

Advanced Information and Knowledge Processing

Series Editors

Professor Lakhmi Jain
Lakhmi.jain@unisa.edu.au
Professor Xindong Wu
xwu@cs.uvm.edu

For other titles published in this series, go to
<http://www.springer.com/4738>

Michel Chein Marie-Laure Mugnier

Graph-based Knowledge Representation

Computational Foundations of Conceptual Graphs

 Springer

Michel Chein
Laboratory of Informatics, Robotics, and
Micro-electronics (LIRMM)
FRANCE

Marie-Laure Mugnier
Laboratory of Informatics, Robotics, and
Micro-electronics (LIRMM)
FRANCE

AI&KP ISSN 1610-3947

ISBN: 978-1-84800-285-2

e-ISBN: 978-1-84800-286-9

DOI 10.1007/978-1-84800-286-9

British Library Cataloguing in Publication Data

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

Library of Congress Control Number: 2008937554

© Springer-Verlag London Limited 2009

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

Springer Science+Business Media
springer.com

Preface

This book provides a definition and study of a knowledge representation and reasoning formalism stemming from conceptual graphs, while focusing on the computational properties of this formalism.

Knowledge can be symbolically represented in many ways. The knowledge representation and reasoning formalism presented here is a graph formalism – knowledge is represented by *labeled graphs*, in the graph theory sense, and reasoning mechanisms are based on *graph operations*, with graph homomorphism at the core.

This formalism can thus be considered as related to semantic networks. Since their conception, semantic networks have faded out several times, but have always returned to the limelight. They faded mainly due to a lack of formal semantics and the limited reasoning tools proposed. They have, however, always rebounded because labeled graphs, schemas and drawings provide an intuitive and easily understandable support to represent knowledge.

This formalism has the visual qualities of any graphic model, and it is *logically founded*. This is a key feature because logics has been the foundation for knowledge representation and reasoning for millennia. The authors also focus substantially on *computational facets* of the presented formalism as they are interested in knowledge representation and reasoning formalisms upon which knowledge-based systems can be built to solve real problems. Since object structures are graphs, naturally graph homomorphism is the key underlying notion and, from a computational viewpoint, this moors calculus to combinatorics and to computer science domains in which the algorithmic qualities of graphs have long been studied, as in databases and constraint networks.

The main notions are intuitively introduced from a knowledge representation viewpoint, but the authors have also tried to give precise definitions of these notions along with complete proofs of the theorems (readers only require a few simple mathematical notions, which are summarized in the appendix).

This book does not present a methodology for knowledge representation nor describe actual applications (except for one chapter) or software tools. It presents theoretical bases that merge numerous previous studies carried out in the conceptual

graph domain since Sowa's seminal book and links basic reasoning issues to other fundamental problems in computer science.

In a nutshell, the authors have attempted to answer the following question: *"How far is it possible to go in knowledge representation and reasoning by representing knowledge with graphs (in the graph theory sense) and reasoning with graph operations?"*

Organization

The book is divided into 3 parts, 13 chapters and 1 appendix.

The first part is devoted to the kernel of the formalism. In Chap. 2, Basic conceptual Graphs (BGs) are defined. BGs are structured by a subsumption relation defined by the homomorphism notion between BGs, as well as by elementary specialization and generalization operations. In Chap. 3, Simple conceptual Graphs (SGs) are introduced. SGs can be sketchily defined as BGs augmented with equality. Chapter 4 is devoted to set and FOL semantics for SGs. The soundness and completeness of homomorphism with respect to entailment relation and FOL deduction is proven. Chapter 5 relates BG homomorphism with homomorphisms of other structures (e.g., hypergraphs, relational structures, conjunctive database queries) and solving a constraint network.

The second part develops computational aspects of basic conceptual graphs. Chapters 6, and 7 are devoted to algorithms for BG homomorphism. Chapter 8 presents other specialization and generalization operations that are not covered in the first part, e.g., least generalization, maximal join and other extended joins.

