[Rot98] U. Rotics. Polynomial time ($o(|v|^7)$) algorithm for the 3-clique-width problem. Preprint, 1998.

[Spi92] J. Spinrad. P_4-trees and substitution decomposition. *Discrete Appl. Math.*, 39:263–291, 1992.

[Wan94] E. Wanke. k-NLC graphs and polynomial algorithms. *Discrete Appl. Math.*, 54:251–266, 1994.

[CMR99] B. Courcelle, J.A. Makowsky, and U. Rotics. Hierarchical decompositions of relational structures and applications to efficient algorithms. in preparartion, 1999.

[CO98] B. Courcelle and S. Olariu. Upper bounds to the clique-width of graphs. submitted for publication (http://dept-info.labri.u-bordeaux.fr/∼courcell/ActSci.html), 1998.

[Cou93] B. Courcelle. Graph grammars, monadic second-order logic and the theory of graph minors. *Contemporary Mathematics*, 147:565–590, 1993.

[GRT97] V. Giakoumakis, , F. Roussel, and H. Thuillier. On P_4-tidy graphs. *Discrete Mathematics and Theoretical Computer Science*, 1:17–41, 1997.

[GV97] V. Giakoumakis and J. Vanherpe. On extended P_4-reducible and extended P_4-sparse graphs. *Theoret. Comput. Sci.*, 180:269–286, 1997.

[JO89] B. Jamison and S. Olariu. P_4-reducible graphs a class of tree representable graphs. *Studies Appl. Math.*, 81:79–87, 1989.

[JO92a] B. Jamison and S. Olariu. A linear-time recognition algorithm for P_4-sparse graphs. *SIAM J. Comput.*, 21:381–406, 1992.

[JO92b] B. Jamison and S. Olariu. A unique tree representation for P_4-sparse graphs. *Discrete Appl. Math.*, 35:115–129, 1992.

[JO95a] B. Jamison and S. Olariu. A linear-time algorithm to recognize P_4-reducible graphs. *Theoret. Comput. Sci.*, 145:329–344, 1995.

[JO95b] B. Jamison and S. Olariu. Linear-time optimization algorithms for P_4-sparse graphs. *Discrete Appl. Math.*, 61:155–175, 1995.

[Bab98b] L. Babel. Recognition and isomorphism of tree–like P_4–connected graphs. Preprint, 1998.

[BM83] H. Buer and R.H. Möhring. A fast algorithm for the decomposition of graphs and posets. *Math. Oper. Res.*, 8:170–184, 1983.

[BO95] L. Babel and S. Olariu. On the isomorphism of graphs with few P_4s. In M. Nagl, editor, *Graph Theoretic Concepts in Computer Science, 21th International Workshop, WG'95*, volume 1017 of *Lecture Notes in Computer Science*, pages 24–36. Springer Verlag, 1995.

[BO98a] L. Babel and S. Olariu. Domination and steiner tree problems on graphs with few P_4s. To appear in the proceedings of WG98, 1998.

[BO98b] L. Babel and S. Olariu. On the p–connectednessof graphs – a survey. Preprint, 1998.

[CER93] B. Courcelle, J. Engelfriet, and G. Rozenberg. Handle-rewriting hypergraph grammars. *J. Comput. System Sci.*, 46:218–270, 1993.

[CH94] A. Cournier and M. Habib. A new linear algorithm for modular decomposition. *Lecture Notes in Computer Science*, 787:68–84, 1994.

[CLS81] D.G. Corneil, H. Lerchs, and L. Stewart. Complement reducible graphs. *Disc. Appl. Math.*, 3:163–174, 1981.

[CMR98a] B. Courcelle, J.A. Makowsky, and U. Rotics. Linear time solvable optimization problems on certain structured graph families, extended abstract. To appear in the proceedings of WG98, 1998.

[CMR98b] B. Courcelle, J.A. Makowsky, and U. Rotics. On the fixed parameter complexity of graph enumeration problems definable in monadic second order logic. in preparartion, 1998.

the first case there is no other vertex z of L such that w and z occur in the same equivalence class of R_{red}, or else z and w are adjacent to both u and v, a contradiction. In the second case there is no vertex z of $L - P_u$ such that z and a vertex of P_u occur in the same equivalence class of R_{red}, or else u must be adjacent to z, a contradiction.
We conclude that for each vertex of $S_{3,blue}$, or for each pair of vertices of $S_{3,blue}$, there is a corresponding vertex of L or a corresponding set of vertices of L, which occur in an equivalence class of R_{red} which is different from all the equivalence classes in which the other vertices of L occur. It follows the vertices of L occur in at least $|S_{3,blue}|/2 \geq n/72$ different equivalence classes of R_{red}. But this contradicts the assumption that the number of equivalence classes of R_{red} is less than $n/72$. □

Theorem 4. *There is a class $\mathcal{C}$ containing infinitely many split graphs, such that for every graph $G = \langle V, E\rangle \in \mathcal{C}$ $cwd(G) \geq (sqrt2|V| - 1)/72$.*

Proof:

Let G_n denote the clique of n vertices, and let $\mathcal{C}$ be the class of split graphs of cliques defined by: $\mathcal{C} = \{splt(G_n) : n \in N,\ n \geq 20\}$. From Theorem 5 and Lemma 4 above it follows that for $n \in N$ $n \geq 20$, $cwd(splt(G_n)) \geq n/72$. Since for every clique G_n $splt(G_n)$ is a split graph, it follows that the class $\mathcal{C}$ satisfies the conditions of the theorem. □

Acknowledgments

We are indebted to Luitpold Babel who made us aware of the $(q, q-3)$ graphs and the prime graphs associated with their corresponding modular decompositions.

References

[Bab98a] L. Babel. On graphs with simple P_4–structure. Preprint, 1998.

CASE 1: Suppose that either $|S_1| < 10n/18$ or $|S_2| < 10n/18$. We assume that $|S_1| < 10n/18$, (the case when $|S_2| < 10n/18$ can be handled similarly). From the definition of S_1 it follows that the number of red vertices is at most: $|S_1|+(|S_1|\times(|S_1|-1))/2$. Since $|S_1| < 10n/18$ we get (by substituting $10n/18$ instead of $|S_1|$ in the above formula) that for $n \geq 20$ the number of the red vertices is at most: $0.165n^2$. However, since the number of red vertices greater than $(n^2)/6$ there must be at least $0.166n^2$ red vertices, a contradiction.

CASE 2: Suppose that both $|S_1| \geq 10/18$ and $|S_2| \geq 10n/18$. Let S_3 denote the intersection of S_1 and S_2. Since the total number of vertices occuring either in S_1 or in S_2 (or in both of them) is equal to n, we get that at least $n/18$ vertices occurs both in S_1 and in S_2. In other words we get that $|S_3| \geq n/18$. Let $S_{3,red}$ ($S_{3,blue}$) denote the set of all the red (blue) vertices occuring in S_3.

CASE 2.1: Suppose that $|S_{3,red}| \geq n/36$. Suppose that there exists three vertices (say x, y and z) in $S_{3,red}$ such x, y and z are in the same equivalence class of R_{red}. Since x is in S_2 there exist a blue vertex w such that w is adjacent to x, since the degree of w is 2, w is not adjacent to either y or z (say z). Hence, there is a blue vertex w which distinguishes x and z. But this contradicts the assumption that x and z are in the same equivalence class of R_{red}.

Hence, every 3 vertices of $S_{3,red}$ can not be in the same equivalence class $of R_{red}$, which implies that the vertices of $S_{3,red}$ occur in at least $|S_{3,red}|/2 \geq n/72$ different equivalence classes of R_{red}. But this contradicts the assumption that the number of equivalence classes of R_{red} is less than $n/72$.

CASE 2.2: Suppose that $|S_{3,red}| < n/36$. Since $|S_3| \geq n/18$ it must be that $|S_{3,blue}| \geq n/36$. Let L be the set of vertices defined by:

$$L = \{v | v \in E, v \text{ is red and } v \text{ is adjacent to some vertex } u \in S_{3,blue}\}.$$

For each vertex u in $S_{3,blue}$ either there exist another vertex v in $S_{3,blue}$ such that u and v have a common neighbor w in L, or the set P_u of all the vertices of L which are adjacent to u is not adjacent to any other vertex in $S_{3,blue}$. In orem. □

6 Some other graph classes of unbounded clique–width

In this section we show how our technique presented in section 5 above can be applied to other graph classes. In particular we show that the split graphs and all the graph classes which contains them are of unbounded clique–width.

