

Advanced Information and Knowledge Processing

Series Editors

Professor Lakhmi Jain

Lakhmi.jain@unisa.edu.au

Professor Xindong Wu

xwu@cems.uvm.edu

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

Youakim Badr • Richard Chbeir • Ajith Abraham •
Aboul-Ella Hassanien
Editors

Emergent Web Intelligence: Advanced Semantic Technologies

 Springer

Editors

Dr. Youakim Badr
INSA de Lyon
Université de Lyon
Département Informatique
7 avenue Jean Capelle
69621 Villeurbanne Cedex
France
youakim.badr@insa-lyon.fr

Dr. Ajith Abraham
Norwegian University
of Science and Technology
Center for Quantifiable Quality
of Service in Communication Systems
O.S. Bragstads plass 2E
7491 Trondheim
Norway
ajith.abraham@ieee.org

Dr. Richard Chbeir
Université de Bourgogne
LE21-UMR CNRS 5158
Faculté des Sciences Mirande
21078 Dijon Cedex
France
richard.chbeir@u-bourgogne.fr

Dr. Aboul-Ella Hassanien
Kuwait University
College of Business & Administration
Dept. Quantitative Methods
& Information Systems
PO Box 5486
13055 Safat
Kuwait
abo@cba.edu.kw

ISSN 1610-3947

ISBN 978-1-84996-076-2

e-ISBN 978-1-84996-077-9

DOI 10.1007/978-1-84996-077-9

Springer London Dordrecht Heidelberg New York

British Library Cataloguing in Publication Data

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

Library of Congress Control Number: 2010926929

© Springer-Verlag London Limited 2010

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 licenses 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

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The success of the World Wide Web depends on the ability of users to store, process and retrieve digital information regardless of distance boundaries, languages and domains of knowledge. The universality and flexibility of the World Wide Web have also enabled the rapid growth of a variety of new services and applications based on human-machine interaction. The semantics of exchanged information and services should be useful not only for human to human communications, but also in that machines would be able to understand and automatically process web content. Semantics give well-defined meaning to web content and enable computers and people to work in cooperation. Today, the crucial challenge becomes the development of languages to express information in a machine processable format. Now more than ever, new advanced techniques and intelligent approaches are required to transform the Web into a universal reasoning and computing machine. Web intelligence attempts to deal with this challenge by exploiting information technologies and artificial intelligence approaches to design the next generation of web-empowered systems and services.

1 Introduction

The semantic Web plays a crucial role in the development of information technologies and services on the World Wide Web. It takes on new challenges in which the meaning of information enable computers to understand the Web content and imitate human intelligence in performing more of the tedious tasks involved in finding, sharing, and combining information on the web. Until now computers have not been able to fully accomplish these tasks without human intervention since web pages are designed to be understood by people, not machines. Tim Berners-Lee originally expressed his vision about the semantic Web. In his word, he wrote: “*[I] have a dream for the Web [in which computers] become capable of analyzing all the data on the Web—the content, links, and transactions between people and computers. A ‘Semantic Web’, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by*

machines talking to machines. The ‘intelligent agents’ people have touted for ages will finally materialize.” Significant research efforts attempt to support this vision and propose a set of design principles, formal specifications, and languages specifically designed for a huge data space. Some of these include Resource Description Framework (RDF), Web Ontology languages (OWL), Semantic-based Query Languages and Logic Models to reason on the structure of knowledge embedded in Web-accessible databases and documents. Despite these initiatives, some of the challenges remain in a bottleneck due to the requirement for automating reasoning systems to deal with inconsistency, vastness, uncertainty, vagueness, and deceit in order to deliver on the promise of the Semantic Web. The discipline of Soft Computing has an evolving collection of methodologies, which aims to exploit tolerance for imprecision, uncertainty, and partial truth to study very complex phenomena: those for which more conventional methods have not yielded low cost and complete solutions. Today, Soft Computing provides an attractive opportunity for developing Web intelligence to represent the ambiguity in human thinking with real life uncertainty, reason on vagueness in ontologies, and makes possible the transition between the Web and its semantic successor. In this context, Semantic Web will enable the emergence of digital ecosystems of software and services delivered by the Internet. It will also extend the Internet with capabilities to reason on its resources and their relationships in order to develop the knowledge-based economy in the 21st century.