The third part pools the kernel extensions. All of these extensions are provided with logically sound and complete reasoning mechanisms based on graph homomorphism. Chapter 9 focuses on nested conceptual graphs, which are hierarchically structured graphs. In Chap. 10, the important rule notion, which allows representation of implicit knowledge, for instance, is studied. The introduction of rules gives the powerfulness of a computability model to the formalism. Positive and negative constraints are considered in Chap. 11, and the BG family of models, combining facts, rules and constraints, is presented. Chapter 12 is devoted to negation in conceptual graphs. The last chapter presents semantic annotations—a currently favored and potentially fruitful application.

Finally, the Appendix summarizes the basic mathematical notions used in the book.

Each chapter begins with an overview of its content. It ends with bibliographical notes devoted mainly to studies in the conceptual graph domain. Concerning the references, the authors chose to cite studies from other domains throughout the text and provide the conceptual graph references in the bibliographical notes.

Implementation

This book is about a KR formalism that can serve as a theoretical foundation for knowledge-based systems. A distinction is made between a KR formalism (e.g., a fragment of first order logic) and a KR programming language (e.g., PROLOG). Even when a KR programming language is based on a KR formalism, the KR language presents variations to the KR formalism (limitations, e.g., Horn clauses or extensions, second order features, etc.), and it has a concrete syntax, contains specific programming constructs (e.g., the cut in PROLOG) and is equipped with software tools. Between a KR formalism and a programming language implementing this formalism, there may be KR tools and platforms. This is the current situation of the graph-based KR formalism presented in the book. CoGITaNT [cog01a] contains a library of C++ classes that implement most of the notions and reasoning mechanisms presented here. CoGITaNT also contains a graphical interface and a client-server architecture. CoGITaNT thus allows a programmer to build knowledge-based systems grounded on the formalism presented in this book. Let us also mention COGUI [cog01b], a graphical user interface dedicated to the construction of a knowledge base and which integrates CoGITaNT.

Audience

The book is intended for computer science students, engineers and researchers. Some of the materials presented in this book have been used for several years at different academic levels, ranging from AI courses for graduate students to professional and research master's level courses. The mathematical prerequisites are minimal, and the Appendix outlines the mathematical notions used in the book.

Here are some suggestions for reading paths of this book depending on readers' specific interests.

Chapter 1 (introduction), Chap. 2 (basic conceptual graphs), Chap. 3 (simple conceptual graphs), the two first sections of Chap. 4 (set and first order logic semantics) are the core chapters and should be considered as the basis.

For programming and algorithmic purposes the following materials can be added to the base: Chap. 6 (basic algorithms for BG homomorphism), the last section of Chap. 5 (relationships with constraint programming techniques), Chap. 7 (techniques for trees and other tractable cases), Chap. 8 (algorithms for other specialization/generalization operations, especially maximal joins), Chap. 10 (rule processing), and Chap. 12 (algorithms for processing BGs with atomic negation).

For modeling purposes the following parts can be added to the base: Chap. 9 (nested conceptual graphs), Chap. 10 (definition and use of rules), Chap. 11 (definition and use of constraints and their combination with rules), Chap. 12 (definition and use of atomic negation) and Chap. 13 (semantic annotations).

For more theory-oriented readers, expressivity, decidability and complexity results, as well as stating equivalence with problems of other domains, are presented throughout the book, except in the last chapter.

Acknowledgments

We are indebted to John Sowa, who is the founding father of conceptual graphs, and the present book would not have existed without his pioneering work [Sow84].

We began working on conceptual graphs in 1992, and over the years many of our colleagues and students have contributed, in one way or another, to this work.

We would like to thank:

Friends and colleagues who have supported our approach over the years: Franz Baader, Marie-Christine Rousset, Pierre Marquis, Fritz Lehmann, Daniel Kayser, Michel Habib, Guy Chaty;

Colleagues or students with whom we have directly collaborated on topics covered in this book: Jean-François Baget, Boris Carbonneill, Olivier Cogis, David Genest, Olivier Guinaldo, Ollivier Haemmerlé, Gwen Kerdiles, Michel Leclère, Anne Preller, Eric Salvat, Geneviève Simonet, Rallou Thomopoulos;

Colleagues from the conceptual graph community with whom we have had fruitful discussions: Tru Cao, Fritjof Dau, Gerard Ellis, Jean Fargues, Bikash Gosh, Fritz Lehmann, Robert Levinson, Dickson Lukose, Guy Mineau, John Sowa, Mark Willems, Vilas Wuwongse.

