THE DESCRIPTION LOGIC HANDBOOK

Description Logics are embodied in several knowledge-based systems and are used to develop various real-life applications. *The Description Logic Handbook* provides a thorough account of the subject, covering all aspects of research in this field; namely, theory, implementation, and applications. Its appeal will be broad, ranging from more theoretically oriented readers to those with more practically oriented interests who need a sound and modern understanding of knowledge representation systems based on Description Logics. As well as general revision throughout the book, this new edition presents a new chapter on ontology languages for the Semantic Web, an area of great importance for the future development of the web. In sum, the book will serve as a unique reference for the subject, and can also be used for self-study or in conjunction with Knowledge Representation and Artificial Intelligence courses.

THE DESCRIPTION LOGIC HANDBOOK

Theory, implementation, and applications

Edited by

FRANZ BAADER
DIEGO CALVANESE
DEBORAH L. McGUINNESS
DANIELE NARDI
PETER F. PATEL-SCHNEIDER

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

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/9780521876254

First published 2003
Second edition published 2007
Reprinted 2008
Paperback edition 2010

Printed in the United Kingdom at the University Press, Cambridge

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

ISBN 978-0-521-87625-4 hardback
ISBN 978-0-521-15011-8 paperback

Contents

List of contributors *page* ix
Preface to the second edition xiii
Preface xv
1 An Introduction to Description Logics *D. Nardi and R. J. Brachman* 1
1.1 Introduction 1
1.2 From networks to Description Logics 5
1.3 Knowledge representation in Description Logics 13
1.4 From theory to practice: Description Logic systems 17
1.5 Applications developed with Description Logic systems 22
1.6 Extensions of Description Logics 32
1.7 Relationship to other fields of Computer Science 39
1.8 Conclusion 42
Part I Theory **45**
2 Basic Description Logics *F. Baader and W. Nutt* 47
2.1 Introduction 47
2.2 Definition of the basic formalism 50
2.3 Reasoning algorithms 81
2.4 Language extensions 98
3 Complexity of Reasoning *F. M. Donini* 105
3.1 Introduction 105
3.2 OR-branching: finding a model 109
3.3 AND-branching: finding a clash 117
3.4 Combining sources of complexity 124
3.5 Reasoning in the presence of axioms 127
3.6 Undecidability 133
3.7 Reasoning about individuals in ABoxes 140
3.8 Discussion 144

3.9 A list of complexity results for subsumption and satisfiability 145
4 Relationships with other Formalisms *U. Sattler, D. Calvanese, and R. Molitor* 149
4.1 AI knowledge representation formalisms 149
4.2 Logical formalisms 161
4.3 Database models 174
5 Expressive Description Logics *D. Calvanese and G. De Giacomo* 193
5.1 Introduction 193
5.2 Correspondence between Description Logics and Propositional Dynamic Logics 195
5.3 Functional restrictions 202
5.4 Qualified number restrictions 209
5.5 Objects 213
5.6 Fixpoint constructs 217
5.7 Relations of arbitrary arity 221
5.8 Finite model reasoning 226
5.9 Undecidability results 232
6 Extensions to Description Logics *F. Baader, R. Küsters, and F. Wolter* 237
6.1 Introduction 237
6.2 Language extensions 238
6.3 Non-standard inference problems 270
Part II Implementation **283**
7 From Description Logic Provers to Knowledge Representation Systems *D. L. McGuinness and P. F. Patel-Schneider* 285
7.1 Introduction 285
7.2 Basic access 287
7.3 Advanced application access 290
7.4 Advanced human access 295
7.5 Other technical concerns 301
7.6 Public relations concerns 301
7.7 Summary 303
8 Description Logic Systems *R. Möller and V. Haarslev* 304
8.1 New light through old windows? 304
8.2 The first generation 305
8.3 Second generation Description Logic systems 313
8.4 The next generation: FaCT, DLP and RACER 324
8.5 Lessons learned 327