Why This Book is Interesting? Industrial and technological demands to develop semantic-driven applications in business, commerce, marketing, finance and education have persuaded academia and scholarly communities across the world to include Web intelligence disciplines in their computer science curriculum. Moreover, many research centers are extensively working on this research field which demonstrates an important interest in building the next Semantic World Wide Web. The number of journals in this area has increased and the number of related conferences organized in the last ten years is overwhelming. However, there are relatively few books about web intelligence and semantic web taking into consideration knowledge discovery, semantic network, ontologies and artificial intelligence techniques such as neural network, fuzzy logic and mining algorithms as a new paradigm. Consequently, the need for a new book meets the increasing demands of academia and research communities and provides advanced techniques and methodologies to help undergraduates, graduates and researchers. The primary target audience for the book includes researchers, scholars, postgraduate students and developers who are interested in exploring various areas and disciplines about how semantic techniques and technologies can bridge the gap between users and applications on the web.

2 Book Organization

This book aims to gather the latest advances and innovative solutions in web intelligence and reporting how future services and web-based applications can gain competitive opportunities by applying different emergent semantic techniques to

real-world problems. The main topics of this edited volume cover various advanced semantic technologies, use tools and software for collaboration and simulations of web intelligence, and finally provide case studies and applications in the field of semantic Web. It presents some of the latest advances of various topics in web intelligence and illustrates how organizations can gain competitive advantages by applying the different emergent techniques in the real-world scenarios. The book contains seventeen self-contained chapters which provide optimal reading flexibility. They are organized into four parts as follows:

- Web, Semantic and Intelligence
- Collaboration, Semantic and Intelligence
- Knowledge, Text, Semantic and Intelligence
- Applications and Case Studies

Part I deals with *Web, Semantic and Intelligence* and consists of four chapters.

In Chap. 1, “*The Dilated Triple*”, Rodriguez et al. present a simple model in which the meaning of a statement is defined in terms of other statements, much like the words of a dictionary are defined in terms of each other. This model serves to strengthen the association between concepts and the contexts in which they are to be understood. It provides a particularly simple means of contextualizing an RDF triple by associating it with related statements in the same graph. This approach, in combination with a notion of graph similarity, is sufficient to select only those statements from an RDF graph, which are subjectively most relevant to the context of the requesting process.

In Chap. 2, “*Semantic Web Technologies and Artificial Neural Networks for Intelligent Web Knowledge Source Discovery*”, Caliusco and Stegmayer present some basic concepts and foundations regarding the new Semantic Web and how it is populated with ontologies and why ontology-matching techniques are needed. The idea of software agents that travel the web carrying query request from users has also been addressed in this Chapter. The web knowledge source discovery task is explained in detail and some motivating scenarios are introduced. To help users avoid irrelevant web search results and wrong decision making, efficient techniques and approaches for developing web intelligence with capabilities for discovering distributed knowledge source are presented.

Wang et al. in Chap. 3, titled “*Computing Similarity of Semantic Web Services in Semantic Nets with Multiple Concept Relations*” propose a novel approach based on application ontologies to improve the selection of semantic web services. After building application ontology by merging semantic service ontologies, the authors represent this application ontology as a fuzzy-weighted semantic net with multiple ontological concept relations, and calculate formal/compound concept similarity on it. The ontological concept similarity is finally used to calculate similarity of semantic services.

Chapter 4, “*General-Purpose Computing on a Semantic Network Substrate*” by Rodriguez presents a model of general-purpose computing on a semantic network substrate. The concepts presented are applicable to any semantic network representation. In the proposed model, the application programming interface, the run-time

program, and the state of the computing virtual machine are all represented in the Resource Description Framework (RDF). The implementations of the concepts presented provide a computing paradigm that leverages the distributed and standardized representational-layer of the Semantic Web.