Our approach has benefited from numerous collaborations on applied research projects, and we would like to thank: Christian Baylon, Alain Bonnet, Bernard Botella, Olivier Corby, Patrick Courounet, Rose Dieng, Steffen Lalande, Philippe Martin, Patrick Taillibert.

We acknowledge Jean-François Baget, David Genest, Alain Gutierrez, Michel Leclère, Khalil Ben Mohamed, Nicolas Moreau, Fatiha Saïs, Eric Salvat, who reviewed parts of this book, with special thanks to Geneviève Simonet (some of her comments could be the seeds for a new book ...).

We would also like to thank David Manley for rectifying our Franglais and Beverly Ford and her team at Springer London for their support.

Obviously, all remaining errors are our fault alone. We welcome corrections and suggestions on all aspects of the book; please send these to us at:

Michel.Chein@lirmm.fr and Marie-Laure.Mugnier@lirmm.fr. A web site companion to the book can also be queried at:
<http://www.lirmm.fr/gbkrbook>

La Boissière,

Michel Chein and Marie-Laure Mugnier
 March 2008

Contents

1	Introduction	1
1.1	Knowledge Representation and Reasoning	1
1.1.1	Knowledge-Based Systems	2
1.1.2	Requirements for a Knowledge Representation Formalism	3
1.2	Conceptual Graphs	8
1.2.1	Basic Notions	8
1.2.2	Subsumption and Homomorphism	9
1.2.3	Formal Semantics	11
1.2.4	Full CGs	12
1.3	A Graph-Based Approach to KR	13
1.3.1	Motivations	13
1.3.2	Extensions of the Basic Formalism	14
1.3.3	Several Approaches to CGs	16

Part I Foundations: Basic and Simple Conceptual Graphs

2	Basic Conceptual Graphs	21
2.1	Definition of Basic Conceptual Graphs (BGs)	22
2.1.1	Vocabulary	22
2.1.2	Basic Conceptual Graphs	25
2.1.3	SubBGs and PseudoBGs	29
2.2	BG Homomorphism	30
2.2.1	Subsumption and Homomorphism	30
2.2.2	Bijjective Homomorphisms and Isomorphisms	33
2.2.3	BG Queries and Answers	34
2.3	BG Subsumption Properties	35
2.3.1	Subsumption Preorder	35
2.3.2	Irredundant BGs	37
2.4	Generalization and Specialization Operations	40
2.4.1	Elementary Generalization Operations for BGs	40
2.4.2	Generalization and Homomorphism	42

2.4.3	Elementary Specialization Operations	45
2.4.4	Specialization and Homomorphism	48
2.5	Normal BGs	49
2.5.1	Definition of Normal BGs	49
2.5.2	Elementary Operations for Normal BGs	52
2.6	Complexity of Basic Problems	54
2.7	Bibliographic Notes	56
3	Simple Conceptual Graphs	59
3.1	Introduction	60
3.2	Vocabulary	62
3.3	Simple Conceptual Graphs (SGs)	66
3.4	Generalization and Specialization Operations	69
3.5	Standard and Normal SGs	71
3.6	Coref-Homomorphism	72
3.7	Antinormal Form and Homomorphism	76
3.8	Bibliographic Notes	80
4	Formal Semantics of SGs	83
4.1	Model Semantic	84
4.1.1	Models of SGs	84
4.1.2	Soundness and Completeness of (coref) Homomorphism	87
4.2	Logical Semantic	89
4.2.1	The FOL Semantic Φ	89
4.2.2	Model Semantic and Φ	91
4.3	Positive, Conjunctive, and Existential Fragment of FOL	92
4.3.1	Ordered Language and \mathcal{L} -Substitution	93
4.3.2	Soundness and Completeness Revisited	94
4.3.3	Relationships Between SGs and FOL(\wedge, \exists)	95
4.3.4	Another FOL Semantic: Ψ	98
4.4	Note on the Relationships Between Description Logics and Conceptual Graphs	100
4.5	Bibliographic Notes	104
5	BG Homomorphism and Equivalent Notions	105
5.1	Conceptual Graphs and Conceptual Hypergraphs	107
5.1.1	Conceptual Hypergraphs	107
5.1.2	Different Kinds of Vocabularies	110
5.2	Graphs	113
5.2.1	From Graphs to BGs	114
5.2.2	From BGs to Graphs	117
5.3	Relational Structures and Databases	119
5.3.1	Relational Structures and BGs	119
5.3.2	Conjunctive Queries and BGs	120
5.4	Constraint Satisfaction Problem	124

