# The Monadic Second-Order Logic of Graphs XV: On a Conjecture by D. Seese

Bruno Courcelle LaBRI, Bordeaux 1 University, 351 cours de la Libération F-33405 Talence courcell@labri.fr

April 6, 2004

#### Abstract

A conjecture by D. Seese states that if a set of graphs has a decidable monadic second-order theory, then it is the image of a set of trees under a transformation defined by monadic second-order formulas. We prove that the general case of this conjecture is equivalent to the particular cases of directed graphs, partial orders and comparability graphs. We present some tools to prove the conjecture for classes of graphs with few cliques or few complete bipartite subgraphs, for line graphs and for interval graphs. We make an essential use of prime graphs, of comparability graphs and of characterizations of graph classes by forbidden induced subgraphs. Our treatment of infinite graphs uses a representation of countable linear orders by binary trees that can be constructed by monadic second-order formulas. By using a counting argument, we show the intrinsic limits of the methods used so far to handle this conjecture.

## 1 Introduction

A conjecture by D. Seese [See] states that if a set of graphs has a decidable monadic theory, then it is the image of a set of trees under a transformation defined by monadic second-order formulas. We call *monadic second-order trans- ductions* such transformations. We will say that it is *tree-definable*, without specifying the logical language which will be monadic second-order logic in this article.

By results of [CE] and [EV], this is equivalent to saying that this set has bounded clique-width. This means that the graphs can be constructed from isolated vertices taken as basic graphs by means of certain graph operations :

#### 5.1 Bounded clique degree and other constraints.

We will use such codings for classes C of simple undirected graphs without loops and isolated vertices. We will call them simply *graphs* in this section. The adaptation to other classes, in particular of directed graphs, will be straightforward.

Let  $\mathcal{B}$  be a class of connected graphs. Let G be a graph. A  $\mathcal{B}$ -subgraph of G is a subgraph isomorphic to some graph in  $\mathcal{B}$ . A maximal  $\mathcal{B}$ -subgraph is one that is maximal for subgraph inclusion. We denote by  $\mathcal{B}(G)$  the set of maximal  $\mathcal{B}$ -subgraphs of G. The  $\mathcal{B}$ -degree of a vertex is the number of maximal  $\mathcal{B}$ -subgraphs to which it belongs. The  $\mathcal{B}$ -degree of a graph is the maximum  $\mathcal{B}$ -degree of its vertices.

If  $\mathcal{B}$  is the set of cliques, we obtain the notion of *clique-degree*. We will also consider the case where  $\mathcal{B}$  is the set  $\mathcal{KB}$  of cliques and complete bipartite graphs. We obtain thus the notion of  $\mathcal{KB}$ -degree. To take a few examples a string graph  $P_n$  has  $\mathcal{KB}$ -degree one for n = 2 or 3, and  $\mathcal{KB}$ -degree 2 for n =4, and  $\mathcal{KB}$ -degree 3 for n > 4. A tree has  $\mathcal{KB}$ -degree at most 3. A complete bipartite graph  $K_{n,p}$  for n, p > 1 has clique-degree  $Max\{n, p\}$ .

A domino is a finite graph of clique degree at most 2. A graph is a domino iff it does not contain as induced subgraph a *claw*  $(K_{1,3})$ , a *gem* (i.e.,  $\mathbf{1} \otimes P_4$ ) or a 4-wheel (i.e.,  $\mathbf{1} \otimes C_4$ ). See Kloks et al. [KKM]. The proof is done for finite graphs, but it is easy to see that it applies also to countable graphs. We will see that the line graphs of triangle-free graphs are dominoes.

A complete subgraph of a graph G can be specified by the set of its vertices. A complete bipartite subgraph can be specified by two disjoint sets of vertices. In order to specify a complete bipartite subgraph, it is not enough to specify its set of vertices. Consider for instance the graph  $K_4 - e$  ( $K_4$  minus one edge) with vertices 1, 2, 3, 4 and missing edge between 1 and 4. It has three complete bipartite subgraphs defined by the pairs ( $\{1, 2, 4\}, \{3\}$ ), ( $\{1, 4\}, \{2, 3\}$ ) and ( $\{1, 3, 4\}, \{2\}$ ).