Definition 20 (Split graph of a clique $splt(G)$**)** *Let* $G = \langle V, E\rangle$ *be a clique, the* split graph of *of* G*, denoted* $splt(G)$*, is is defined as the graph* $\langle V', E'\rangle$*, such that:* $V' = V \cup E$ *and*

$$E' = E \cup \{(v,e) \mid v \in V, e \in E \text{ and } v \text{ is one of the endpoints of } e\}$$

Lemma 4 *Let* $n \in N$ *be such that* $n \geq 20$*, let* $G = \langle V, E\rangle$ *be a clique of* n *vertices, and let* $G' = splt(G)$ *be the split graph of* G*, then* $2colw(G') \geq n/72$*.*

Proof:
Let $n \in N$ be such that $n \geq 20$, let $G = \langle V, E\rangle$ be a clique of n vertices, and let $G' = \langle V', E'\rangle$ be the split graph of G. Suppose that $2colw(G') < n/72$. Then there is a partition of the vertices of G' into two disjoint sets V'_{red} and V'_{blue} of red and blue vertices respectively, such that $n(n+1)/6 \leq |V'_{red}| \leq n(n+1)/3$, and the number of equivalence classes of R_{red} (see definition 17 above) is less then $n/72$.
We say that a vertex v of G' is *spanned–by–red–edges* (*spanned–by–blue–edges*) if $v \in V$ and there is at least one red (blue) vertex $e \in E$ such that v is adjacent to e in G'. Note that a vertex v can be spanned–by–red–edges and also spanned–by–blue–edges. Let S_1 (S_2) denote the set of all the vertices of G' which are spanned–by–red–edges (spanned–by–blue–edges). We consider the following two cases:

Since we have considered all possible cases, and got a contradiction in each case, we conclude that our assumption that $2colw(G) < n/3$ was not correct. In other words we conclude that $2colw(G) \geq n/3$. □

Theorem 2. *There is a class $\mathcal{C}$ containing infinitely many $(6,3)$ graphs, such that for every graph $G = \langle V, E\rangle \in \mathcal{C}$, $cwd(G) \geq \sqrt{|V|/27}$.*

Proof:

Let $\mathcal{C}$ be the class of extended square grids defined by: $\mathcal{C} = \{H_{n,1,1} : n \in N,\ n \geq 3\}$. From Theorem 5 and Lemma 2 above it follows that for $n \in N$ $n \geq 3$, $cwd(H_{n,1,1}) \geq n/3$. Since by Fact 2 above every $H_{n,1,1}$ is a $(6,3)$ graph the class $\mathcal{C}$ satisfies the conditions of the theorem. □

5.3 (q, q) graphs for $q \geq 4$

In this section we show that the class of (q,q) graphs are not of bounded clique–width. For that we shall consider the extended square grids $H_{n,q,q}$ (see definition 19) above which are (q,q) graphs by Fact 3 above.

Lemma 3 *For every $n \in N$ such that $n \geq 3$, and for every $q \geq 1$, $2colw(H_{n,q,q}) \geq n/3q$.*

Proof:

The proof is similar to the proof of Lemma 2 above. □

Theorem 3. *For every $q \geq 4$, there is a class $\mathcal{C}(q)$ containing infinitely many (q,q) graphs, such that for every graph $G = \langle V, E\rangle \in \mathcal{C}$ $cwd(G) \geq \sqrt{|V|/27q^3}$.*

Proof:

For $q \geq 4$ let $\mathcal{C}(q)$ be the class of extended square grids defined by: $\mathcal{C}(q) = \{H_{n,q,q} : n \in N,\ n \geq 3\}$. From Theorem 5 and Lemma 3 above it follows that for $n \in N$ $n \geq 3$, $cwd(H_{n,q,q}) \geq n/3q$. Since by Fact 3 above every $H_{n,q,q}$ is a (q,q) graph the class $\mathcal{C}(q)$ satisfies the conditions of the the-of L occurs in different equivalence classes of R_{red}, which implies that the number of these equivalence classes is at least n, a contradiction.

CASE 2: Suppose that for some i in $\{1, 3, \ldots, 2n-1\}$ there is there is no vertex $v \in S$ occuring in column i of G. There are two possible cases:

CASE 2.1: Suppose that all the vertices of column i are red. We say that v is a red (resp. blue) row–alternating vertex if v is a red (resp. blue) vertex such that there is a blue (resp. red) vertex u adjacent to v such that v and u are in the same row in G. We construct a set Q by the following procedure:

(i) Set $j = 1$, and set $Q = \emptyset$.

(ii) If $j = 2n + 1$ stop.

(iii) If not all the vertices at row j are red then let v be any red row–alternating vertex occuring in row j. Set $Q = Q \cup \{v\}$.

(iv) If all the vertices at row j are red and not all the vertices at row $j-1$ are red then let v by any red vertex occuring in row j such that its neighbor at row $j-1$ is blue. Set $Q = Q \cup \{v\}$.

(v) set $j = j + 2$ and go to step (ii) above

Suppose that the number of vertices occuring in Q is $< n/3$ it follows that there more than $2n/3$ rows in which we could not choose a vertex to add to Q by the above procedure. In other words there are at least $2n/3$ odd rows j in {3,5, ...2n-1} such that all the vertices at row j and row $j-1$ are red. Counting the number of red vertices in these rows, we get that $|V_{red}| \geq 2n^2 - 2n/3 > 2n^2 - 4n/3$, a contradiction.

Hence $|Q| \geq n/3$. But since all the vertices of Q occur in different equivalence classes of R_{red} it follows that the number of equivalence classes of R_{red} is at least $n/3$ a contradiction.

CASE 2.2: Suppose that all the vertices of column i are blue. In this case we can use a similar argument to case 2.1 above to show that this case is not possible.

Definition 19 (Extended square grid $H_{n,l,m}$) *Let G_n be an $n \times n$ square grid. The* extended square grid *$H_{n,l,m}$ is the graph obtained from G_n by replacing every edge (u, v) occuring in a row (resp. column) of G_n by a simple path of length $l+1$ (resp. $m+1$) such that the endpoints of the path are u and v.*

Counting the number of possible P_4s in every q vertices of an extended square grid, it can be showed that:

Fact 2 *For $n \geq 2$ every extended square grid $H_{n,1,1}$ is a $(6,3)$ graph.*

Fact 3 *For $n \geq 2$ every extended square grid $H_{n,q,q}$ is a (q,q) graph.*

Hence, to show that the class of $(6,3)$ graphs is not of bounded clique–width we shall show that:

Lemma 2 *For every $n \in N$ such that $n \geq 3$, $2colw(H_{n,1,1}) \geq n/3$.*

Proof:
Let $n \in N$ be such that $n \geq 3$, and let $G = \langle V, E \rangle$ be the extended square grid $H_{n,1,1}$. First note that $|V| = 3n^2 - 2n$. Suppose that $2colw(G) < n/3$. Then there is a partition of the vertices of G into two disjoint sets V_{red} and V_{blue} of red and blue vertices respectively, such that $n^2 - 2n/3 \leq |V_{red}| \leq 2n^2 - 4n/3$, and the number of equivalence classes of R_{red} (see definition 17 above) is less then $n/3$. Note that G has 2n-1 rows and 2n-1 columns. We number the rows and columns of G from 1 to $2n-1$. Note that all the even rows and columns contains just vertices of degree 2. We say that v is a red (resp. blue) column–alternating vertex if v is a red (resp. blue) vertex such that there is a blue (resp. red) vertex u adjacent to v such that v and u are in the same column in G. Let S denote the set of all the red column–alternating vertices. We consider the following two cases:
CASE 1: Suppose that for every i in 1,3, …,2n-1 there is one vertex $v \in S$ occuring in column i in G. Then there is a set $L = \{v_1, v_3, \ldots, v_{2n-1}\}$ of n vertices of S such that for every i, j in $\{1, 3, \ldots, 2n-1\}$ such that $i \neq j$ v_i and v_j occur in different columns in G. It is easy to see that all the vertices

at the same distance from column i (in both directions), then chose arbitrarily either u or v. Set $Q = Q \cup \{v\}$.

(iv) set $j = j + 1$ and go to step (ii) above

Suppose that the number of vertices occuring in Q is $< n/3$ it follows that there are more than $2n/3$ red rows. Hence, $|V_{red}| > 2n^2/3$, a contradiction. Therefore we can assume that $|Q| \geq n/3$.
Suppose that two vertices occuring in Q (say x and y) are in the same equivalence class of R_{red}. We can assume without loss of generality that x and y occurs in columns j_1 and j_2 respectively, such that either $j \leq j_1 \leq j_2$, or $\leq j_2 \leq j_1 \leq j$. Let i denote the row in which y occurs. In both cases the blue vertex w which occurs in row i and is adjacent to y distinguishes x and y. But this contradicts the assumption that x and y are in the same equivalence class of R_{red}.
We conclude that all the vertices of Q occur in different equivalence classes of R_{red}, which implies that the number of equivalence classes of R_{red} is $\geq n/3$, a contradiction.
CASE 2.2: Suppose that all the vertices of column i are blue. In this case we can use a similar argument to case 2.1 above to show that this case is not possible.