Part II consists of four chapters and deals with *Collaboration, Semantic and Intelligence*.

In Chap. 5, “*Agent Technology Meets the Semantic Web: Interoperability and Communication Issues*”, Karanastasi and Matsatsini review the recent research on agent technologies and how this technology can serve the scopes of the Semantic Web project including web Agent and the Characteristics of Multi-Agent Systems, Agent Communication Languages, Knowledge and Query Manipulation Language, FIPA Agent Communication Language, and Ontologies.

In Chap. 6, “*Mining of Semantic Image Content Using Collective Web Intelligence*”, Leung et al. describe an indexing method, whereby the aggregate intelligence of different Web users is continuously transferred to the Web. Such intelligence is codified, reinforced, distilled and shared among users so as to enable the systematic mining and discovery of semantic image contents. The described method allows the collaborative creation of image indexes, which is able to instill and propagate deep knowledge and collective wisdom into the Web concerning the advanced semantic characteristics of Web images.

In Chap. 7, “*Suited Support for Distributed Web Intelligence Cooperative Work*”, Decouchant et al. present the PINAS platform, which provides means for supporting cooperative work on the Web. Using cooperative applications that are built employing the services of this infrastructure, several users can access and modify replicated shared entities in a consistent and controlled way. PINAS provides suited features, such as: user identification, multi-site user definition, user and entity naming, shared entity fragmentation and replication, storage, consistency, and automatic distributed updating. The authors propose seamless extensions to standard Web services that can be fully integrated within the Web environment.

In Chap. 8, “*Web services and Software Agents for Tailorable Groupware Design*”, Cheaib et al. present a new groupware architecture model called UD3 that explicitly introduces the notion of tailor ability in designing collaborative applications. This model is based on the integration of web services and software agent technologies, thus using protocols of each while reinforcing their individual strengths in the context of tailorable groupware design. Web services are dynamically invoked by software agents in order to bring new behaviors, and hence, enhancing the collaboration process by dynamically adapting the services offered in the system to the users’ preferences and not the other way around.

Part III consists of five chapters and focus on *Knowledge, Text, Semantic and Intelligence*.

In Chap. 9, “*Toward Distributed Knowledge Discovery on Grid Systems*”, Khac et al. present a distributed data mining (DDM) system based on Grid environments to execute new distributed data mining techniques on very large and distributed heterogeneous datasets. The architecture and motivation for the design are presented. The authors developed prototypes for each layer of the system to evaluate the system features, test each layer as well as whole framework and building simulation

and DDM test suites. Knowledge map layer, key layer of this system, is integrated in this framework.

In Chap. 10, “*Metamodel of Ontology Learning from Text*”, Wisniewski presents the metamodel of the ontology learning from text. The approach is based on the survey of the existing methods, while evaluation is provided in the form of a reference implementation of the introduced metamodel. The author has applied a qualitative evaluation by implementing some of the current state-of-the-art methods and illustrates how they can be described with a metamodel notation.

Ambiguity is a challenge faced by systems that handle natural language. To assuage the issue of linguistic ambiguities found in text classification, Chap. 11, “*An Analysis of Constructed Categories for Textual Classification Using Fuzzy Similarity and Agglomerative Hierarchical Methods*” by Guelpeli et al. proposes a text categorizer using the methodology of Fuzzy Similarity. The clustering algorithms Stars and Cliques are adopted in the Agglomerative Hierarchical method and they authors identify the groups of texts by specifying some type of relationship rule to create categories based on the similarity analysis of the textual terms.

In Chap. 12, “*Emergent XML Mining: Discovering an Efficient Mapping from XML Instances to Relational Schemas*”, Ishikawa proposes an adaptable approach to discovery of database schemas for well-formed XML data such as EDI, news, and digital libraries, which we interchange, filter, or download for future retrieval and analysis. The generated schemas usually consist of more than one table. Author’s approach controls the number of tables to be divided by use of statistics of XML so that the total cost of processing queries is reduced. To achieve this, three functions namely NULL expectation, Large Leaf Fields, and Large Child Fields are introduced for controlling the number of tables to be divided. The author also describes the concept of short paths contained by generated database schemas and their effects on the performance of query processing.