Let G be covered by (i.e., be the union of) a set of complete bipartite graphs (represented by)  $(X_1, Y_1), ..., (X_p, Y_p)$ , and a set of cliques (represented by)  $Z_1, ..., Z_m$ , where  $X_1, ..., X_p, Y_1, ..., Y_p, Z_1, ..., Z_m$  are sets of vertices. Let us build a directed colored graph  $H = \gamma(G)$  as follows :

 $V_{H} \ = V_{G} \ \cup W \text{ where } W = \{u_{1},...,u_{p},z_{1},...,z_{m}\} \text{ is a set of new vertices},$ 

 $E_H$  consists of edges that link  $u_j$  to each vertex of  $X_j$ , each vertex of  $Y_j$  to  $u_j$ , for each j = 1, ..., p, and  $z_i$  to each vertex of  $Z_i$ , for each i = 1, ..., m.

Furthermore, we color  $u_1, ..., u_p$  by 1, and  $z_1, ..., z_m$  by 2. The other vertices have no color.

We consider  $H = \gamma(G)$  as a coding of G. The graph G can be recovered from H as follows :

One creates an edge between u and v where u and v are uncolored vertices whenever there is a path in H such that :  $u \longrightarrow w \longrightarrow v$  with w colored by 1. One creates an edge between u and v where u and v are uncolored vertices whenever there are edges in H such that :  $u \longleftarrow w \longrightarrow v$  with w colored by 2. One deletes all vertices colored by 1 or 2. It is clear that we obtain G in this way by an MS transduction denoted by  $\delta$ . Some edges of G may belong to several of the covering cliques and bipartite complete subgraphs. Note also that the colors 1 and 2 make it possible to distinguish easily the vertices of G from the auxilliary vertices of the set W.

Let us now assume that G has  $\mathcal{KB}$ -degree at most k. It is covered by  $\mathcal{KB}(G)$ . Let  $H = \gamma(G)$ . Each vertex of G has degree at most k in H. Hence H has an orientation of indegree at most k (not the one we defined above) and is thus uniformly k-sparse (See Courcelle [03a]). We denote by C(x) the set of elements of  $\mathcal{KB}(G)$  that are cliques containing a vertex x, and by B(x) the set of those that are complete bipartite subgraphs containing x.

In order to apply Proposition 5.1 we need only prove that  $\gamma$  is an MS transduction.

We can do that with the additional assumption that the given graph G is given with an  $\omega$ -linear order  $\leq$  of its set of vertices. Assuming this, we can order lexicographically the sets of subgraphs C(x) and B(x).

For each *i* one can build MS formulas  $\gamma_i(x, Z)$  and  $\beta_i(x, X, Y)$  expressing respectively that Z is the *i*-th element in C(x) and that (X, Y) is the *i*-th one in B(x). For each *i* and *j*, one can build an MS formula  $\overline{\gamma}_{i,j}(x, y)$  expressing that the *i*-th element in C(x) is the same as the *j*-th element in C(y), and an MS formula  $\overline{\beta}_{i,j}(x, y)$  expressing that the *i*-th element in B(x) is the same as the *j*-th element in B(y).

One can then build H by taking as elements  $u_h$  of W the equivalence classes of pairs (x, i) for the equivalence relation  $\approx$  such that  $(x, i) \approx (y, j)$ iff the *i*-th element in B(x) is the same as the *j*-th element in B(y), and as elements  $z_h$ , the equivalence classes of pairs (x, i) such that  $(x, i) \approx' (y, j)$  iff the *i*-th element in C(x) is the same as the *j*-th element in C(y). This can be done using the formulas  $\overline{\gamma}_{i,j}$  and  $\overline{\beta}_{i,j}$ 

done using the formulas  $\overline{\gamma}_{i,j}$  and  $\overline{\beta}_{i,j}$ The edges between the vertices of W and those of G can be defined by using the MS formulas  $\gamma_i(x, Z)$  and  $\beta_i(x, X, Y)$  which say which subgraph a pair (x, i) does represent.