Since we have considered all possible cases, and got a contradiction in each case, we conclude that our assumption that $2colw(G) < n/3$ was not correct. In other words we conclude that $2colw(G) \geq n/3$. □

Remark 1 *Using a more complicated argument it can be shown that for every $n \times n$ square grid G, $cwd(G) \geq n$. Since it is not hard to see that every $n \times n$ square grid can be build by an $n+2$–expression, it follows that for every such graph the clique–width is between n and $n+2$. We suspect that for $n \geq 3$ the clique width of an $n \times n$ square grid is exactly $n+2$, but we leave it as an open question.*

CASE 1: Suppose that for $1 \leq i \leq n$ there is at least one vertex $v \in S$ occuring in column i in G. Then there is a set $L = \{v_1, \ldots, v_n\}$ of n vertices of S such that for $1 \leq i \leq n$ and for $1 \leq j \leq n$ such that $i \neq j$ v_i and v_j occur in different columns in G.

Suppose that there are 3 vertices x, y and z of L which are in the same equivalence class of R_{red}. Clearly two of these 3 vertices (say x and y) occur at two non-consecutive columns in G. Let i denote the columns of G which contains x. Since x is in S there is a blue vertex w which is adjacent to x and is included in column i. Since the columns of x and y are not consecutive it follows that w is not adjacent to y. In other words the blue vertex w distinguishes x and y. But this contradicts the assumption that x and y are in the same equivalence class of R_{red}.

Hence, every 3 vertices of L can not be in the same equivalence class of R_{red}, which implies that the vertices of L occur in at least $n/2$ different equivalence classes of R_{red}. But this contradict the assumption that the number of equivalence classes of R_{red} is less than $n/3$.

CASE 2: Suppose that for some $1 \leq i \leq n$, there is no vertex $v \in S$ occuring in column i of G. There are two possible cases:

CASE 2.1: Suppose that all the vertices of column i are red. We say that v is a red (blue) row–alternating vertex if v is a red (blue) vertex such that there is a blue (red) vertex u adjacent to v such that v and u are in the same row in G. Let P denote the set of all the red row–alternating vertices. We construct a set $Q \subset P$ by the following procedure:

(i) Set $j = 1$, and set $Q = \emptyset$.

(ii) If $j = n + 1$ stop.

(iii) If not all the vertices at row j are red then let v be the red row–alternating vertex occuring in row j, such that the column in which v occur is the closest column to column i, which contains a red row–alternating vertex. If there are two such vertices v and u occuring

The following fact shows that the other direction of Theorem 5 above does not hold:

Fact 1 *There is a graph* $G = \langle V, E \rangle$ *such that the clique–width of* G *is* $\geq \sqrt{|V|}/6$ *and the 2–color–width of* G *is equal to 1.*

Proof:
Let $G = \langle V, E \rangle$ be the graph obtained by taking the disjoint union of n^2 isolated vertices with and an $n \times n$ square grid. By Lemma 1 below $cwd(G) \geq n/3$ which implies that $cwd(G) \geq \sqrt{|V|}/6$. By coloring all the n^2 isolated vertices of G with red and all the n^2 vertices of the square grid of G with blue, we we obtain that the number of the equivalence classes of R_{red} is one, which implies that the 2–color–width of G is equal to 1. □

5.2 $(6, 3)$ graphs

In this section we shall show that the class of $(6, 3)$ graphs is of unbounded clique–width. Before handling the case of the $(6, 3)$ graphs we show how our techniques can be used for the class of square grids:

Lemma 1 *Let* $n \in N$ *be such that* $n \geq 4$*, and let* G *be an* $n \times n$ *square grid, then* $2colw(G) \geq n/3$*.*

Proof:
Let $n \in N$ be such that $n \geq 4$, and let G be an $n \times n$ square grid, suppose that $2colw(G) < n/3$. Then there is a partition of the vertices of G into two disjoint sets V_{red} and V_{blue} of red and blue vertices respectively, such that $n^2/3 \leq |V_{red}| \leq 2n^2/3$, and the number of equivalence classes of R_{red} (see definition 17 above) is less then $n/3$.
We say that v is a red (blue) column–alternating vertex if v is a red (blue) vertex such that there is a blue (red) vertex u adjacent to v such that v and u are in the same column in G. Let S denote the set of all the red column–alternating vertices. We consider the following two cases:

finally obtain a sub–tree T_x of T such that $nl(T)/3 \leq nl(T_x) \leq 2nl(T)/3$. □

We say that a class of graphs $\mathcal{C}$ is of unbounded clique–width (2–color–width) if there is no fixed $k \in N$, such that for every graph $G \in C$, the clique–width (the 2–color–width) of G is at most k.

Theorem 5 *For every graph G, if the 2–color–width of G is $> k$ then the clique–width of G is $> k$.*

Proof:

Let $G = \langle V, E \rangle$ be any graph such that $2col(G) > k$. Suppose that $cwd(G) \leq k$. Then there is a k–expression t which defines G. By proposition 11 above there is a node a in $tree(t)$ (see definition 7 above) such that $nl(tree(t))/3 \leq nl(tree(a,t)) \leq 2nl(tree(t))/3$, where $tree(a,t)$ denotes the sub–tree of tree(t) rooted at a (cf. definition 8 above) and $nl(tree(t))$ denotes the number of leaves in $tree(t)$.

Clearly, $nl(tree(t)) = |V|$. Let V_{red} and V_{blue} be the partition of the vertices of $|V|$ into two disjoint sets such that all the vertices of V_{red} are colored with red and occurs in $tree(a,t)$ and all the vertices of V_{blue} are colored with blue and do not occur in $tree(a,t)$. Since $nl(tree(t)) = |V|$ and $nl(tree(a,t)) = |V_{red}|$, it follows that $|V|/3 \leq |V_{red}| \leq 2|V|/3$. Let R_{red} be the relation defined in definition 17 above. Since $2colw(G) > k$, there are at least $k+1$ equivalence classes in R_{red}. Hence there is a set $S = \{v_1, \ldots, v_{k+1}\}$ of $k+1$ vertices of G such that for $1 \leq i \leq k+1$ and for $1 \leq j \leq k+1$ such that $i \neq j$, v_i and v_j do not occur in the same equivalence class of R_{red}.

Suppose that v_i and v_j has the same label at a, it follows that there is no blue vertex w which distinguishes v_i and v_j, a contradiction to the assumption that v_i and v_j are in different equivalence classes of R_{red}. Hence, all the vertices occuring at S must have different labels at a. In other words the vertices of S are labeled with $k+1$ different labels at a. But this contradicts the assumption that t is a k–expression. □

Definition 17 (The equivalence relation R_{red}) *Let $G = \langle V, E\rangle$ be a graph, and let V_{red} and V_{blue} be a partition of V into two disjoint sets of vertices colored by red and blue respectively. We define the relation R_{red} such that a pair of vertices (u, v) is in R_{red} if and only if u and v are both red and there is no blue vertex x which distinguishes u and v. Clearly R_{red} is an equivalence relation.*

Definition 18 (The 2–color–width of a graph G, $2colw(G)$) *Let $G = \langle V, E\rangle$ be any graph, the* 2–color–width *of G, denoted as $2colw(G)$ is defined as the smallest number $l \in N$, such that there is a partition of the vertices of G into two disjoint sets V_{red} and V_{blue} such that $|V|/3 \leq |V_{red}| \leq 2|V|/3$ and R_{red} has l equivalent classes.*

Recall that for any tree T and an internal node a of T we denote by T_a the sub–tree of T rooted at a. For every tree T, we denote by $nl(T)$, the number of leaves of T.

In proving Theorem 5 below we shall use the following proposition:

Proposition 11 *Let $T = \langle V, E\rangle$ be any binary tree, then there is an internal vertex of the tree a, such that $nl(T)/3 \leq nl(T_a) \leq 2nl(T)/3$.*

Proof:

Let $T = \langle V, E\rangle$ be any binary tree, let a be the highest node in T which has two sons, and let b and c be the sons of a. If both $nl(T_b) \geq nl(T)/3$ and $nl(T_c) \geq nl(T)/3$ then we are done since both T_b and T_c satisfies the conditions of the proposition.

Hence, we can assume without loss of generality that $nl(T_b) < nl(T)/3$ and $nl(T_c) > 2nl(T)/3$. Let d be the highest node in T_c which has two sons (d may be equal to c), and let e and f be the sons of d in T_c. We assume without loss of generality that $nl(T_e) \leq nl(T_f)$.

If $nl(T_f) \leq 2nl(T)/3$ we are done, since T_f satisfies the conditions of the proposition.

Hence we assume that $nl(T_f) > 2nl(T)/3$. Noting that $nl(T_f) < nl(T_c)$, it follows that repeating the above argument at most $nl(T)/3$ times we shall at most q vertices have clique–width $\leq q$, by proposition 3 above it follows that G has clique–width $\leq q$. A q-expression defining G can be constructed in time $O(|V|+|E|)$ as follows:

(i) Construct the modular decomposition of G, $T(G)$ in time $O(|V|+|E|)$ by classical methods, as shown in [GV97].