In Chap. 13, “*XML Based Information Systems and Formal Semantics of Programming Languages*”, Despeyroux illustrates how techniques used to define the formal semantics of programming languages can be used in the context of the Web. The author also explores how techniques used in this context can be used to enforce the quality of information systems.

Part IV consists of four chapters and deals with *Applications and Case Studies*.

In Chap. 14, “*Modeling and Testing of Web Based Systems*”, Cavalli et al. present two methodologies to attain automatic test cases generation: The first uses extended finite state machines to model Web services composition described in BPEL, while the other uses UML to model Web applications. Together with the formal models of the web systems, this chapter presents methods for conformance and non-regression test generation.

As web applications are becoming ever larger, more complex and thus more demanding for their users, there is a growing need for customer support. Very often, it is provided by support centers via phone. However, the media break between browser and phone hampers the common understanding of user and consultant. As a result, support becomes ineffective and expensive, and users get frustrated. Screen sharing solutions are one possible solution for this problem, but they have major

disadvantages like high bandwidth requirements, slow performance and, most importantly, the need for a client-side installation. These drawbacks are addressed by VCS, a concept and system for instant co-browsing, that runs directly within the users browser. It equally allows all participants of a support session to see and navigate the same web page on their screens, being aware of what the other person is currently doing on the page. People can directly interact with each other, jointly complete tasks and solve. The event-based nature of the synchronization approach to be presented further facilitates adaptation, so that users with heterogeneous end devices may collaborate. Niederhausen et al., in Chap. 15, “*Web-Based Support by Thin-Client Co-Browsing*”, present VCS and also discuss the special challenges that this approach entails.

In Chap. 16, “*NetPay Micro-Payment Protocols for Three Networks*”, Xiaoling and Grundy describe the NetPay micro-payment protocol that is actually extended from its original pay-per-click for web content to peer-to-peer networks and mobile device networks. The authors outline the key motivation for NetPay, the basic micro-payment protocol using e-coins and e-wallets, and the three variants of the protocol for different domains.

Chapter 17, “*Enforcing Honesty in Fair Exchange Protocols*” by Alaraj and Munro surveys the field of Fair Exchange Protocols and then presents a special type of protocol between a customer (C) and a merchant (M) that enforces one of them to be honest. It makes minimal use of a Trusted Third Party (TTP). The protocol has the features that it: (1) only comprises three messages to be exchanged between C and M; (2) guarantees strong fairness for both C and M; (3) allows both parties to be sure of the item that they will receive from the other party; and (4) resolves disputes automatically online.

3 Acknowledgments

We would like to thank the authors who provided excellent chapters and timely revisions. We are also grateful for their trust in us and patience during the review process. We would like to express our sincere thanks to the reviewers for their tremendous effort and challenging task of choosing high quality chapters, and their valuable criticism that greatly improved the quality of final chapter versions. The editors also would like to thank, Professor Lakhmi C. Jain, the editor-in-chief of the Advanced Information and Knowledge Processing (AI and KP) series of Springer for editorial assistance and excellent cooperative collaboration to produce this important scientific work. We hope this volume motivates its readers to take the next steps beyond building models to implementing, evaluating, comparing, and extending proposed approaches and applications. We finally hope that readers will share our excitement to present this volume on **Web Intelligence: Advanced Semantic Technologies** and find it useful.