Since the equivalence relations  $\approx$  and  $\approx'$  are MS-definable, their quotients can be constructed by an MS transduction (subsection 2.2).

Hence, we have  $H = \gamma'(G, \leq)$  where  $\gamma'$  is an MS transduction taking as input G augmented by an  $\omega$ -linear order  $\leq$ . Hence we have proved the following.

**Proposition 5.2 :** There exists an MS transduction  $\gamma'$  that transforms  $(G, \leq)$  into  $\gamma(G)$ , where G is a graph of  $\mathcal{KB}$ -degree at most k and  $\leq$  is any  $\omega$ -linear order on its set of vertices.

The only purpose of the linear order  $\leq$  is to order lexicographically the sets of vertices. The graph H, although defined in terms of  $\leq$  as  $\gamma'(G, \leq)$ , is independent of the choice of  $\leq$ . We obtain for graphs of clique degree, or of  $\mathcal{KB}$ -degree at most k a weak version of the Conjecture :

**Theorem 5.3:** If a set L of graphs of clique degree at most k, or of  $\mathcal{KB}$ degree at most k has a decidable MS-OI satisfiability problem, then it has
bounded clique-width.

**Proof**: Let L be of  $\mathcal{KB}$ -degree at most k, let L' be the corresponding set of  $\omega$ -linearly ordered graphs. Let  $M = \gamma'(L')$ . Let us prove that this set has a decidable MS satisfiability problem. Every MS formula  $\varphi$  on these graphs translates back via  $\gamma'$  into an MS formula  $\varphi^{\#}$  on  $\omega$ - linearly ordered graphs. This formula is order-invariant since the graphs  $\gamma'(G, \leq)$  do not depend on the particular choice of  $\leq$ . Hence its satisfiability in L can be decided by the hypothesis. So can be the MS satisfiability problem for M. Hence M has bounded clique-width since its members are uniformly k-sparse. So has  $L = \delta(M)$  where  $\delta$  is an MS transduction.

If L has clique degree at most k, the proof is similar. In the construction of  $\gamma$ , we omit everything concerning complete bipartite graphs .

**Corollary 5.4**: If a class of graphs has clique degree at most k, or has  $\mathcal{KB}$ -degree at most k and if an  $\omega$ -linear order is MS-definable on the graphs of this class, then it satisfies the Conjecture.

It was proved in [Cou95b] that the Conjecture holds for finite chordal graphs of bounded clique degree. We obtain here a similar result which is not exactly an extension of this result to graphs which are not chordal because we use the auxiliary  $\omega$ -linear orders so that the hypothesis is stronger.

**Question 5.5 :** Can one construct  $\gamma$  in the proof of Proposition 5.2 without using the auxiliary linear order ?

If this is possible we obtain a proof of the Conjecture for classes of graphs of bounded clique degree and of bounded  $\mathcal{KB}$ -degree. This proof technique can be adapted to other classes of graphs than cliques and complete bipartite graphs. It suffices to have a monadic second-order descriptions of the subgraphs in terms of a fixed number of sets of vertices and a representation with few edges so as to obtain graphs uniformly k-sparse.

### 5.2 Line graphs

We now consider the case of line graphs.

Let G be a simple loop-free undirected graph without isolated vertices. For the purpose of using MS logic with edge set quantifications, we will represent it by the structure  $Inc(G) = \langle V_G \cup E_G, inc_G \rangle$ . Note that  $(e, x, y) \in inc_G$  iff  $(e, y, x) \in inc_G$ .

The line graph of G, denoted by L(G) is the simple undirected graph with set of vertices  $E_G$  and edges  $e \longrightarrow f$  iff e and f are edges of G sharing a vertex. We say that H is a line graph if it is L(G) for some graph G. We let  $\mathcal{LG}$  be the class of line graphs.

