

Review of⁵
Graph Structure and Monadic Second-Order Logic
A Language-Theoretic Approach
by **Bruno Courcelle and Joost Engelfriet**
Cambridge University Press, 2012
728 pages, hardcover, \$160.00

Review by
Michaël Cadilhac michael@cadilhac.name
Wilhelm-Schickard-Institut für Informatik, Universität Tübingen

1 Introduction

This book is built around three axes: specification of graphs akin to context-free grammars, efficient model-checking on graphs, and logical description of graphs. The study is motivated by two classes of “context-free” graphs: the HR (Hyperedge Replacement) and VR (Vertex Replacement) equational sets. Those sets are linked to two important complexity measures on graphs which are used to devise fixed-parameter tractable algorithms for monadic second-order (MSO) model-checking. The book focuses on the interplay between algorithmics, algebra, and logic on graphs, and does so in a very detailed, precise, and exhaustive way. Statistically speaking, there is roughly one theorem (or proposition, etc.) and 9 cross-references per page. The book contains no exercises but an extensive list of open problems.

2 Summary

One great didactic value of the book is its first chapter, which constitutes an extended survey (2⁶ pages) of the book’s matter. This chapter helps a lot in making the book pleasant to read. It lays down the structure of the following 600 pages in a very accessible, illustrated, and clear way, hence making both skimming or exhaustive reading of the remainder of the book more natural. This chapter has a similar content to a survey written by Courcelle in 1997.

Chapter 2 presents the two graph algebras that will be the main focus of the book. Both can be specified as a generalization of context-free grammars to graphs, with different generalizations of concatenation on words. These are (1) the *HR graph algebra*, consisting of injectively and partially labeled hypergraphs, built from trivial graphs by identifying nodes with the same labels, and (2) the *VR graph algebra*, consisting of partially labeled simple graphs, built from trivial graphs by adding edges between nodes with given labels. As equational (“context-free”) sets, HR graphs correspond in a precise sense to bounded tree-width graphs, and, as a more artificial yet still natural counterpart, VR graphs correspond to a measure called *bounded clique-width*. It is still open whether this measure has an independent combinatorial characterization. Those two width measures will lead to fixed-parameter tractable algorithms for MSO problems on graphs.

Chapter 3 deals with many-sorted algebras in a more general context. Their equational and recognizable (“regular”) sets are studied, hence laying the ground for the study of such sets for the HR and VR graph algebras. Within this general framework, terms and automata on terms (tree automata) are introduced. The

⁵©2014, Michaël Cadilhac

classical result that context-free languages are preserved under intersection with regular languages carries through this setting.

Chapter 4 studies the {equational, recognizable} sets of graphs of the {HR, VR} algebras. Decidability results are presented, and the links between HR-equationality and tree-width on the one hand, and VR-equationality and clique-width on the other hand are made precise. Some equivalents of Myhill-Nerode theorem for the recognizable sets are shown. Connecting HR and VR, it is shown that HR-equational simple graphs are VR-equational, and that the converse holds for sparse graphs (a similar result holds for recognizable sets, but in the opposite direction).

Chapter 5 introduces MSO logic, and shows a variant of Büchi's theorem on the equivalence of MSO and recognizable. More precisely, MSO graphs are VR-recognizable, and if quantification on the set of edges is allowed, they are HR-recognizable. This in turn implies the decidability of MSO satisfiability for the class of graphs of tree-width and clique-width at most k , for each k . The chapter is placed in a more general framework.

Chapter 6 develops fixed-parameter tractable algorithms, with tree- and clique-width as parameters, for the previously mentioned MSO satisfiability problem. Of particular interest are alternative proofs of previous important theorems, leading to a better algorithmic understanding of the results.

Chapter 7, a very important chapter, focuses on MSO transductions, i.e., transformations specified by MSO formulas. One of the main results in this line is that a set of graphs is VR-equational iff it is the image of the set of trees under an MSO transduction; a similar characterization exists for HR-equational sets, thus emphasizing the strong links between the two formalisms. Some consequences on the decidability of MSO satisfiability and on the logical characterization of recognizability are given.

Chapter 8 considers MSO transductions of terms (trees) and words. In particular, implementations of such transductions with finite-state transducers are presented (relying on two-way transducers for words, and compositions of tree-walking transducers — and other devices — for terms). This leads to automata-theoretic characterizations of the VR-equational sets of *terms* and *words*.

Chapter 9 is about more general relational structures, and strives to generalize the preceding results, in particular algorithmic results, to that setting — some of them being already presented in previous chapters for the algebra of all relational structures. Considering the incidence graphs of relational structures, the results about graphs of bounded tree-width and MSO are generalized. Going further, a new complexity measure generalizing clique-width is introduced, leading to important open questions.

The book is concluded with an opening on the future, including an extensive list of open problems.

3 Opinion

This book is an impressive work, in that it is precise yet didactic, and extensive yet focused.

The prospective reader should have a good understanding of formal language theory, some knowledge on tree-automata, and a ground basis of logic. I believe this is a necessary *and sufficient* basis to be interested in this book. Anyone to which the title sounds interesting could (should) read Chapter 1 (the overview) and the Conclusion. This gives a broad picture of the field and the results of the past (and next) 25 years.

The writing style is, to say the least, precise. This implies a huge number of cross-references which are not always easy to follow. In such cases, completeness has been preferred to succinctness, making this book a great reference. However, when cross-references become too numerous, the authors do not hesitate to repeat succinctly a definition, helping the flow of reading. Most chapters are filled with (precisely delimited) digressions, which may open the view of the reader and contain possible new directions.