4 About the Editors

Youakim Badr received his Ph.D. in Information Systems from the French National Institute for Applied Sciences in Lyon (INSA of Lyon). In 2004, he joined the faculty of the INSA of Lyon as Assistant Professor of Computer Science. Dr. Badr has worked extensively in the field of coupling XML documents and Object-Relational Databases. Through his research he has acquired skills in fields such as Interoperability, Modeling, System Architectures and Networking, and their application to various domains such as Business Processes, Supply Chains, Productions Systems and Virtual Enterprises. His current academic research interests include systems in both the service sector and ICT. In particular, he studies the ecosystem of services and the multidisciplinary modeling approach to design services through the integration of ICT, strategy and processes. He leads the Service-Oriented Enterprise research team which combines industrial and computer engineering approaches. Dr. Badr is vigorously involved in a series of international conferences. He served as General Co-Chair of ICDIM'07, CSTST'08, Programme Chair of INCOS'09, Track chair of IEEE DEST'10, AINA'10, ICETET'09, ICITST'08 and International Program Member of IAS'08, SITIS'07, JFO'07, WCNC'07 and ECWS'06. He is a professional member of ACM, IEEE Services Computing Community, MIL-RLabs/France coordinator, IEEE-SMC Technical Committee on Soft Computing, Digital Ecosystems Community, a member of OW2 and the Service Sciences working group of the Networked European Software and Services Initiative (NESSI).

Richard Chbeir received his Ph.D. in Computer Science from the INSA of Lyon, France in 2001. He is member of IEEE and ACM since 1999. He is currently an Associate Professor in the Computer Science Department of the Bourgogne University, Dijon, France. His research interests are in the areas of distributed multimedia database management, XML similarity and rewriting, spatio-temporal applications, indexing methods, and multimedia access control models. Dr. Chbeir has published (more than 40 peer-reviewed publications) in international journals and books (IEEE Transactions on SMC, Information Systems, Journal on Data Semantics, Journal of Systems Architecture, etc.), conferences (ER, WISE, SOFSEM, EDBT, ACM SAC, Visual, IEEE CIT, FLAIRS, PDCS, etc.), and has served on the program committees of several international conferences (SOFSEM, AINA, IEEE SITIS, ACM SAC, IEEE ISSPIT, EuroPar, SBBB, etc.). He has been organizing many international conferences and workshops (ACM SAC, ICDIM, CSTST, SITIS, etc.). He is currently the Vice-Chair of the ACM SIGAPP and the Chair of its French Chapter.

Ajith Abraham received his Ph.D. degree in Computer Science from Monash University, Melbourne, Australia. His research and development experience includes over 18 years in the Industry and Academia. He works in a multidisciplinary environment involving machine intelligence, network security, sensor networks, e-commerce, Web intelligence, Web services, computational grids, data mining, and applications to various real-world problems. He has given more than 30 plenary lectures and conference tutorials in these areas. He authored or coauthored more than 500 publications. He works with the Norwegian University of Science and Technology, Norway and also coordinate the activities of the Machine Intelligence Research

Labs (MIR Labs), which has representation in 47 countries. He is the Co-Chair of the IEEE Systems Man and Cybernetics Society Technical Committee on Soft Computing. He is the founder of several conference series, which are now sponsored by IEEE and also serves the editorial board of over 30 international Journals.

Aboul Ella Hassanien received his B.Sc. with honours in 1986 and M.Sc. degree in 1993, both from Ain Shams University, Faculty of Science, Ain Sham University, Egypt. On September 1998, he received his doctoral degree from the Department of Computer Science, Graduate School of Science and Engineering, Tokyo Institute of Technology, Japan. He is currently a Professor at Cairo University, Faculty of Computer and Information. He has authored/coauthored over 120 research publications in peer-reviewed reputed journals and conference proceedings. He serves on the editorial board and reviewer of number of journals and on the program committee of several international conferences and he has editing/written more than 18 books. He has received the excellence younger researcher award from Kuwait University.