Several graphs may have the same line graph. Fo example  $K_3$  is the line graph of both  $K_3$  and  $K_{1,3}$ . A triangle-free graph is a graph without  $K_3$  as a subgraph. We let  $\mathcal{TF}$  be the class of triangle free graphs.

In this article, we have shown that proving the Conjecture for a quite large variety of graph classes is as difficult as for the general case. These classes are listed in Theorem 3.8, and one should also add to the list the class of comparability graphs. We have also established a few new cases of validity of the Conjecture, in particular for line graphs, interval graphs, dimension 2 posets and quasi-series-parallel posets, and given a method potentially applicable for classes of graphs covered by few cliques and few complete bipartite graphs.

Forbidden configurations have been very useful especially for dealing with comparability graphs. This suggests to compare graph classes in terms of the forbidden induced subgraphs characterizing them. What follows is more a research program than a set of achievements.

### 10.1 Forbidden induced subgraphs.

Let  $\mathcal{Z}$  be a set of finite simple graphs, either directed or not. We let  $Forb(\mathcal{Z})$  be the class of finite simple graphs that do not have any induced subgraph isomorphic to any graph in  $\mathcal{Z}$ . If  $\mathcal{Z}$  is finite, then  $Forb(\mathcal{Z})$  is first-order definable. This is the case for cographs (excluding  $P_4$ ), for line graphs (there are 9 excluded induced subgraphs) to take a few examples. If  $\mathcal{Z}$  is MS definable, then so is  $Forb(\mathcal{Z})$ . This is the case for the class of comparability graphs, of convex bipartite graphs, of interval graphs as we have seen, and of perfect graphs (as a consequence of the proof of the Strong Berge Conjecture). We propose the following classification of graph classes characterized by finite or infinite sets of forbidden induced subgraphs, in relation with the Conjecture. To simplify the discussion, we only consider finite graphs.

For every set  $\mathcal{Z}$  of forbidden induced subgraphs, we ask the following questions concerning the class  $Forb(\mathcal{Z})$ :

Has it has bounded clique-width ?

If it has unbounded clique-width, then we ask the following :

Does it satisfy the Conjecture ?

Is it equivalent to the full class Finite Undirected Graphs.?

It may of course happen that we cannot answer any of these questions. We collect a few known answers.

We need some definitions. For positive integers q and t, a (q, t)-graph is a graph in which every set of q vertices has at most t induced subgraphs isomorphic to  $P_4$ . This can be expressed, for each (q, t) in terms of finitely many forbidden induced subgraphs. This notion was introduced by Babel and Olariu [BO]. Let **Bcwd** be set of pairs (q, t) such that (q, t)-graphs have clique-width < k for some k depending only on q and t. Makowsky and Rotics proved ([MakRot]) that

**Bcwd** contains the pairs (q, t) for  $t \le q-3$  and  $q \ge 7$ , does not contain (q, q-1) for  $q \ge 4$ , neither (q, q-3) for q = 4, 5, 6. The other cases are unknown. The structural descriptions of these graph classes should help to settle more cases.

**Proposition 10.1 :** 1) The classes  $Forb(\mathcal{Z})$  have bounded clique-width in the following cases :

 $\mathcal{Z}$  contains  $P_4$ 

 $\mathcal{Z}$  contains the finitely many graphs characterizing the (q, t)- graphs, for each (q, t) in **Bcwd**.

 $\mathcal{Z}$  contains the graphs characterizing the bipartite graphs without a skew star (Lozin [Loz]).

2) The classes  $Forb(\mathcal{Z})$  satisfy the Conjecture if either

 $\mathcal{Z}$  contains the finite set of graphs that characterizes the graphs of degree at most d, for each d,

 $\mathcal{Z}$  contains the finite sets of graphs that characterize line graphs or directed line graphs,

 $\mathcal{Z}$  contains the infinite set characterizing interval graphs.

3) The classes  $Forb(\mathcal{Z})$  are equivalent to the class Finite Undirected Graphs if

 $\mathcal{Z}$  contains no split graph.

#### **Proofs** :

1) The graphs without  $P_4$  are the cographs, which are the graphs of cliquewidth at most 2. The other results are due to [MakRot].