5.4.1	Definition of CSP	124
5.4.2	From CSP to BGs	127
5.4.3	From BGs to CSP	128
5.5	Bibliographic Notes	132

Part II Computational Aspects of Basic Conceptual Graphs

6	Basic Algorithms for BG Homomorphism	135
6.1	Algorithms for BG Homomorphisms	135
6.1.1	Basic Backtrack Algorithms	136
6.1.2	Backtrack Improvements	141
6.2	Constraint Processing	151
6.2.1	A Panorama of Constraint Processing Techniques	151
6.2.2	Arc-Consistency	154
6.2.3	Forward Checking	159
6.3	Label Comparison	161
6.3.1	Basic Data Structures and Algorithms	163
6.3.2	Related Problems	164
6.3.3	Tree Orders	166
6.3.4	Partition in Chains	167
6.3.5	Lattices	168
6.4	Bibliographic Notes	169
7	Tractable Cases	171
7.1	Introduction	171
7.1.1	About Tractability	172
7.1.2	The Structure of the <i>Target</i> BG is of No Help	172
7.2	Tractability Based on the Multigraph-Acyclicity of the <i>Source</i> BG	174
7.2.1	Acyclic Multigraphs and Trees	174
7.2.2	BGs Trivially Logically Equivalent to Acyclic BGs	184
7.3	Tractability Based on the Hypergraph-Acyclicity of the Source BG	185
7.3.1	Use of a Join Tree	187
7.3.2	Construction of a Join Tree	191
7.3.3	Equivalence with the Existential Conjunctive Guarded Fragment	193
7.4	Generalizations of Graph-Acyclicity and Hypergraph-Acyclicity	198
7.4.1	Graphs and Treewidth	198
7.4.2	Hypergraphs and Hypertreewidth	200
7.4.3	Expressivity Results	201
7.5	What About Labels?	202
7.6	Complementary Notes	204

8	Other Specialization/Generalization Operations	207
8.1	The Least Generalization and Greatest Specialization of Two BGs	208
8.2	Basic Compatibility Notions and Maximal Joins	212
8.2.1	Compatible Node Set	212
8.2.2	Maximal Join	215
8.3	Compatible Partitions and Extended Join	222
8.3.1	Compatible C-Partition and R-Partition	222
8.3.2	Extended Join	226
8.3.3	Join According to a Compatible Pair of C-Partitions	226
8.4	\mathcal{G} -Specializations	228
8.4.1	Surjective Homomorphism	228
8.4.2	Union	229
8.4.3	Inductive Definition of BGs	231
8.4.4	\mathcal{G} -Specializations	232
8.5	Type Expansion and Contraction	235
8.6	Bibliographic Notes	242

