GRAPH STRUCTURE AND MONADIC SECOND-ORDER LOGIC

A Language-Theoretic Approach

The study of graph structure has advanced significantly in recent years: finite graphs can now be described algebraically, enabling them to be constructed out of more basic elements. One can obtain algebraic characterizations of tree-width and clique-width, two graph complexity measures that are important for the construction of polynomial-time graph algorithms. Separately the properties of graphs can be studied in a logical language called monadic second-order logic. In this book, these two features of graph structure are brought together for the first time in a presentation that unifies and synthesizes research over the last 25 years. The authors not only provide a thorough description of the theory, but also detail its applications, on the one hand to the construction of graph algorithms, and on the other to the extension of formal language theory to finite graphs. This extension combines algebraic notions (equational and recognizable sets) and logical ones (graph transformations specified by logical formulas). Applications of these tools to languages of words and terms are also presented.

Consequently the book will be of interest to graduate students and researchers in graph theory, finite model theory, formal language theory and complexity theory.

Encyclopedia of Mathematics and Its Applications

This series is devoted to significant topics or themes that have wide application in mathematics or mathematical science and for which a detailed development of the abstract theory is less important than a thorough and concrete exploration of the implications and applications.

Books in the **Encyclopedia of Mathematics and Its Applications** cover their subjects comprehensively. Less important results may be summarized as exercises at the ends of chapters. For technicalities, readers can be referred to the bibliography, which is expected to be comprehensive. As a result, volumes are encyclopedic references or manageable guides to major subjects.

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

All the titles listed below can be obtained from good booksellers or from Cambridge University Press.
For a complete series listing visit www.cambridge.org/mathematics.

94 S. R. Finch *Mathematical Constants*
95 Y. Jabri *The Mountain Pass Theorem*
96 G. Gasper and M. Rahman *Basic Hypergeometric Series, 2nd edn*
97 M. C. Pedicchio and W. Tholen (eds.) *Categorical Foundations*
98 M. E. H. Ismail *Classical and Quantum Orthogonal Polynomials in One Variable*
99 T. Mora *Solving Polynomial Equation Systems II*
100 E. Olivieri and M. Eulália Vares *Large Deviations and Metastability*
101 A. Kushner, V. Lychagin and V. Rubtsov *Contact Geometry and Nonlinear Differential Equations*
102 L. W. Beineke and R. J. Wilson (eds.) with P. J. Cameron *Topics in Algebraic Graph Theory*
103 O. J. Staffans *Well-Posed Linear Systems*
104 J. M. Lewis, S. Lakshmivarahan and S. K. Dhall *Dynamic Data Assimilation*
105 M. Lothaire *Applied Combinatorics on Words*
106 A. Markoe *Analytic Tomography*
107 P. A. Martin *Multiple Scattering*
108 R. A. Brualdi *Combinatorial Matrix Classes*
109 J. M. Borwein and J. D. Vanderwerff *Convex Functions*
110 M.-J. Lai and L. L. Schumaker *Spline Functions on Triangulations*
111 R. T. Curtis *Symmetric Generation of Groups*
112 H. Salzmann *et al. The Classical Fields*
113 S. Peszat and J. Zabczyk *Stochastic Partial Differential Equations with Lévy Noise*
114 J. Beck *Combinatorial Games*
115 L. Barreira and Y. Pesin *Nonuniform Hyperbolicity*
116 D. Z. Arov and H. Dym *J-Contractive Matrix Valued Functions and Related Topics*
117 R. Glowinski, J.-L. Lions and J. He *Exact and Approximate Controllability for Distributed Parameter Systems*
118 A. A. Borovkov and K. A. Borovkov *Asymptotic Analysis of Random Walks*
119 M. Deza and M. Dutour Sikirić *Geometry of Chemical Graphs*
120 T. Nishiura *Absolute Measurable Spaces*
121 M. Prest *Purity, Spectra and Localisation*
122 S. Khrushchev *Orthogonal Polynomials and Continued Fractions*
123 H. Nagamochi and T. Ibaraki *Algorithmic Aspects of Graph Connectivity*
124 F. W. King *Hilbert Transforms I*
125 F. W. King *Hilbert Transforms II*
126 O. Calin and D.-C. Chang *Sub-Riemannian Geometry*
127 M. Grabisch *et al. Aggregation Functions*
128 L. W. Beineke and R. J. Wilson (eds.) with J. L. Gross and T. W. Tucker *Topics in Topological Graph Theory*
129 J. Berstel, D. Perrin and C. Reutenauer *Codes and Automata*
130 T. G. Faticoni *Modules over Endomorphism Rings*
131 H. Morimoto *Stochastic Control and Mathematical Modeling*
132 G. Schmidt *Relational Mathematics*
133 P. Kornerup and D. W. Matula *Finite Precision Number Systems and Arithmetic*
134 Y. Crama and P. L. Hammer (eds.) *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*
135 V. Berthé and M. Rigo (eds.) *Combinatorics, Automata and Number Theory*
136 A. Kristály, V. D. Rădulescu and C. Varga *Variational Principles in Mathematical Physics, Geometry, and Economics*
137 J. Berstel and C. Reutenauer *Noncommutative Rational Series with Applications*
138 B. Courcelle and J. Engelfriet *Graph Structure and Monadic Second-Order Logic*
139 M. Fiedler *Matrices and Graphs in Geometry*
140 N. Vakil *Real Analysis through Modern Infinitesimals*
141 R. B. Paris *Hadamard Expansions and Hyperasymptotic Evaluation*
142 Y. Crama and P. L. Hammer *Boolean Functions*
143 A. Arapostathis, V. S. Borkar and M. K. Ghosh *Ergodic Control of Diffusion Processes*
144 N. Caspard, B. Leclerc and B. Monjardet *Finite Ordered Sets*
145 D. Z. Arov and H. Dym *Bitangential Direct and Inverse Problems for Systems of Integral and Differential Equations*
146 G. Dassios *Ellipsoidal Harmonics*