(ii) From the modular decomposition $T(G)$ construct an expression consisting of a sequence of vertex substitutions which defines G, as shown in the proof of proposition 3 (see [CMR98a]). Since the number of vertices in $T(G)$ is $O(|V|)$ (as proved in [Spi92]), this step can be done in time $O(|V|)$.

(iii) Convert the expression of vertex substitutions obtained at the previous step, to a q-expression for G as shown in the proof of proposition 2 (see [CMR98a]). This step can be done in time $O(|V|)$, since each graph H used in the substitutions is either an edgeless graph, a clique, a prime spider, a disc, a prime p–tree or a graph with at most q vertices, a q-expression which defines H can be constructed in $O(|V(H)|)$ time, as can be shown easily for the first 2 cases and was shown in [CMR98a] for the prime spiders, in Propositions 5 and 6 above for the discs or in propositions 7, 8, 9 and 10 for the prime p–trees.

□

5 $(6,3)$ and (q,q) for $q \geq 4$ graphs are of unbounded clique–width

5.1 The 2–color width of graphs

We say that a vertex x *distinguishes* y *and* z if x is adjacent to y and is not adjacent to z, or vise versa.

Proof:

Follows immediately from claim 3 above since the 6–expression t_k defines the spiked p–chain Q_k. For the complexity, since the 6–expression t_k can be constructed in k steps using claim 3 above and in each step the amount of additional work is bounded by a constant, it follows that t_k can be constructed in $O(|V|)$ time. □

Proposition 10 *Every complement of a spiked p–chain Q_k has clique–width ≤ 6 and a 6–expression defining it can be constructed in time $O(|V|)$.*

Proof:

Similar to the proof of proposition 9 above. □

4 $(q, q-3)$ graphs for $q \geq 7$ are of clique–width $\leq q$

In this section we show that:

Theorem 1. *For every $(q, q-3)$ graph G such that $q \geq 7$, G has clique–width $\leq q$, and a q–expression defining it can be constructed in time $O(|V|+|E|)$.*

Proof:

Let G be a $(q, q-3)$ graph for $q \geq 7$ and let $T(G)$ be the modular decomposition of G. By proposition 3 above in order to show that $cwd(G) \leq q$, it is enough to show that for each internal node h of $T(G)$, $cwd(G(h)) \leq q$, where $G(h)$ is the representative graph of h in $T(G)$. If h is a P-node (S-node) then $G(h)$ is an edgeless graph (a clique), and has a clique width equals to 1 (2). If h is an N-node then by proposition 4 above $G(h)$ is is isomorphic to either a prime spider, a disc, a prime p–tree or a graph with at most q vertices. Since prime spiders have clique–width ≤ 4 (cf. [CMR98a]), discs have clique width ≤ 4 (by Propositions 5 and 6 above), prime p–trees have clique–width ≤ 6 (by Propositions 7, 8, 9 and 10 above), and a graph with

Hence, the 6 expression t_i defined below satisfies the conditions of the claim:

$$t_i = \quad \rho_{6\to 2}(\eta_{6,5}(\eta_{6,1}(\eta_{6,2}(6(z_i)\oplus \rho_{6\to 5}(\rho_{5\to 3}(\eta_{6,5}(\eta_{6,1}(6(v_i) \oplus t_{i-1}))))))))$$

CASE 2: Suppose i is odd.

In this case from observation 1 above and from the inductive hypothesis on t_{i-1} it follows that a 6–expression t_i which defines the graph G_i and satisfies the conditions of the claim can be constructed by the following steps:

(i) Add the vertex v_i and label it with 6.

(ii) Connect all the vertices labeled with 6 to all the vertices labeled with 1. This will connect v_i to the vertices in $S_{even,i-3}$.

(iii) Rename the label 5 with 1. This will change the label of v_{i-1} from 5 to 1.

(iv) Rename the label 6 with 5. This will change the label of v_i from 6 to 5.

(v) Add the vertex z_i and label it with 6.

(vi) Connect all the vertices labeled with 6 to all the vertices labeled with 1 or 2. This will connect z_i to the vertices in $S_{even,i-1} \cup K_{even,i-1}$.

(vii) Rename the label 6 with 4. This will change the label of z_i from 6 to 4.

Hence, the 6 expression t_i defined below satisfies the conditions of the claim:

$$t_i = \quad \rho_{6\to 4}(\eta_{6,1}(\eta_{6,2}(6(z_i)\oplus \rho_{6\to 5}(\rho_{5\to 1}((\eta_{6,1}(6(v_i) \oplus t_{i-1}))))))$$

□

Proposition 9 *Every spiked p–chain Q_k has clique–width $\leq$ 6 and 6–expression defining it can be constructed in time $O(|V|)$.*

Proof:
We shall prove the claim by induction on i. The claim trivially holds for the case when $i = 3$. Suppose the claim holds for $j \leq i-1$. Then there is a 6 expression t_{i-1} which satisfies the conditions of the claim. We shall show how to construct the 6–expression t_i which defines the labeled graph G_i such that the conditions of the claim are satisfied. First note that we use t_i to build the graph G_i in which we assume that all the vertices $\{z_2, \ldots, z_i\}$ exists. If any of these vertices does not exist then the 6–expression obtained from t_i by omitting all the vertices which does not exist in G_i from the expression is the required 6–expression which defines G_i and satisfies the conditions of the claim.
We consider the following two cases:
CASE 1: Suppose i is even.
In this case from observation 1 above and from the inductive hypothesis on t_{i-1} it follows that a 6–expression t_i which defines graph G_i and satisfies the conditions of the claim can be constructed by the following steps:

(i) Add the vertex v_i and label it with 6.

(ii) Connect all the vertices labeled with 6 to all the vertices labeled with 1 or 5. This will connect v_i to the vertices in $S_{even,i-2} \cup \{v_{i-1}\}$.

(iii) Rename the label 5 with 3. This will change the label of v_{i-1} from 5 to 3.

(iv) Rename the label 6 with 5. This will change the label of v_i from 6 to 5.

(v) Add the vertex z_i and label it with 6.

(vi) Connect all the vertices labeled with 6 to all the vertices labeled with 1 or 2 or 5. This will connect z_i to the vertices in $S_{even,i} \cup K_{even,i-2}$.

(vii) Rename the label 6 with 2. This will change the label of z_i from 6 to 2.

denote by K_{even} (resp. K_{odd}) the set of even (resp. odd) vertices of K. We denote by S_i (resp. K_i) the set of vertices $\{v_1, \ldots, v_i\}$ (resp. $\{z_1, \ldots, z_i\}$). We denote by $S_{even,i}$ (resp. $S_{odd,i}$) the set of even (resp. odd) vertices in S_i. Likewise we denote by $K_{even,i}$ (resp. $K_{odd,i}$) the set of even (resp. odd) vertices in K_i. We denote by G_i the subgraph of G induced by $S_i \cup K_i$. We assume also for simplicity that for $k-4 \leq i \leq k$ the set K_i is equal to the set K_{k-5}. The following observation follows from the above definitions.

Observation 1 *Let G be a spiked p–chain Q_k, then G_i can be obtained from G_{i-1} by adding the two vertices v_i and z_i and adding the following edges:*

- *If i is even then connect v_i to all the vertices in $S_{even,i-2} \cup \{v_{i-1}\}$.*
- *If i is odd then connect v_i to all the vertices in $S_{even,i-3}$.*
- *If i is even then connect z_i to all the vertices in $S_{even,i} \cup K_{even,i-2}$*
- *If i is odd then connect z_i to all the vertices in $S_{even,i-1} \cup K_{even,i-1}$*

Claim 3 *Let G be a spiked p–chain Q_k, then for $4 \leq i \leq k$ there is a 6–expression t_i which defines the labeled graph G_i such that:*

- *If i is even then*
  - *All the vertices of $S_{even,i-2}$ (resp. $K_{even,i}$) are labeled with 1 (resp. 2).*
  - *All the vertices of $S_{odd,i-1}$ (resp. $K_{odd,i-1}$) are labeled with 3 (resp. 4).*
  - *v_i is labeled with 5.*

  If i is odd then
  - *All the vertices of $S_{even,i-1}$ (resp. $K_{even,i-1}$) are labeled with 1 (resp. 2).*
  - *All the vertices of $S_{odd,i-2}$ (resp. $K_{odd,i}$) are labeled with 3 (resp. 4).*
  - *v_i is labeled with 5.*

can be constructed in $O(|V|)$ time. □

Proposition 8 *Every complement of a spiked p–chain P_k which is not a complement of a simple path has clique–width exactly 4 and a 4–expression defining it can be constructed in time $O(|V|)$.*

Proof:
Similar to the proof of proposition 7 above. □

3.3 A spiked p–chain Q_k and its complement

Below we recall from [Bab98b] the definitions of the graphs called p–chain Q_k and spiked p–chain Q_k.
Let G be a graph and let $v_1, \ldots, v_k$ be an ordering of the vertices of G. We denote by $N(v_i)^+$ (resp. $\overline{N}(v_i)^+$) the set of all neighbors (resp. non-neighbors) of v_i with index larger than i. Then G is called a *p–chain* Q_k if G has k vertices $v_1, \ldots, v_k$ and the edges of G are defined as follows:

- $N(v_i)^+ = \{v_{i+1}\}$ for i odd.
- $\overline{N}(v_i)^+ = \{v_{i+1}\}$ for i even.