9 Implementation and Optimization Techniques *I. Horrocks* 329
9.1 Introduction 329
9.2 Preliminaries 331
9.3 Subsumption-testing algorithms 336
9.4 Theory versus practice 341
9.5 Optimization techniques 347
9.6 Discussion 371
Part III Applications **375**
10 Conceptual Modeling with Description Logics *A. Borgida and R. J. Brachman* 377
10.1 Background 377
10.2 Elementary Description Logic modeling 379
10.3 Individuals in the world 381
10.4 Concepts 384
10.5 Subconcepts 387
10.6 Modeling relationships 390
10.7 Modeling ontological aspects of relationships 392
10.8 A conceptual modeling methodology 399
10.9 The ABox: modeling specific states of the world 399
10.10 Conclusions 401
11 Software Engineering *C. A. Welty* 402
11.1 Introduction 402
11.2 Background 402
11.3 LaSSIE 403
11.4 CODEBASE 408
11.5 CSIS and CBMS 409
12 Configuration *D. L. McGuinness* 417
12.1 Introduction 417
12.2 Configuration description and requirements 419
12.3 The PROSE and QUESTAR family of configurators 433
12.4 Summary 434
13 Medical Informatics *A. Rector* 436
13.1 Background and history 437
13.2 Example applications 441
13.3 Technical issues in medical ontologies 447
13.4 Ontological issues in medical ontologies 453
13.5 Architectures: terminology servers, views, and change management 456
13.6 Discussion: key lessons from medical ontologies 457

14 OWL: a Description-Logic-Based Ontology Language for the Semantic Web *I. Horrocks, P. F. Patel-Schneider, D. L. McGuinness, and C. A. Welty* 458
14.1 Background and history 458
14.2 Steps towards integration with the Semantic Web: OIL and DAML+OIL 461
14.3 Full integration into the Semantic Web: OWL 467
14.4 Summary 484
15 Natural Language Processing *E. Franconi* 487
15.1 Introduction 487
15.2 Semantic interpretation 488
15.3 Reasoning with the logical form 492
15.4 Knowledge-based natural language generation 497
16 Description Logics for Databases *A. Borgida, M. Lenzerini, and R. Rosati* 500
16.1 Introduction 500
16.2 Data models and Description Logics 504
16.3 Description Logics and database querying 513
16.4 Data integration 517
16.5 Conclusions 523
Appendix Description Logic Terminology *F. Baader* 525
A.1 Notational conventions 525
A.2 Syntax and semantics of common Description Logics 526
A.3 Additional constructors 531
A.4 A note on the naming scheme for Description Logics 534

Bibliography 537
Index 593

Contributors

Franz Baader
Institut für Theoretische Informatik
Fakultät Informatik
TU Dresden
01062 Dresden, Germany
`http://wwwtcs.inf.tu-dresden.de/~baader/`

Alex Borgida
Department of Computer Science
Rutgers University
Piscataway, NJ 08855, U.S.A.
`http://www.cs.rutgers.edu/~borgida/`

Ronald J. Brachman
Yahoo!
`http://www.brachman.org/`

Diego Calvanese
Faculty of Computer Science
Free University of Bozen-Bolzano
Piazza Domenicani 3, 39100 Bolzano, Italy
`http://www.inf.unibz.it/~calvanese/`

Giuseppe De Giacomo
Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, 00198 Roma, Italy
`http://www.dis.uniroma1.it/~degiacomo/`

Francesco M. Donini
Political Science Department, Information Systems Research Group
University of Tuscia
Via San Carlo 32, 01100 Viterbo, Italy
`http://dee.poliba.it/dee-web/doniniweb/donini.html`

Enrico Franconi
Faculty of Computer Science
Free University of Bozen-Bolzano
Piazza Domenicani 3, 39100 Bolzano, Italy
http://www.inf.unibz.it/~franconi/

Volker Haarslev
Computer Science Department
Concordia University
1455 de Maisonneuve Blvd. W., Montreal, Quebec H3G IM8, Canada
http://www.cs.concordia.ca/~faculty/haarslev/

Ian Horrocks
School of Computer Science
University of Manchester
Manchester M13 9PL, U.K.
http://www.cs.man.ac.uk/~horrocks/