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Graph Structure and Monadic Second-Order Logic

A Language-Theoretic Approach

BRUNO COURCELLE

Université de Bordeaux

JOOST ENGELFRIET

Universiteit Leiden

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521898331

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Courcelle, B.
Graph structure and monadic second-order logic : a language-theoretic approach / Bruno Courcelle,
Joost Engelfriet.
p. cm. – (Encyclopedia of mathematics and its applications ; 138)
Includes bibliographical references and index.
ISBN 978-0-521-89833-1 (hardback)
1. Logic, Symbolic and mathematical – Graphic methods. I. Engelfriet, Joost. II. Title.
QA9.C748 2012
511.3–dc23 2012008159

ISBN 978-0-521-89833-1 Hardback

Additional resources for this publication at www.labri.fr/perso/courcell/TheBook.html

To Dominique and Louco

Cambridge University Press
978-0-521-89833-1 - Graph Structure and Monadic Second-Order Logic: A Language-Theoretic Approach
Bruno Courcelle and Joost Engelfriet
Frontmatter
More information

Contents

Cambridge University Press
978-0-521-89833-1 - Graph Structure and Monadic Second-Order Logic: A Language-Theoretic Approach
Bruno Courcelle and Joost Engelfriet
Frontmatter
More information

Foreword
by Maurice Nivat

The genesis of this great and beautiful book spans more than 20 years. It collects and unifies many theoretical notions and results published by Bruno Courcelle and others in a large number of articles.

The concept of a language to communicate with a computer, a machine or any kind of device performing operations is at the heart of Computer Science, a field that has truly thrived with the emergence of symbolic programming languages in the 1960s. Formalizing the algorithms that enable computers to calculate an intended result, to control a machine or a robot, to search and find the relevant information in response to a query, and even to imitate the human brain in actions such as measuring risk and making decisions, is the main activity of computer scientists as well as of ordinary computer users.

The languages designed for these tasks, which number by thousands, are defined in the first place by syntactic rules that construct sets of words and to which are then attached meanings. This understanding of a language was first conceived by structural linguists, in particular Nicolaï Troubetskoï, Roman Jacobson and Noam Chomsky, and has transformed Linguistics, the study of natural languages, by giving it new directions. It has also been extended to programming languages, which are artificial languages, and to the Lambda Calculus, one of many languages devised by logicians, among whom we can cite Kurt Gödel, Alonzo Church and Alan Turing, who aspired to standardize mathematical notation and to mechanize proofs. This same idea has inspired all research on computation theory and programming. Thanks to the results of this research, planes can fly with continuously monitored flight parameters, providing us with unprecedented reliability: this is so because millions of lines of code have been formally proved to be correct.

Words are strings of symbols taken from finite alphabets. They constitute the basic elements. They can represent all the information one might wish to capture, use, process, disseminate or share in a world that is fast becoming more and more "digital," as Gérard Berry emphasized recently in his lectures at the Collège de France.

Most information, though represented always by words, is nevertheless structured hierarchically and can thus be presented in a natural way as a tree or as a graph. Most of the countless electronic chips that make up computers but are also used in an ever-growing number of other machines as well, from washing machines to nuclear power plants, calculate on graphs: by connecting vertices, their edges can represent virtually any relationship of subordination, analogy, neighborhood or causality. From the early 1960s, the algorithms for graphs (and for trees, which are particular graphs) have been developed swiftly, and most of the current computing applications are based on these algorithms. Thousands of them have been designed by numerous researchers and engineers, and they fuel a burgeoning literature.

It was around 1980 that Bruno Courcelle, a former student of the prestigious Ecole Normale Supérieure (Rue d'Ulm in Paris), a logician by training, and a young but already established researcher, tackled a seemingly impossible task: to build a theory of tree languages that would classify all of these algorithms and present them in a unified and rational way. Bruno Courcelle is not a "problem solver" who happened to discover more-or-less elegant and clever answers to questions; he likes well-founded and harmonious theories, and is always looking for unifying concepts. Armed with a knowledge of logic and with a familiarity with Fundamental Computer Science, and in particular with Formal Language Theory which he gained during his years as a researcher at INRIA (Institut National de Recherche en Informatique et Automatique) while preparing his thesis, Bruno Courcelle got down to work with perseverance and determination.