Part III Extensions

9	Nested Conceptual Graphs	247
9.1	Introduction	248
9.2	Nested Basic Graphs (NBGs)	249
9.3	Nested Graphs (NGs)	256
9.4	Nested Typed Graphs	258
9.5	The Semantics Φ_N	262
9.5.1	Definition of Φ_N	263
9.5.2	Soundness and Completeness	265
9.6	Representation of Nested Typed Graphs by BGs	267
9.7	Bibliographic Notes	270
10	Rules	273
10.1	Definition and Logical Semantics of a Rule	273
10.1.1	Logical Semantics of a Rule	276
10.1.2	Rule as a Bicolored Graph	277
10.2	Forward Chaining	278
10.2.1	Rule Application	279
10.2.2	Derivation and Deduction	281
10.2.3	Non-Redundant Rule Application	284
10.2.4	Soundness and Completeness of Forward Chaining	286
10.3	Backward Chaining	291
10.3.1	Outline of the Backward Chaining Mechanism	292
10.3.2	Piece Resolution	294
10.3.3	Soundness and Completeness of Backward Chaining	298
10.4	Computational Complexity of \mathcal{FR} -DEDUCTION with Rules	301
10.4.1	Semi-Decidability of \mathcal{FR} -DEDUCTION with Rules	301

10.4.2	Decidable Cases	303
10.5	Bibliographic Notes	307
11	The BG Family: Facts, Rules and Constraints	311
11.1	Overview of the BG Family	311
11.2	\mathcal{FC} : Facts and Constraints	313
11.2.1	Positive and Negative Constraints	313
11.2.2	Translation to FOL	316
11.2.3	Complexity of Consistency and Deduction	318
11.3	Combining Rules and Constraints	321
11.3.1	\mathcal{FRC} : Constraints and Inference Rules	321
11.3.2	\mathcal{FEC} : Constraints and Evolution Rules	325
11.3.3	\mathcal{FREC} : Constraints, Inference and Evolution Rules	326
11.3.4	Complexity of Combining Rules and Constraints	326
11.4	Complexity in $\mathcal{FRC}/\mathcal{FEC}/\mathcal{FREC}$ for Particular Cases of Rules and Constraints	327
11.4.1	Particular Cases of Rules	327
11.4.2	Particular Cases of Constraints	330
11.5	Bibliographic Notes	334
12	Conceptual Graphs with Negation	337
12.1	Full Conceptual Graphs	338
12.1.1	Existential Graphs: a Diagrammatical System for Logics	338
12.1.2	Full Conceptual Graphs (FCGs)	340
12.1.3	Logical Semantics of FCGs	342
12.1.4	Equivalence of CGs and FOL	344
12.1.5	FCG Calculus	346
12.1.6	Dau's FCGs	347
12.1.7	Kerdiles' FCGs	348
12.2	Conceptual Graphs with Atomic Negation	350
12.2.1	Polarized Graphs	350
12.2.2	Handling Coreference and Difference	355
12.2.3	PG-DEDUCTION and Equivalent Problems	359
12.2.4	Complexity of PG-DEDUCTION	360
12.2.5	Special Cases with Lower Complexity for PG-DEDUCTION	362
12.2.6	Algorithmic Improvements	367
12.2.7	Querying Polarized Graphs	371
12.2.8	Note on Negation and Rules	374
12.3	Bibliographic Notes	375
13	An Application of Nested Typed Graphs: Semantic Annotation Bases	377
13.1	Annotation	377
13.1.1	Annotations, Metadata and Resources	377
13.1.2	Examples of Annotation Base Uses	378
13.1.3	Components of an Annotation System	380

13.2	Annotation Base	381
13.2.1	Exact Knowledge	382
13.2.2	Modules	383
13.2.3	Plausible Knowledge	384
13.3	Querying an Annotation Base	385
13.3.1	Exact Search	386
13.3.2	Approximate Search	386
13.4	Annotation and the Semantic Web	388
13.5	Conclusion	390
A	Mathematical Background	393
A.1	Sets and Relations	394
A.1.1	Sets and Elements	394
A.1.2	Relations and Mappings	395
A.2	Graphs	397
A.2.1	Directed Graphs	397
A.2.2	Homomorphism	399
A.2.3	Different Sorts of Graphs	400
A.2.4	Hypergraphs	403
A.3	Ordered Sets	403
A.3.1	Basic Notions	403
A.3.2	Lattices	405
A.4	First Order Logic (FOL)	406
A.4.1	Syntax	406
A.4.2	Semantics and Models	408
A.4.3	Clausal Form	409
A.4.4	Resolution and SLD-Resolution	409
A.5	Algorithm and Problem Complexity	410
	References	413
	Index	423