A graph G is called a *spiked p–chain* Q_k if G is a p–chain $Q_k = (v_1, \ldots, v_k)$, $k \geq 6$, with additional vertices $z_2, z_3, \ldots, z_{k-5}$ such that

- $N(z_i) = \{v_2, v_4, \ldots v_{i-1}, v_{i+1}\} \cup \{z_2, z_4 \ldots, z_{i-1}\}$ for i odd, and
- $\overline{N}(z_i) = \{v_1, v_3, \ldots v_{i-1}, v_{i+1}\} \cup \{z_3, z_5 \ldots, z_{i-1}\}$ for i even.

Any of the vertices $z_2, z_3, \ldots z_{k-5}$ may be missing. We say that G is an even (resp. odd) spiked p–chain Q_k if k is even (resp. odd).
Let G be a spiked p–chain Q_k we denote by S the set of vertices $v_1, \ldots, v_k$ of G and we denote by K the set of vertices $z_2, \ldots, z_{k-5}$ of G. We denote by S_{even} (resp. S_{odd}) the set of even (resp. odd) vertices of S. Likewise we

spiked p–chain P_k (cf. [Bab98b]) if it is a simple path $P_k = (v_1, \ldots, v_k)$ for $k \geq 6$ with the possibility of adding one or two vertices x and y such that: $N(x) = \{v_2, v_3\}$ and $N(y) = \{v_{k-1}, v_k\}$ and x and y do not belong to a common P_4. In this section we show that a spiked p–chain P_k which is not simple path has clique–width exactly 4, and the same holds for its complement.

Proposition 7 *Every spiked p–chain P_k which is not a simple path has clique–width exactly 4 and a 4–expression defining it can be constructed in time $O(|V|)$.*

Proof:
Let G be a spiked p–chain P_k such that G is the simple path $\{v_1, \ldots, v_k\}$ with the addition of the two vertices x and y such that $N(x) = \{v_2, v_3\}$ and $N(y) = \{v_{k-1}, v_k\}$. The case when either x or y is missing can be handled similarly. We first note that the clique–width of the subgraph of G induced by the 6 vertices: $v_1, \ldots, v_5$ and x is of clique width greater than 3. This can be proved by considering all the possible ways to define this graph using 3–expression and showing that this is not possible. Hence, the clique width of G is ≥ 4. We show below that G can be defined by a 4–expression which implies that the clique–width of G is exactly 4.
By claim 1 above there is a 4–expression t_{k-6} which defines the labeled simple path P_{k-6} such that the two endpoints of the path are labeled with 1 and 2 and all the other vertices of the path are labeled with 3. Clearly there are 4 expressions t_{left} and t_{right} which defines the labeled subgraphs of G induced by the vertices $\{x, v_1, v_2, v_3\}$ and $\{y, v_{k-2}, v_{k-1}, v_k\}$ respectively such that the vertices v_3 and v_{k-2} are labeled with 4 and all the other vertices are labeled with 3. It follows that the following 4–expression, denoted as e defines the graph G:

$$e = \eta_{4,2}(t_{right} \oplus \rho_{4\to3}(\eta_{4,1}(t_{left} \oplus t_{k-6})))$$

Since t_{k-6} can be constructed in $k-6$ steps and t_{right} and t_{left} can be constructed in constant time, it follows that the 4–expression e which defines G

at least 3 in the graph defined by s_k. Also no vertex of S is adjacent to any vertex in $tree(c, s_k)$, since all the vertices of $tree(c, s_k)$ other than v and x have label 3 at a. It follows that the graph defined by s_k is disconnected, since there is no vertex in this graph which is adjacent to any vertex in S. This contradicts the assumption that the graph defined by s_k is a C_k.

CASE 2.2: Suppose that u is adjacent to just one vertex in $tree(c, s_k)$ and v is adjacent to just one vertex in $tree(b, s_k)$.

From this assumption it follows that all the vertices of the graph defined by s_k other than u and v must have the same label at a which is different from 1 or 2. Let denote this label by 3. Let S_1 and S_2 denote the sets of vertices occuring at $tree(b, s_k)$ and $tree(c, s_k)$ respectively. Since all the vertices of S_1 and S_2 other than u and v have label 3 at v, it follows that there is just one edge e crossing between S_1 and S_2 in the graph defined by s_k. Thus, the graph defined by s_k is not 2–connected, since removing the edge e will disconnect the graph. Since a cycle C_k is 2–connected, this contradicts the assumption that the graph defined by s_k is a C_k.

Since we have considered all possible cases we conclude that there is no 3–expression which defines a simple cycle C_k having at least 7 vertices. Since we have shown above that every such cycle can be defined by a 4–expression, it follows that the clique–width of every such cycle is exactly 4. □

Recall that we denote by $\overline{C_n}$ the complement of the simple cycle C_n.

Proposition 6 *Every complement of a simple cycle $\overline{C_n}$ having at list 7 vertices, has a clique–width exactly 4 and a 4–expression defining it can be constructed in time $(O(|V|))$.*

Proof:

Similar to the proof of proposition 5 above. □

3.2 A spiked p–chain P_k and its complement

Recall that for $n \in N$ we denote by P_k the simple path of length $k-1$. Recall also that $N(x)$ denote the set of all neighbors of x. A graph G is called the same label (say 3), and the two endpoints of the path either have the same label 2 or have the two labels 1 and 2. Since $k-1 \geq 6$, this contradicts either claim 2 above or corollary 1 above.

CASE 2: Suppose that u is not the only vertex of the cycle occuring in $tree(b, s_k)$ and v is not the only vertex of the cycle occuring in $tree(c, s_k)$.

CASE 2.1: Suppose that either u is adjacent to two vertices in $tree(c, s_k)$ or v is adjacent to two vertices in $tree(b, s_k)$. We assume without loss of generality that u is adjacent to two vertices in $tree(c, s_k)$. Clearly, one of these two vertices is v and let x be the other vertex occuring in $tree(c, s_k)$ which is adjacent to u.

CASE 2.1.1: Suppose that x and v have different labels at a. Let 3 denote the label of x at a. By the above assumption there is another vertex y occuring in $tree(b, s_k)$.

If y is labeled with 1 at a (i.e the same label as u) then the 4 vertices u,v,x and y induce a C_4 in the graph defined by s_k, a contradiction since this graph is a C_k, for $k \geq 7$.

If y is labeled with 2 or 3 at a, then u have degree at least 3 in the graph defined by s_k, a contradiction since this graph is a C_k.

Hence the label of y at a must be different from 1, 2 or 3, a contradiction to the assumption that s_k is a 3–expression.

CASE 2.1.2: Suppose that x and v have the same label at a. Recall that we denote this label by 2. By the above assumption there is another vertex y occuring in $tree(b, s_k)$.

If y is labeled with 1 at a (i.e the same label as u) then the 4 vertices u,v,x and y induce a C_4 in the graph defined by s_k, a contradiction since this graph is a C_k, for $k \geq 7$.

If y is labeled with 2 at a, then u have degree at least 3 in the graph defined by s_k, a contradiction since this graph is a C_k.

Hence, y and all the vertices of the graph other than u,v or x must have label 3 at a. Let S denote the set of all the vertices occuring in $tree(b, s_k)$ excluding u. Clearly, no vertex of S is adjacent to u, or else u will have degree

Proposition 5 *Every simple cycle C_n having at list 7 vertices, has a clique–width exactly 4 and a 4–expression defining it can be constructed in time $O(|V|)$.*

Proof:
We first show that for $n \geq 3$, the clique–width of every simple cycle C_n is ≤ 4. Let $n \in N$, such that $n \geq 3$. By claim 1 above there is a 4–expression t_{n-1} which defines the labeled path P_{n-1}, such that the two endpoints of P_{n-1} are labeled with 1 and 2 and all the internal vertices of the path are labeled with 3. Then the following 4–expression, denoted by s_n defines C_n:

$$s_n = \eta_{4,2}(\eta_{4,1}(4(x) \oplus t_{n-1}))$$

From the construction of t_{n-1} in claim 1 above it follows that the 4–expression s_n defining C_n can be obtained in time $O(|V|)$.
We shall show below that every simple cycle of at least 7 vertices can not be defined by any 3–expression. Suppose that there is a 3–expression s_k which defines a cycle C_k having at least 7 vertices. Let $tree(s_k)$ be the tree corresponding to s_k, (see definition 7 above), let a be the highest node in $tree(s_k)$ which corresponds to a $\oplus$ operation, and let b and c be the two sons of a. Since a is the highest $\oplus$ operation in $tree(s_k)$, it follows that all the vertices of the cycle C_k occur in the leaves of $tree(a, s_k)$ (see definition 8 above).