Ralf Küsters
Institut für Informatik und Praktische Mathematik
Christian-Albrechts-Universität zu Kiel
Olshausenstraße 40, 24098 Kiel, Germany
http://www.ti.informatik.uni-kiel.de/~kuesters/

Maurizio Lenzerini
Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, 00198 Roma, Italy
http://www.dis.uniroma1.it/~lenzerini/

Deborah L. McGuinness
Knowledge Systems Laboratory
Gates Building 2A, Stanford University
Stanford, CA 94305-9020, U.S.A.
http://ksl.stanford.edu/people/dlm/

Ralf Molitor
Swiss Life
IT Research and Development Group
General Guisan Quai 40, CH-8002 Zürich, Switzerland
http://research.swisslife.ch/~molitor/

Ralf Möller
Computer Science Department
University of Hamburg
Vogt-Kölln-Straße 30, 22527 Hamburg, Germany
http://kogs-www.informatik.uni-hamburg.de/~moeller/

Daniele Nardi
Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, 00198 Roma, Italy
http://www.dis.uniroma1.it/~nardi/

Werner Nutt
Faculty of Computer Science
Free University of Bozen-Bolzano
Piazza Domenicani 3, 39100 Bolzano, Italy
http://www.inf.unibz.it/~nutt/

Peter F. Patel-Schneider
Bell Labs Research
600 Mountain Avenue
Murray Hill, NJ 07974, U.S.A.
http://www.bell-labs.com/user/pfps/

Alan Rector
Medical Informatics Group
Department of Computer Science
University of Manchester
Manchester M13 9PL, U.K.
http://www.cs.man.ac.uk/mig/

Riccardo Rosati
Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, 00198 Roma, Italy
http://www.dis.uniroma1.it/~rosati/

Ulrike Sattler
School of Computer Science
University of Manchester
Manchester M13 9PL, U.K.
http://www.cs.man.ac.uk/~sattler/

Christopher A. Welty
Knowledge Structures Group
IBM Watson Research Center
19 Skyline Dr., Hawthorne, NY 10532, U.S.A.

Frank Wolter
Department of Computer Science
University of Liverpool
Ashton Building, Ashton Street
Liverpool L69 3BX, U.K.
http://www.csc.liv.ac.uk/~frank/

Preface to the second edition

Since the publication of the first edition of *The Description Logic Handbook* in 2003, the interest in Description Logics (DL) has steadily increased. This applies both to the number of active DL researchers working on DL theory and implementations of reasoning services, and to the number of applications based on DL technology. One effect of this growing interest was that the first edition of the *Handbook* has gone through quite a number of reprints. Another effect is, of course, that in the last three years there have been interesting new developments in the three areas (theory, implementation, and applications) that the *Handbook* covers. Despite that, we feel that most chapters of the *Handbook* still provide a good introduction to the field and lay a solid foundation that enables the reader to understand and put into context the research articles describing results since 2003. For this reason, we have decided to leave most of the chapters unchanged.

The principal exception is Chapter 14, which in the first edition was entitled "Digital Libraries and Web-Based Information Systems." This chapter provided a selected history of the use of Description Logics in web-based information systems, and the developments related to emerging web ontology languages such as OIL and DAML+OIL. Since the writing of this chapter, the new language OWL has been developed and recommended by the World Wide Web consortium as the standard web ontology language for the Semantic Web. In the second edition, Chapter 14, now co-authored by Peter Patel-Schneider, concentrates on OWL, which is reflected by its new title: "OWL: a Description-Logic-Based Ontology Language for the Semantic Web." The chapter still briefly reviews some early efforts that combine Description Logics and the Web, including predecessors of OWL such as OIL and DAML+OIL. But then it goes on to describe OWL in some detail, including the various influences on its design, its relationship with RDFS, its syntax and semantics, and a range of tools and applications.

A minor change was made in Chapter 2. In fact, Proposition 2.9 in the first edition, which tried to give a syntactic criterion for the existence of fixpoint models of cyclic terminologies, turned out to be wrong.[1] In the second edition, it has been replaced by a correct criterion, now given in Proposition 2.10. The new material starts with (the new) Proposition 2.8 and ends with Proposition 2.10.