5 List of Reviewers

Ahmed El Oualkadi (Universit Catholique de Louvain, Belgium)
 Akira Asano (Hiroshima University, Japan)
 Alfredo Cuzzocrea (University of Calabria, Italy)
 Andries Engelbrecht (University of Pretoria, South Africa)
 Bernard Grabot (LGP-ENIT, France)
 Carlos Alberto Reyes-Garcia (Instituto Nacional de Astrofisica Optica Y Electronica, Mexico)
 Chi Shen (University of Kentucky, USA)
 Chrisa Tsinaraki (Technical University of Crete, Greece)
 Christine Verdier (University of Grenoble, France)
 Danielle Boulanger (University of Jean Moulin-MODEME, France)
 Deborah Dahl (Conversational Technologies, USA)
 Elizabeth Goldbarg (Federal University of Rio Grande do Norte, Brazil)
 Estevam Hruschka Jr. (Federal University of Sao Carlos, Brazil)
 Etienne Kerre (University of Gent, Belgium)
 Gabriel Luque (University of Malaga, Spain)
 Georgios Ch. Sirakoulis (Democritus University of Thrace, Greece)
 Hiranmay Ghosh (Tata Consultancy Services, India)
 Hiroshi Ishikawa (Shizuoka University, Japan)
 Ignacio Ponzoni (Universidad Nacional del Sur, Argentina)
 James Lu (Emory University, USA)
 Jaroslaw Kozlak (University of Science and Technology Krakow, Poland)
 Javier J. Sanchez-Medina (University of Las Palmas de Gran Canaria, Spain)
 Kazushi Ohya (Tsurumi University, Japan)
 Kubilay Ecerkale (Hava Harp Okulu, Turkey)
 Mario Koeppen (Kyushu Institute of Technology, Japan)
 Mario Ventresca (University of Waterloo, Canada)
 Maytham Safar (Kuwait University, Kuwait)
 Mei-Ling Shyu (University of Miami, USA)

Michael Blumenstein (Griffith University, Gold Coast, Australia)
 Monica Chis (University of Cluj-Napoca, Romania)
 Nadine Cullot (Bourgogne University, France)
 Oscar Corcho (University of Manchester, UK)
 Paolo Merialdo (Universita' degli Studi Roma Tre, Italy)
 Patrick Siarry (Universit Paris 12, LiSSi, France)
 Patrizia Grifoni (National Research Council of Italy, Italy)
 Raquel Barco (Universidad de Malaga, Spain)
 Sadok Ben Yahia (University of Tunis, Tunisia)
 Saravanan Muthaiyah (George Mason University, USA)
 Sebastin Lozano (University of Seville, Spain)
 Selma Ayse Ozel (Cukurova University, Turkey)
 Siti Mariyam Shamsuddin (University Technology of Malaysia, Malaysia)
 Thanasis Daradoumis (Open University of Catalonia, Spain)
 Thierry Badard (Universite Laval, Canada)
 Thomas Hanne (Applied Sciences Northwestern Switzerland University, Switzerland)
 Tianrui Li (Southwest Jiaotong University, China)
 Tomasz Smolinski (Emory University, USA)
 Urszula Markowska-Kaczmar (Wroclaw University of Technology, Poland)
 Ying Ding (Indiana University, USA)
 Yinghua Ma (Shanghai JiaoTong University, China)
 Zhigang Zeng (Wuhan University of Technology, China)
 Zhi-Hong Deng (Peking University, China)

Villeurbanne Cedex, France
 Dijon Cedex, France
 Trondheim, Norway
 Safat, Kuwait

Youakim Badr
 Richard Chbeir
 Ajith Abraham
 Aboul-Ella Hassanien

Contents

Part I Web, Semantic and Intelligence

1	The Dilated Triple	3
	Marko A. Rodriguez, Alberto Pepe, and Joshua Shinavier	
1.1	Introduction	3
1.2	The Dilated Triple Model	6
1.3	Contextualizing a Relationship	7
1.4	Comparing Contexts	10
1.5	Conclusion	13
	References	13
2	Semantic Web Technologies and Artificial Neural Networks for Intelligent Web Knowledge Source Discovery	17
	M.L. Caliusco and G. Stegmayer	
2.1	Introduction	17
2.2	Foundations	19
2.2.1	Ontologies and Ontology-Matching	19
2.2.2	Software Agents	20
2.2.3	Artificial Neural Networks	21
2.3	ANNs and the Semantic Web: Literature Review	22
2.3.1	Searching and Query Answering on the Semantic Web	22
2.3.2	Ontology-Matching and ANN Models	23
2.4	Web Knowledge Source Discovery	24
2.4.1	A Knowledge Source Discovery Agent	26
2.5	The ANN-Based Ontology-Matching Model Inside the KSD Agent	29
2.5.1	ANN-Based Ontology-Matching Model: Training Phase	32
2.5.2	ANN-Based Ontology-Matching Model: Matching Phase	33
2.6	Conclusions	35
	References	35