Upon his arrival at Bordeaux-1 University in 1979 (LaBRI, the Laboratoire Bordelais de Recherche en Informatique, was created in 1988), Bruno Courcelle found an excellent work environment. The concept of attribute grammar, which is important in compilation, provided the model that he has used to develop an algebraic approach to graph grammars and a logical approach to the proof of properties of the graphs they generate. The first published work he devoted to attribute grammars is the source of the theory presented here, based on Logic and Universal Algebra.

The impact of his work surpassed all expectations, even taking into consideration the remarkable qualities of method and rigor that characterize Bruno Courcelle. For when the first elements of his theory began to spread among those who work on designing and improving graph algorithms, these researchers realized that Bruno Courcelle had provided a convenient formal framework in which many problems could be solved. In particular, Bruno's theory brought a logical lightening to the profound works of Paul Seymour and his collaborators on graph minors. Other researchers have been inspired by his theory to study new problems and invent new algorithms. A daunting theory that was originally seen as arcane and abstract proved to be rich and fertile. In 2004, Bruno Courcelle was awarded by INIST (an institute depending on the Centre National pour la Recherche Scientifique) and the ISI-Web of Science (Thomson-Reuters) the prize of the "most cited researcher in Computer Science in France."

I am not going to analyze his work further; in any case Chapter 1 is a long overview that is perfectly readable, even by those who are well versed neither in algebra nor in mathematical logic. I would rather emphasize that the work of a computer scientist, as of any scientist, can be very diverse. The quest for new results is one endeavor, but bringing up to date the underlying structures, unifying concepts and simplifying the presentation of results, is quite another. The rapid, sometimes frantic, growth of publications in Computer Science has led most researchers to choose the former in pursuit of better results and more efficient algorithms. It gives me great pleasure to preface a work that is of the opposite nature: a long book produced by a comprehensive study, which leads to a very interesting result: the formation of a theory that brings order, that explains and simplifies a vast collection of results obtained by others and, at the same time, that proposes methods, yields results and raises new questions.

While still a student I was very struck when I first read André Lichnérowicz's book on linear algebra. I had already taken courses in linear algebra which, I confess, were not very helpful, and suddenly this book made everything clear. The mysterious operations which we were taught to perform on the square tables called "determinants" started making sense; the concepts of both vector space and the dimension of a vector subspace finally allowed me to understand what it meant to agonize over a determinant and, moreover, why this notion is important. Lichnérowicz's book is a classic that has enabled generations of students to learn linear algebra with ease, and it has become a mathematical tool widely used by engineers and technicians who are not professional mathematicians. I believe that this book will get the same reputation, quickly become a classic and provide an easy access to the burgeoning world of graph algorithms and its numerous applications throughout the sciences and beyond.

The comments above were written two years ago, when Bruno Courcelle's book was only 500 pages long, and I cannot change what I wrote then: it is a great and beautiful book that is going to take its place very soon on the library shelves of all the departments of Computer Science around the world. But now the book is 700 pages long and has two authors, Bruno and Joost Engelfriet. What happened is that Bruno sent the previous version to Joost to read and suggest corrections and improvements. Joost is a very old acquaintance of both Bruno and myself, and we have always known him as one of the most knowledgeable researchers in the field of grammars, automata and transducers on words and trees. And Joost had so many things to suggest that it is another book that I present today: thicker, with new results and a number of proofs that have been replaced by simpler and more elegant ones. Obviously the cooperation between Bruno and Joost was a very fruitful one indeed.

Knowing Joost as I do, this is not a surprise: when I asked him to referee papers submitted to the journal *Theoretical Computer Science*, in most cases Joost's report was longer and sometimes richer than the refereed article. His comments always led to a major improvement of the original text. Clearly Bruno's manuscript inspired Joost. And we all have to be grateful to him for, as usual, his comments and the work he did on the manuscript resulted in a major improvement and a sizable enlargement.

Thus today I am very happy to thank the two authors of this beautiful book, which I consider to be a wonderful source of knowledge in Computer Science. It is a theoretical book, and for that reason some people may find it hard to read, but reading it is worth the pain, because the formalism introduced and the methods presented have already led to many new algorithms on graphs (as the number of citations of Bruno's published papers show) and they will lead to many others in the future. To anyone interested in graph algorithms I can only recommend that they read this book first.

For indeed this book lies at the very heart of Computer Science, which is the expressiveness of the languages used to represent and manipulate information and information structures, graphs being among the most widely used information structures. Progress in the efficiency, liability and simplicity of algorithms comes mainly from the use of better representations, better structures and a better understanding of the different ways in which one can describe sets of data and express their properties. This book provides a huge number of conceptual tools to design and study graph algorithms that no one should ignore.

In the name of the young but fast-growing science that in French we call *Informatics*, in the name of all future researchers in this field, I just say to Bruno and Joost: Thanks, you have done a good job!