We are indebted to David Tranah, our editor at Cambridge University Press, for his patience during the preparation of this second edition, but also for the gentle pressure he exerted, without which this second edition would probably not have been completed.

[1] This problem was independently detected by several Ph.D. students, including Yuming Shen, Hongkai Liu, and Boontawee Suntisrivaraporn. Thank you for your careful reading!

Preface

Knowledge Representation is the field of Artificial Intelligence that focuses on the design of formalisms that are both epistemologically and computationally adequate for expressing knowledge about a particular domain. One of the main lines of investigation has been concerned with the principle that knowledge should be represented by characterizing classes of objects and the relationships between them The organization of the classes used to describe a domain of interest is based on a hierarchical structure, which not only provides an effective and compact representation of information, but also allows the relevant reasoning tasks to be performed in a computationally effective way.

The above principle drove the development of the first frame-based systems and semantic networks in the 1970s. However, these systems were in general not formally defined and the associated reasoning tools were strongly dependent on the implementation strategies. A fundamental step towards a logic-based characterization of required formalisms was accomplished through the work on the KL-ONE system, which collected many of the ideas stemming from earlier semantic networks and frame-based systems, and provided a logical basis for interpreting objects, classes (or concepts), and relationships (or links, roles) between them. The first goal of such a logical reconstruction was the precise characterization of the set of constructs used to build class and link expressions. The second goal was to provide reasoning procedures that are sound and complete with respect to the semantics. The article "The tractability of subsumption in Frame-Based Description Languages" by Ron Brachman and Hector Levesque, presented at AAAI 1984, addressing the tradeoff between the expressiveness of KL-ONE–like languages and the computational complexity of reasoning, is usually regarded as the origin of research on *Description Logics*.

Subsequent research came under the label *terminological systems* to emphasize the fact that classes and relationships were used to establish the basic terminology adopted in the modeled domain. Still later, the emphasis was on the set of concept forming constructs admitted in the language, giving rise to the name *concept languages*. Recently, attention has moved closer to the properties of the underlying logical systems, and the term *Description Logics* has become popular.

Research on Description Logics has covered theoretical aspects, implementation of knowledge representation systems (modern frame-based systems) and the use of such systems to realize applications in several areas. This pattern of development is an example of one of the standard research methodologies, as is recognized by the Artificial Intelligence community. The key element has been the very close interaction between theory and practice. On the one hand, there are various implemented systems based on Description Logics, offering a palette of description formalisms with differing expressive power, and which are employed in various application domains (such as natural language processing, configuration of technical systems, databases). On the other hand, the formal and computational properties (like decidability, complexity) of various description formalisms have been studied in detail. These investigations are usually motivated by the use of certain constructors in systems or the need for these constructors in specific applications, and the results of such investigations have strongly influenced the design of new systems.

The Description Logics research community currently consists of at least 100 active researchers. In addition, other communities are now becoming interested in Description Logics, most notably the Databases community and, more recently, the Semantic Web one. After more than a decade of research on Description Logics there is a substantial body of work and well-established technical literature. However, there is no comprehensive presentation of the major achievements in the field, although survey papers have been published and workshop proceedings are available.

Now, since 1989 a workshop dedicated to Description Logics has been held, initially every two years but annually from 1994. At the 1997 workshop a Working Group was formed to develop a proposal for a book that would provide a systematic introduction to Description Logics, covering all aspects of the research in the field, namely: theory, implementation, and applications. Following the spirit that fostered this research, *The Description Logic Handbook* would provide a thorough introduction to Description Logics both for the more theoretically oriented reader interested in the

formal study of Description Logics and for the more practically oriented reader aiming at a principled usage of knowledge representation systems based on Description Logics. Although some refinements have been made to the initial proposal to embody recent developments in the field, the final structure of the *Handbook* reflects the original intentions.

The *Handbook* is organized into three parts plus an initial chapter providing a general introduction to the field.

Part I addresses the theoretical work in Description Logics and includes five chapters. Chapter 2 introduces Description Logics as a formal language for representing knowledge and reasoning about it. Chapter 3 addresses the computational complexity of reasoning in several Description Logics. Chapter 4 explores the relationship with other representation formalisms, within and outside the field of Knowledge Representation. Chapter 5 covers extensions of the basic Description Logics introduced in Chapter 2 by very expressive constructs that require advanced reasoning techniques.