2) Graphs of bounded degree are uniformly d-sparse, the result follows from Proposition 2.5.2. The other cases have been considered in Theorems 5.9 and 8.3.

3) The class of split graphs is closed under taking induced subgraphs. If  $\mathcal{Z}$  contains no split graph, then  $Forb(\mathcal{Z})$  contains all split graphs, hence it is equivalent to the full class Finite Undirected Graphs by the results of section 3.

#### 10.1.1 More open questions

Here are some open questions.

**Question 10.2** : Do we have : Finite Graphs  $\implies$  Countable Graphs ?

**Question 10.3** : Do we have : Finite Graphs  $\implies$  Finite Relational Structures ?

**Question 10.4** : What about the classes of finite partial orders of dimension k for k at least 3 ?

Concerning Question 10.2, it is proved in [Cou04] that a countable graph has clique-width at most f(k) if its finite induced subgraphs have clique-width at most k, where f is a fixed function. It follows that, if the Conjecture holds for finite graphs, then we have the following result :

If a set of graphs has a decidable  $MS^{f}$ -satisfiability problem, then it has bounded clique-width,

where  $MS^f$  is MS logic with an atomic formula Fin(X) expressing that a set X is finite. The proof is easy : using this predicate, we get that if a set of graphs L has a decidable  $MS^f$ -satisfiability problem, so has the set of its finite induced subgraphs. This set has bounded clique-width and so has L by the "compactness" result of [Cou04]. Finiteness can be expressed in terms of an arbitrary  $\omega$ -linear order. It follows that  $MS^f$  is intermediate between MS and MS-OI.

Finally, one may strengthen the definition of tree-definable as follows. Let us say that a set of graphs L is *strongly tree-definable* if there exists an MS coding  $(\gamma, \delta)$  of L into a set of labelled trees T. In such a case, the MS satisfiability problems for L and for the subset of T that is its image under the MS transduction  $\gamma$  are interreducible.

A stronger form of Seese's Conjecture could be :

**Conjecture 10.5** : If a set of graphs has a decidable MS satisfiabiliy problem, then it is strongly tree-definable.

D. Lapoire has proved [Lap] that for each k, one can define a (fairly complicated) MS coding of finite graphs of tree-width at most k into a set of finite labelled trees which encode some tree-decompositions of width k of the input graphs. It is important to note that the trees in T are unordered and of unbounded degree. All existing proofs of particular cases of the Conjecture reduce via MS transductions to that of graphs of bounded tree-width. Hence for all these cases the strong form of the Conjecture holds.

**Acknowledgements :** Thanks to A. Blumensath for his critical reading of a first draft of this article.

# 11 References

## References

[BO] L. Babel, S. Olariu, On the structure of graphs with few P4's. *Discrete Applied Maths*, 84 (1998) 1-13.

- [BVS] A. Brandstädt, Van Bang Le, J. Spinrad, Graph classes, a survey, SIAM monographs on discrete mathematics, 1999.
- [Cou89] B. Courcelle, The monadic second-order logic of graphs II: Infinite graphs of bounded width, Mathematical Systems Theory, 21(1989)187-221.
- [Cou91] B. Courcelle, The monadic second-order logic of graphs V: On closing the gap between definability and recognizability. *Theoretical Computer Science*, 80 (1991) 153-202.
- [Cou92] B. Courcelle. The monadic second-order logic of graphs VII: Graphs as relational structures. *Theoretical Computer Science*, 101 (1992) 3–33.
- [Cou94] B. Courcelle. Monadic second-order graph transductions: A survey. Theoretical Computer Science, 126:53–75, 1994.
- [Cou95] B Courcelle: Structural properties of contexte-free sets of graphs generated by vertex replacement, *Information and Computation* 116(1995) 275-293.
- [Cou95b] B. Courcelle, The monadic second order logic of graphs VIII: Orientations, Annals of Pure Applied Logic 72 (1995) 103-143.
- [Cou96] B. Courcelle, The monadic second-order logic of graphs X: Linear orders, Theoretical Computer Science, 160 (1996) 87-143.
- [Cou97] B. Courcelle. The expression of graph properties and graph transformations in monadic second-order logic. In G. Rozenberg, editor, Handbook of graph grammars and computing by graph transformations, Vol. 1: Foundations, pages 313–400. World Scientific, 1997.
- [Cou02] B. Courcelle, A monadic second-order definition of the structure of convex hypergraphs, Information and Computation 178 (2002) 391-411
- [Cou03a] B. Courcelle, The monadic second-order logic of graphs XIV: Uniformly sparse graphs and edge set quantifications. Theoretical Computer Science 299 (2003) 1-36.
- [Cou04] B. Courcelle, Clique-width of countable graphs: a compactness property. Discrete Mathematics 276 (2004) 127-148.
- [CE] B. Courcelle and J. Engelfriet. A logical characterization of the sets of hypergraphs defined by hyperedge replacement grammars. *Mathematical Systems Theory*, 28:515–552, 1995.
- [CouMak] : B. Courcelle, J. Makowsky, Fusion in relational structures and the verification of monadic second-order properties, Math. Struct. in Comp. Sci. 12 (2002) 203-235.