Since C_k is connected it follows that there are two adjacent vertices u and v occuring in $tree(b, s_k)$ and $tree(c, s_k)$ respectively. Clearly u and v must have different label at a. Let 1 and 2 denote the labels of u and v at a respectively. We consider the following cases:
CASE 1: Suppose that either u is the only vertex of the cycle occuring in $tree(b, s_k)$ or v is the only vertex of the cycle occuring in $tree(c, s_k)$. We assume without loss of generality that u is the only vertex of the cycle occuring in $tree(b, s_k)$.
In this case the graph defined by the 3–expression corresponding to $tree(c, s_k)$ is a labeled path P_{k-1}, such that all the internal vertices of the path have

CASE 2.1.3: Suppose that u is not adjacent to any vertex occuring in $tree(c, t_k)$, and v is not adjacent to any vertex occuring in $tree(b, t_k)$. Since the graph P_k is connected there are two vertices x,z occuring in $tree(b, t_k)$. and $tree(c, t_k)$, respectively such that x is adjacent to z. x and z must have different labels at a, or else they can not be made adjacent. Moreover, x and z can not be labeled with 1 or 2 at a, or else the graph defined by t_k will include an internal node which has the same label as the two endpoints of P_k. It follows that u,v,x and z have 4 different labels at a, a contradiction to the assumption that t_k is a 3–expression.

CASE 2.2: suppose that u and v have the same label at a. This case can be handled similarly to case 1.2 above.

Since we have considered all possible cases we conclude that there is no 3–expression t_k which defines the labeled path P_k such that its two endpoints have the same label and all the internal vertices has another label, a contradiction. □

Corollary 1 *for every $n \in N$, such that $n \geq 6$ there is no 3–expression t_n, which defines the labeled path P_n such that the two endpoints of the path u and v are labeled with 1 and 2 respectively, and all the other vertices of the path are labeled with 3.*

Proof:

Suppose there is a 3–expression r_n which defines the labeled path P_n such that the two endpoints of the path u and v are labeled with 1 and 2 respectively, and all the other vertices of the path are labeled with 3. Let t_n be the 3–expression defined by:

$$t_n = \rho_{3\to 2}(\rho_{2\to 1}(r_n))$$

It is easy to see that the 3–expression t_n defines the labeled path P_n such that the two endpoints of the path are labeled with 1 and all the other vertices of the path are labeled with 2, a contradiction to claim 2 above. □

has the same label as the two endpoints of P_k. It follows that u,v,x and z have 4 different labels at a, a contradiction to the assumption that t_k is a 3–expression.

CASE 1.2: suppose that u and v have the same label at a. Let 1 denote the label of u and v at a. Since there is no vertex in P_k which is adjacent to its two endpoints u and v, it follows that there is no vertex x in $tree(c, t_k)$ which is adjacent to u or to v. Since the graph P_k defined by the 3–expression t_k is connected there are two vertices x,z occuring in $tree(b, t_k)$. and $tree(c, t_k)$, respectively such that x is adjacent to z. As in case 1.1.2 it can be shown that the 3 vertices u,x and z must have different labels at a. Let 2 and 3 denote the labels of x and z at a respectively.

Let W denote the set of all vertices of P_k other than u,v x and z. Clearly, no vertex in W can be labeled with 1 at a. Suppose that all the vertices in W are labeled just with 2 and 3 at a. Since there are at least two vertices in W, it follows that either there is one vertex in P_k which is adjacent to more than two vertices or that there is at least an induced C_4 of P_k, which is not possible. Hence, at least one vertex in W must have a label different than 1 2 or 3 at a, in contradiction to the assumption that t_k is a 3–expression.

CASE 2: Suppose that u and v the two endpoints of the path P_k, occur in $tree(b, t_k)$ and $tree(c, t_k)$ respectively.

CASE 2.1: Suppose that u and v have different labels at a. Let 1, and 2 be the two labels of u and v at a, respectively.

CASE 2.1.1: Suppose that u is adjacent to some vertex x occuring in $tree(c, t_k)$. Then x must have label different than 1 and 2 at a. Let 3 denote the label of x at a. Since u is an endpoint vertex, x is the only vertex which is adjacent to u. Thus, all the other vertices of $tree(a, t_k)$ can not have label 3 at a. It follows that there is at least one vertex y which must have a label other than 1 2 or 3 at a, a contradiction to the assumption that t_k is a 3–expression.

CASE 2.1.2: Suppose that v is adjacent to some vertex x occuring in $tree(b, t_k)$. This case is similar to case 2.1.1 above.

Proof:
Assume that the claim does not hold for some $n = k$. Then there is a 3–expression t_k which defines the labeled path P_k such that the two endpoints of the path are labeled with 1, and all the other vertices of the path are labeled with 2. Let $tree(t_k)$ be the tree corresponding to t_k, (see definition 7 above), let a be the highest node in $tree(t_k)$ which corresponds to a $\oplus$ operation, and let b and c be the two sons of a. Since a is the highest $\oplus$ operation in $tree(t_k)$, it follows that all the vertices of the path P_k occur in the leaves of $tree(a, t_k)$ (see definition 8 above). We consider the following cases:

CASE 1: Suppose that u and v the two endpoints of the path P_k, occur either in $tree(b, t_k)$ or in $tree(c, t_k)$. We assume without loss of generality that both u and v occur in $tree(b, t_k)$.

CASE 1.1: suppose that u and v have different labels at a. Let 1 and 2 be the two labels of u and v at a, respectively.

CASE 1.1.1: Suppose that one of the vertices u and v (say u) is adjacent to some vertex x occuring in $tree(c, t_k)$. Then x must have label different than 1 and 2 at a. Let 3 denote the label of x at a. Since u is an endpoint vertex, x is the only vertex which is adjacent to u. Hence, all the other vertices of $tree(a, t_k)$ can not have label 3 at a. Moreover, all the other vertices of $tree(a, t_k)$ can not have label 1 or 2 at a, or else the labeled graph defined by t_k will have an internal vertex which has the same label as the two endpoints. Since there are at least 6 vertices in P_k, we get that there is another vertex say z, having label at a which is different from 1,2 or 3. But this contradicts the assumption that t_k is a 3–expression.

CASE 1.1.2: Suppose that u and v are not adjacent to any vertex x occuring in $tree(c, t_k)$. Since the graph P_k defined by the 3–expression t_k is connected there are two vertices x,z occuring in $tree(b, t_k)$. and $tree(c, t_k)$, respectively such that x is adjacent to z. x and z must have different labels at a, or else they can not be made adjacent. Moreover, x and z can not be labeled with 1 or 2, or else the graph defined by t_k will include an internal node which

Proposition 4 (Babel [Bab98a]) *Let G be a $(q, q-3)$ graph for $q \geq 7$ and let h be an internal N-node of $T(G)$, then $G(h)$ is isomorphic to either a prime p–tree, a disc, a prime spider, or a graph with at most q vertices.*

3 Clique–width of prime graphs of $(q, q-3)$ graphs

3.1 A simple cycle C_n and its complement $\overline{C_n}$

Recall that for $n \in N$, such that $n \geq 3$, C_n denotes a simple cycle of length n. Clearly, C_3 and C_4 are cographs and have clique–width exactly 2. For C_5 and C_6 there is a 3–expression which defines them, and since they are not cographs they have clique–width exactly 3. In this section we show that for $n \geq 7$, C_n has clique–width exactly 4. We shall use the following two claims:

Claim 1 *For every $n \in N$, such that $n \geq 2$ there is a 4–expression t_n, such that the labeled graph defined by t_n is a path P_n such that the two endpoints of the path are labeled with 1 and 2 and all the other vertices of the path are labeled with 3.*

Proof:
We shall prove the claim by induction on n. The claim trivially holds for $n = 2$. Assume that the claim holds for $n = k-1$, and let t_{k-1} be a 4–expression which defines the labeled path P_{k-1} such that the two endpoints of the path are labeled with 1 and 2 and all the other vertices of the path are labeled with 3. Then the following 4–expression t_k defines the path P_k and satisfies the conditions of the claim:

$$t_k = \rho_{4\to 1}(\rho_{1\to 3}(\eta_{4,1}(4(x) \oplus t_{k-1}))))$$

□

Claim 2 *for every $n \in N$, such that $n \geq 6$ there is no 3–expression t_n, which defines the labeled path P_n such that the two endpoints of the path are labeled with 1 and all the other vertices of the path are labeled with 2.*

2.3 The modular decomposition of $(q, q-3)$ graphs

In this section we recall from [Bab98a] the list of possible prime graphs obtained by the modular decomposition of a $(q, q-3)$ graph. For that we shall need the following definitions.

Recall that the neighborhood $N(v)$ of a vertex v of G is defined as the set of vertices of G adjacent to v, i.e.: $N(v) = \{u | (u, v) \in E\}$.