Chapter 6 considers extensions of Description Logics by representation features and non-standard inference problems not available in the basic framework.

Part II is concerned with the implementation of knowledge representation systems based on Description Logics. Chapter 7 describes the features that need to be provided, in addition to the inference engine for a particular Description Logic, to build a knowledge representation system. Chapter 8 reviews implemented knowledge representation systems based on Description Logics that have played or play an important role in the field. Chapter 9 describes the implementation of the reasoning services which form the core of Description Logic knowledge representation systems.

Part III addresses the deployment of Description Logics in the design and implementation of fielded applications. Chapter 10 discusses the issues involved in the development of an ontology for some universe of discourse, which is to become a conceptual model or knowledge base represented and reasoned with using Description Logics. Chapter 11 presents applications of Description Logics in the area of software engineering. Chapter 12 introduces the problem of configuration and the largest and longest lived family of Description Logic-based configurators. Chapter 13 is concerned with the use of Description Logics in various kinds of applications in medical informatics – terminology, intelligent user interfaces, decision support and semantic indexing, language technology, and systems integration. Chapter 14 reviews the applications of Description Logics in web-based information systems, and the more recent developments related to languages for the Semantic Web.

Chapter 15 analyzes the uses of Description Logics for natural language processing to encode syntactic, semantic, and pragmatic elements needed to drive semantic interpretation and natural language generation processes. Chapter 16 surveys the major classes of application of Description Logics and their reasoning facilities to the issues of data management, including the expression of the conceptual domain model/ontology of the data source, the integration of multiple data sources, and the formulation and evaluation of queries.

The syntax and semantics for Description Logics is summarized in an Appendix, which has been used as a reference to unify the notation throughout the book. Finally, an extended, integrated bibliography is provided and, within each chapter, comprehensive guides through the relevant literature are given.

The chapters are written by some of the most prominent researchers in the field, introducing the basic technical material before taking the reader to the current state of the subject. The chapters have been reviewed in a two step process, which involved two or three reviewers for each chapter. We have relied on the work of several external reviewers, selected both within the Description Logic community, and outside the field, to increase the readability for non experts. In addition, each chapter has been read also by authors of other chapters, to improve the overall coherence.

As such, the book is conceived as a unique reference for the subject. Although not intended as a textbook, the *Handbook* can be used as a basis for specialized courses on Description Logics. In addition, some of the chapters can be used as teaching material in Knowledge Representation courses. The *Handbook* is also a comprehensive reference to the subject in more introductory courses in the field of Artificial Intelligence.

We want to acknowledge the contribution and help of several people. First of all, the authors, who have successfully accomplished the hardest task of writing the chapters, carefully addressing the reviewers' comments as well as the issues raised by the effort in making the presentation and notation uniform. Second, we thank the reviewers for their precious work, which led to significant improvements in the final outcome. The external reviewers were:

Premkumar T. Devanbu,
Peter L. Elkin,
Jerome Euzenat,
Erich Grädel,
Michael Gruninger,

Frank van Harmelen,
Jana Koehler,
Diane Litman,
Robert M. MacGregor,
Amedeo Napoli,
Hans-Jürgen Ohlbach,
Marie-Christine Rousset,
Nestor Rychtyckyj,
Renate Schmidt,
James G. Schmolze,
Roberto Sebastiani,
Michael Uschold,
Moshe Y. Vardi,
Grant Weddell,
Robert A. Weida.

A special thank you goes also to Christopher A. Welty who, besides serving as a reviewer, also coordinated the reviewing process for some of the chapters. Third, we express our gratitude to the Description Logics community as a whole (see also the Description Logics homepage at `http://dl.kr.org/`) for the outstanding research achievements and for applying the pressure that enabled us to complete the *Handbook*. Finally, we are indebted to Cambridge University Press, and, in particular, to David Tranah, for giving us the opportunity to put the *Handbook* together and for the excellent support in the editing process.