- [CMR] B. Courcelle, J. Makowsky, U. Rotics, Linear time solvable optimization problems on graphs of bounded clique-width, Theory Comput. Systems 33(2000) 125-150.
- [BC-SO] B. Courcelle and S. Olariu. Upper bounds to the clique-width of graphs. Discrete Applied Maths, 101 (2000) 77-114.
- [DJGT] R. Diestel, T. Jensen, K. Gorbunov, C. Thomassen, Highly connected sets and the excluded grid theorem, J. Comb. Th. Series B, 75 (1999) 61-73.
- [EV] J. Engelfriet and V. van Oostrom. Logical description of context-free graph-languages. Journal of Computer and System Sciences, 55:489– 503, 1997.
- [Gol] M. Golumbic, Algorithmic graph theory and perfect graphs, Academic Press, 1980.
- [GR] M. Golumbic, U. Rotics, On the clique-width of some perfect graph classes, Int. J. Found. Comput. Sci. 11 (2000) 423-443.
- [HJ] M. Habib, R.Jegou, N-free posets as generalizations of series-parallel posets, Discrete Applied Maths 12(1985), 279-291.
- [Kel] D. Kelly, Comparability graphs, in I. Rival ed., Graphs and order, Reidel, Dordrecht, 1985, pp. 3-40.
- [KKM] T.Kloks, D.Krzatsch, H. Müller, Dominoes, Lecture Notes Comput. Sci. Vol. 903, Springer, 1995, pp. 106-120.
- [Lap] D. Lapoire, Recognizability equals monadic second-order definability for sets of graphs of bounded tree-width, Proceedings of STACS 1998, Lec. Notes Comp. Sci., Vol. 1373, pp.618-628.
- [Loz] V. Lozin, Bipartite graphs without a skew star, Discrete Mathematics 257 (2002) 83-100.
- [MakRot] J. Makowsky, U. Rotics, On the Clique-Width of Graphs with Few P4's, Int. J. Found. Comput. Sci. 10(1999) 329-348.
- [MR] R. Mőhring and F. Radermacher, Substitution decomposition for discrete structures and connections with combinatorial optimization, Annals of Discrete Mathematics 19 (1984) 257-356.
- [RS-GM5] N. Robertson, P.Seymour, Graph Minors V, Excluding a planar graph, J. Comb. Th. Series B, 41 (1986) 92-114.

- [See] D. Seese. The structure of the models of decidable monadic theories of graphs. *Annals of Pure Applied Logic*, 53:169–195, 1991.
- [Tro] W. Trotter, Combinatorics and partially ordered sets Dimension theory, J. Hopkins University Press, Baltimore, London, 1992.
- [TM] W. Trotter, J. Moore, Characterization problems for graphs, partially ordered sets, lattices, and families of sets, Discrete Mathematics 16(1976) 361-381.
- [Wil] H. Wilf, Generating Functionology, Academic Press, 1994