are better than the previously known bound on clique-width of proper interval graphs.

References

- [1] A. Brandstädt, V. B. Le, and J. P. Spinrad. Graph classes : A survey. SIAM Monographs on Discrete Mathematics and Applications, 1999.
- [2] D. G. Corneil. A simple 3-sweep LBFS algorithm for the recognition of unit interval graphs. Disc. Appl. Math. 138, pp. 371–379, 2004.
- [3] D. G. Corneil, H. Kim, S. Natarajan, S. Olariu, and A. P. Sprague. Simple linear time recognition of unit interval graphs. Inf. Proc. Let. 55, pp. 99–104, 1995.
- [4] D. G. Corneil and U. Rotics. On the relationship between clique-width and treewidth. SIAM J. Comput. 34, pp. 825–847, 2005.
- [5] B. Courcelle, J. Engelfriet, and G. Rozenberg. Handle-rewriting hypergraph grammars. J. Comput. System Sci. 46, pp. 218–270, 1993.

- [6] B. Courcelle and S. Olariu. Upper bounds to the clique width of graphs. *Discrete Applied Mathematics* 101, pp. 77–114, 2000.
- [7] X. Deng, P. Hell, and J. Huang. Linear-time representation algorithms for proper circular-arc graphs and proper interval graphs. *SIAM J. Comput.* 25, pp. 390–403, 1996.
- [8] M. R. Fellows, F. A. Rosamond, U. Rotics, and S. Szeider. Clique-width Minimization is NP-hard. *Proceedings of the thirty-eighth annual ACM Symposium on Theory of Computing, STOC’06*, pp. 354–362, 2006.
- [9] P. W. Goldberg, M. C. Golumbic, H. Kaplan, and R. Shamir. Four strikes against physical mapping of DNA. *J. Comput. Bio.* 2, pp. 139–152, 1995.
- [10] M. C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs*. Second edition. *Annals of Discrete Mathematics* 57. Elsevier, 2004.
- [11] M. C. Golumbic and U. Rotics. On the clique-width of some perfect graph classes. *Int. J. Found. Comp. Sci.* 11, pp. 423–443, 2000.
- [12] P. Hell and J. Huang. Certifying LexBFS Recognition Algorithms for Proper Interval Graphs and Proper Interval Bigraphs. *SIAM J. Disc. Math.* 18, pp. 554–570, 2005.
- [13] P. Heggernes, D. Meister, and Ch. Papadopoulos. A new representation of proper interval graphs with an application to clique-width. Technical report no. 354, Department of Informatics, University of Bergen, Norway, 2007.
- [14] C. M. Herrera de Figueiredo, J. Meidanis, and C. Picinin de Mello. A linear-time algorithm for proper interval graph recognition. *Inf. Proc. Lett.* 56, pp. 179–184, 1995.
- [15] P. J. Looges and S. Olariu. Optimal greedy algorithms for indifference graphs. *Computers & Mathematics with Applications* 25, pp. 15–25, 1993.
- [16] D. Meister. Recognition and computation of minimal triangulations for AT-free claw-free and co-comparability graphs. *Disc. Appl. Math.* 146, pp. 193–218, 2005.
- [17] B. S. Panda and S. K. Das. A linear time recognition algorithm for proper interval graphs. *Inf. Proc. Lett.* 87, pp. 153–161, 2003.
- [18] F. S. Roberts. Indifference graphs. In: F. Harary (Ed.), *Proof techniques in graph theory*, pp. 139–146, Academic Press, New York, 1969.
- [19] J. Spinrad. *Efficient Graph Representations*. Fields Institute Monographs 19. ACM, 2003.