Definition 16 (Prime spider) *A graph G is a* prime spider *if the vertex set of G can be partitioned into sets S, K and R such that:*

(i) S is a stable set (i.e. no vertex in S is adjacent to the other), K is a clique and $|S| = |K| \geq 2$.

(ii) R contains at most one vertex, i.e. $|R| \leq 1$, and if R contains one vertex say r, then r is adjacent to all the vertices in K and is not adjacent to any of the vertices in S.

(iii) There exist a bijection f between S and K such that either $N(x) = \{f(x)\}$ for all vertices x in S or else $N(x) = K - \{f(x)\}$ for all vertices x in S.

The triple (S, K, R) is called the spider partition *of G.*

A *disc* (cf. [Bab98a]) is a simple cycle C_n or its complement for $n \geq 5$.

The graph R_5 is obtained by adding the edge (v_1, v_3) to the path of length 4 consisting of the vertices $v_1, \ldots, v_5$. The graph R_6 is obtained from R_5 by adding one vertex v_6 and connecting it just to v_2. The graph R_7 is obtained from R_6 by adding one vertex v_7 and connecting it to v_3 and v_4.

The definitions of a *spiked p–chain* P_k and a *spiked p–chain* Q_k are given in sections 3.2 and 3.3 respectively.

We say that a graph G is a *prime p–tree* if it is either a P_4, R_5, $\overline{R_5}$, R_6, $\overline{R_6}$, R_7, $\overline{R_7}$, a spiked p–chain P_k, a complement of a spiked p–chain P_k, a spiked p–chain Q_k or a complement of a spiked p–chain P_k.

The following proposition follows from [Bab98a]:

Definition 14 (The module $M(h)$ and the representative graph $G(h)$) *Let h be an internal node of $T(G)$, we denote by $M(h)$ the module corresponding to h which consists of the set of vertices of G appearing in the leaves of the subtree of $T(G)$ rooted at h. Let $\{h_1, \ldots, h_r\}$ be the set of sons of h in $T(G)$, we denote by $G(h) = \langle V(h), E(h) \rangle$ the* representative graph *of the module $M(h)$ defined by: $V(h) = \{h_1, \ldots, h_r\}$ and*

$$E(h) = \{(h_i, h_j) \mid \exists u, v(u \in M(h_i) \wedge v \in M(h_j) \wedge (u, v) \in E)\}$$

Note that by the definition of a module, if a vertex of $M(h_i)$ is adjacent to a vertex of $M(h_j)$ then every vertex of $M(h_i)$ will be adjacent to every vertex of $M(h_j)$. From the construction of $T(G)$ it follows that:

Proposition 1 *Let G be any graph and let h be an internal node of $T(G)$:*

(i) if h is an S–node then $G(h)$ is a complete graph.

(ii) if h is a P–node then $G(h)$ is edge-less.

(iii) if h is an N–node then $G(h)$ is a prime graph.

Definition 15 ($G[H/v]$) *Let G and H be two disjoint graphs and let v be a vertex of G. We denote by $G[H/v]$ the graph K obtained by the substitution in G of H for v. Formally, $V(K) = V(G) \cup V(H) - \{v\}$, and*

$$\begin{aligned} E(K) = \quad & E(H) \cup \{e : e \in E(G) \text{ and } e \text{ is not incident with } v\} \cup \\ & \{(u, w) : u \in V(H),\ w \in V(G) \text{ and } w \text{ is adjacent to } v \text{ in } G\} \end{aligned}$$

Proposition 2 (Courcelle and Makowsky and Rotics [CMR98a]) *For every disjoint graphs G,H, and for every vertex v of G, $cwd(G[H/v]) = Max\{cwd(G), cwd(H)\}$.*

Recall that for any graph G, we denote by $T(G)$ the modular decomposition of G (which is a tree), and for each internal node h of $T(G)$ we denote by $G(h)$ the representative graph of h defined in definition 14 above.

Proposition 3 (Courcelle and Makowsky and Rotics [CMR98a]) *For every graph G, $cwd(G) = Max\{cwd(H) :\ H$ is a representative graph of an internal node h in the modular decomposition of $G\}$.*

2.2 Clique–width and the modular decomposition of graphs

In this section we recall the connection established in [CMR98a] between the well known concept of the modular decomposition of graphs and the clique–width property of graphs.

The modular decomposition of a graph G, is tree denoted as $T(G)$, together with a set of prime graphs associated with the internal nodes of the tree labeled by N. We start by presenting the basic definitions and properties of the modular decomposition of graphs. It is well known (for example see [CH94]) that for each graph G, the modular decomposition of the graph $T(G)$ is unique up to isomorphism, and can be obtained in linear ($O(|E|)$ time. In our presentation of the modular decomposition below we shall mainly be concerned with its properties, rather than in the way in which it can be constructed. More details on the exact algorithms which can be used for constructing the modular decomposition of graphs can be found in [GV97, BM83, CH94].

Definition 12 (Module, strong module, prime graph) *A subset M of vertices of a graph G is called a* module *of G if every vertex outside M is either adjacent to all vertices in M or to none of them. A module M is called* strong*, if for any module M_1 either $M \cap M_1 = \emptyset$, or one module contains the other. For every graph $G = \langle V, E \rangle$, the* trivial *modules of G are the set V of all the vertices of G, and all the sets of single vertices of G of the form $\{v\}$, where v is any vertex of G. A graph G is called* prime *if it does not have any non-trivial module.*

Definition 13 ($T(G)$ – the modular decomposition of G) *The* modular decomposition *of a graph G, is a tree denoted as $T(G)$. The leaves of $T(G)$ are the vertices of G, and the set of leaves associated with the subtree rooted at an internal node, induce a strong module of G. An internal node is labeled by either P, S or N standing for Parallel, Series and Neighborhood, respectively, and it can be shown that for every graph G the tree $T(G)$ is unique up to isomorphism. More details on how the tree $T(G)$ is constructed can be found in [GV97, BM83, CH94].*

A polynomial time algorithm for recognizing the class $\mathcal{C}(3)$ is presented in [Rot98].

In the following sections when considering a k–expression t which defines a graph G, it will often be useful to consider the tree structure, denoted as $tree(t)$, corresponding to the k–expression t. For that we shall need the following definitions.

Definition 7 ($tree(t)$) *Let t be any k–expression, and let G be the graph denoted by t. We denote by $tree(t)$ the parse tree constructed from t in the usual way. The leaves of this tree are the vertices of G, and the internal nodes corresponds to the operations of t, and can be either binary corresponding to $\oplus$ or unary corresponding to η or ρ.*

***Definition 8 (**$tree(a,t)$**, sub–expression**(a,t)**)** Let t be any k–expression, a be any node in t, we denote by $tree(a,t)$ the subtree of $tree(t)$ rooted at a. We denote by sub–expression(a,t) the k–expression corresponding to $tree(a,t)$.*

Definition 9 (t_1 is a sub–expression of t_2) *Let t_1 be a k–expression and let t_2 be an l–expression, $k \leq l$. We say that t_1 is a sub–expression of t_2 if there exists a node a such that $tree(t_1)$ is the sub–tree of $tree(t_2)$ rooted at a. In other words $tree(t_1)$ is equal to $tree(a, t_2)$.*

Definition 10 (num–vertices(t)) *Let t be any k–expression, we denote by* num–vertices(t) *the number of vertices of the graph defined by t. In other words num–vertices(t) is the number of leaves in $tree(t)$.*

Definition 11 (The label of a vertex v at an internal node a) *Let t be any k–expression, and let G be the graph defined by t. Let a be any internal node of $tree(t)$ and let v be any vertex of G occuring in $tree(a,t)$, i.e. v is a leaf of $tree(a,t)$. The labels of v may change by the ρ operations in t. However, whenever an operation is applied on a sub–expression t_1 of t which contains v, the label of v (like the labels of all the other vertices occuring in t_1) is well defined. The* label of v at a *is defined as the label that v has when the operation a is applied on the subtree of $tree(t)$ rooted at a.*

Definition 4 ($\rho_{i\to j}(G)$) *For a* k*-graph* G *as above we denote by* $\rho_{i\to j}(G)$ *the renaming of* i *into* j *in* G *such that:*

$$\rho_{i\to j}(G) = \langle V, E, V'_1, \dots V'_k\rangle, \text{ where}$$

$V'_i = \emptyset$*,* $V'_j = V_j \cup V_i$*, and* $V'_p = V_p$ *for* $p \neq i, j$*.*

These graph operations have been introduced in [CER93] for characterizing graph grammars. For every vertex v of a graph G and $i \in \{1, \dots k\}$, we denote by $i(v)$ the k-graph consisting of one vertex v labeled by i.

Example 1 *A clique with four vertices* u, v, w, x *can be expressed as:*

$$\rho_{2\to 1}(\eta_{1,2}(2(u) \oplus \rho_{2\to 1}(\eta_{1,2}(2(v) \oplus \rho_{2\to 1}(\eta_{1,2}(1(w) \oplus 2(x)))))))$$

Definition 5 (**k–expression**) *With every graph* G *one can associate an algebraic expression built using the 3 type of operations mentioned above which defines* G*. We call such an expression a* k*–expression defining* G*, if all the labels in the expression are in* $\{1, \dots, k\}$*. Clearly, for every graph* G*, there is an* n*–expression which defines* G*, where* n *is the number of vertices of* G*.*

Definition 6 (**The clique–width of a graph G, $cwd(G)$**) *Let* $\mathcal{C}(k)$ *be the class of graphs which can be defined by* k*–expressions. The clique-width of a graph* G*, denoted* $cwd(G)$*, is defined by:* $cwd(G) = Min\{k : G \in \mathcal{C}(k)\}$*.*

The clique–width is a complexity measure on graphs somewhat similar to tree width, which yields efficient graph algorithms provides the graph is given with its k–expression (for fixed k). A related notion has been introduced by Wanke [Wan94] in connection with graph grammars. $\mathcal{C}(1)$ is the class of edge-less graphs.

Cographs are exactly the graphs of clique width at most 2, and trees have clique width at most 3 (cf. [CO98]).

Problem 2 *Find characterization of graphs of clique width at most* $k, k \geq 3$*. Does there exist a polynomial time algorithms for recognizing the classes* $\mathcal{C}(k),\ k \geq 4$ *?*

theory of square grids. As a by product of Theorems 2 and 4 above we obtain another proof of these results, but with an explicit lower bound on the clique–width. Our proof is direct and does not relay on the notions of graph grammars and the undecidability of the MSOL theory of square grids.

2 Background

2.1 Graph operations and clique-width

In this section we define the notions of graph operations and clique-width, as presented in [CO98].

Definition 1 (**k–graph**) *A* k-graph *is a labeled graph with (vertex) labels in $\{1, 2, \ldots, k\}$. A k-graph G, is represented as a structure $\langle V, E, V_1, \ldots V_k \rangle$, where V and E are the sets of vertices and edges respectively, and $V_1, \ldots, V_k$ form a partition of V, such that V_i is the set of vertices labeled i in G. Note that some V_i's may be empty. A non-labeled graph $G = \langle V, E \rangle$, will be considered as a 1-graph such that all the vertices of G are labeled by 1.*

Definition 2 (**$G \oplus H$**) *For k-graphs G, H such that $G = \langle V, E, V_1, \ldots, V_k \rangle$ and $H = \langle V', E', V'_1, \ldots, V'_k \rangle$ and $V \cap V' = \emptyset$ (if this is not the case then replace H with a disjoint copy of H), we denote by $G \oplus H$, the disjoint union of G and H such that:*

$$G \oplus H = \langle V \cup V', E \cup E', V_1 \cup V'_1, \ldots, V_k \cup V'_k \rangle$$

Note that $G \oplus G \neq G$.

Definition 3 (**$\eta_{i,j}(G)$**) *For a k-graph G as above we denote by $\eta_{i,j}(G)$, where $i \neq j$, the k-graph obtained by connecting all the vertices labeled i to all the vertices labeled j in G. Formally:*

$$\eta_{i,j}(G) = \langle V, E', V_1, \ldots V_k \rangle \text{ , where}$$

$$E' = E \cup \{(u, v) : u \in V_i,\ v \in V_j\}$$

above can not be proved on these graph classes. In particular we show that:

Theorem 2 *There is a class $\mathcal{C}$ containing infinitely many $(6,3)$ graphs, such that for every graph $G = \langle V, E\rangle \in \mathcal{C}$, $cwd(G) \geq \sqrt{|V|/27}$.*

Theorem 3 *For every $q \geq 4$, there is a class $\mathcal{C}(q)$ containing infinitely many (q,q) graphs, such that for every graph $G = \langle V, E\rangle \in \mathcal{C}$, $cwd(G) \geq \sqrt{|V|/27q^3}$.*

Clearly, a (q,t) graph is also a (q',t') graph for $q \geq q'$ and for $t' \leq t$. Hence, by Theorems 1 - 3 we have settled the clique–width question on the (q,t) graph classes for all the possible combinations of q and t, except for the following which is still open:

Problem 1 *Are the classes of $(q, q-1)$ graphs and $(q, q-2)$ graphs for $q \geq 7$ of bounded clique–width?*

For proving Theorems 2 and 3 above we define (cf. definition 18 below) the *2–color–width* property of a graph. We shall show (cf. Theorem 5 below) that for every graph G, if G has 2–color–width $\geq k$ then G has clique–width $\geq k$. However, the other direction (cf. Fact 1 below) does not hold: there is a graph G which has clique–width $\geq k$ but has 2–color–width 1. We believe that this new concept of unbounded 2–colored–width is significant, since it characterizes a big subclass of the class of graphs of unbounded clique–width without using the notions of graph operations, k–expressions and clique width.

Using the same technique for other graph classes, we show that the class of split graphs is not of bounded clique–width. In particular, we show that:

Theorem 4 *There is a class $\mathcal{C}$ containing infinitely many split graphs, such that for every graph $G = \langle V, E\rangle \in \mathcal{C}$, $cwd(G) \geq (\sqrt{2|V|} - 1)/72$.*

Courcelle showed in [Cou93] that the classes of square grids and chordal graphs are not of bounded clique–width, using the notion of graph grammars and based on the the undecidability of the Monadic Second Order Logic

1 Introduction

The study of graph classes having few P_4s have been very active in recent years. Example for such graph classes are the classes of cographs, (extended) P_4–sparse graphs, (extended) P_4–reducible graphs and P_4–tidy, studied in [CLS81, JO89, JO92b, JO92a, JO95a, JO95b, GRT97, GV97]. Babel and Olariu introduced in [BO95] the class of (q,t) graphs which for $t = q - 3$ extends all the graph classes mentioned above. In such a graph no set with at most q vertices is allowed to induced more than t distinct P_4s. Clearly, we assume that $q \geq 4$. In a series of papers (cf. [BO95, BO98a, BO98b, Bab98a, Bab98b]) Babel and Olariu studied the classes of $(q, q-4)$ and $(q, q-3)$ graphs.

The notion of clique-width of graphs was first introduced by Courcelle, Engelfriet and Rozenberg in [CER93], as graphs which can be defined by k-expressions based graph operation which use k vertex labels. A detailed study of clique–width is [CO98]. Clique–width has analogous properties as treewidth: If the clique–width of a class of graphs $\mathcal{C}$ is bounded by k (and the k–expression can be computed from its corresponding graph in time $T(|V|)$) then every decision, optimization, enumeration or evaluation problem on $\mathcal{C}$ which can be defined by a Monadic Second Order formula ψ can be solved in time $c_k \cdot O(|V|) + T(|V|)$ where c_k is a constant which depends only on ψ and k and v is the number of vertices of the input. For details, cf. [CMR98a, CMR98b, CMR99].

In this paper we study the clique–width of the (q,t) graphs for almost all combinations of q and t. We first show that:

Theorem 1 *For every $(q, q-3)$ graph G such that $q \geq 7$, G has clique–width $\leq q$, and a q–expression defining it can be constructed in time $O(|V| + |E|)$.*

The proof of Theorem 1 above is based on the the results of Babel (cf. [Bab98a]) which studied the prime graphs of the class of $(q, q-3)$ graphs.

We continue by showing that the class of $(6,3)$ graphs and the classes of (q,q) graphs for $q \geq 4$ are not of bounded clique–width. Hence Theorem 1

Contents

On the Clique–Width of Graphs with Few P_4's

J.A. Makowsky*and U. Rotics

October 13, 1998

Department of Computer Science
Technion–Israel Institute of Technology
32000 Haifa, Israel
{janos,rotics}@@cs.technion.ac.il

Abstract

Babel and Olariu (1995) introduced the class of (q,t) graphs in which every set of q vertices has at most t distinct induced P_4s.
Graphs of clique-width at most k were introduced by Courcelle, Engelfriet and Rozenberg (1993) as graphs which can be defined by k-expressions based on graph operations which use k vertex labels.
In this paper we study the clique–width of the (q,t) graphs, for almost all possible combinations of q and t.
On one hand we show that every $(q, q-3)$ graph for $q \geq 7$, has clique–width $\leq q$ and a q–expression defining it can be obtained in linear time.
On the other hand we show that this result does not hold for the class of (q,q) graphs for any q, and for the class of $(q, q-3)$ graphs for $q \leq 6$. More precisely, we show that for every q, for every $n \in N$ there is a graph H_n which is a (q,q) graph having n vertices and the clique–width of H_n is at least $(\sqrt{n/3q})/3q$.

*Partially supported by a Grant of the Israeli Ministry of Science for French-Israeli Cooperation (1994), a Grant of the German-Israeli Binational Foundation (1995-1996), and by the Fund for Promotion of Research of the Technion–Israeli Institute of Technology