3 Computing Similarity of Semantic Web Services in Semantic Nets with Multiple Concept Relations 37
 Xia Wang, Yi Zhao, and Wolfgang A. Halang

3.1 Introduction 37

3.2 Ontologies in the SWS Context 39

 3.2.1 Ontology Levels 39

 3.2.2 Ontologies in WSMO 40

3.3 An Ontology-Based Selection Framework of SWSs 41

 3.3.1 Building Application Ontologies 42

3.4 Semantic Net with Multiple Concept Relations 43

 3.4.1 Ontological Concept Similarity Algorithm (OCSA) 44

 3.4.2 Algorithm and Implementation 46

 3.4.3 Service Similarity 47

3.5 Experimental Evaluation and Discussion 49

3.6 Related Work 52

3.7 Conclusion 54

 References 54

4 General-Purpose Computing on a Semantic Network Substrate 57
 Marko A. Rodriguez

4.1 Introduction 57

 4.1.1 General-Purpose Computing and the Virtual Machine 58

 4.1.2 The Semantic Web and RDF 59

 4.1.3 Object-Oriented Programming and OWL 62

 4.1.4 Contributions 64

4.2 A High-Level Perspective 65

 4.2.1 The Ontological Level 67

 4.2.2 The Instance Level 68

 4.2.3 The Machine Level 68

4.3 The Neno Language 69

 4.3.1 The Universally Unique Identifier Address Space 70

 4.3.2 Class Declarations in Neno Source Code 71

 4.3.3 Starting a Program in Neno 84

 4.3.4 Typical Use Case 85

4.4 The Fhat Virtual Machine Architecture 86

 4.4.1 Fhat Components 86

 4.4.2 Migrating Fhat Across Different Host CPUs 91

 4.4.3 Fhat Reflection 93

 4.4.4 r-Fhat 93

4.5 The Fhat Instruction Set 94

 4.5.1 The Method 94

 4.5.2 A Method Instance 97

4.6 Conclusion 98

 References 100

Part II Collaboration, Semantic and Intelligence

5 Agent Technology Meets the Semantic Web: Interoperability and Communication Issues 105
 Anastasia Karanastasi and Nikolaos Matsatsinis

5.1 Introduction 105

5.2 Main Differences Between Web Services and Agents 107

5.3 Web Agents 107

5.4 Characteristics of Multi-Agent Systems 109

5.5 Agent Communication Languages 110

5.5.1 Knowledge and Query Manipulation Language (KQML) . . 111

5.5.2 FIPA ACL 111

5.5.3 Comparing KQML and FIPA ACL 112

5.6 Ontologies 113

5.6.1 OWL 115

5.7 Specifications for Open Multi-Agent Systems 116

5.8 Interoperability Issues 117

5.9 Conclusions 118

References 119

6 Mining of Semantic Image Content Using Collective Web Intelligence 121
 C.H.C. Leung, J. Liu, A. Milani, and W.S. Chan

6.1 Cumulative Web Intelligence for Image Mining 121

6.2 Attaining Intelligence Through Adaptive Indexing 122

6.3 Reinforcing Collective Judgment 124

6.3.1 Incrementing the Score 124

6.3.2 Decrementing the Score 125

6.4 Expanding Intelligence Through Evolution 126

6.4.1 Augmentation of an Existing Index Term 127

6.4.2 Addition of New Index Terms 127

6.4.3 Maximal Indexing 129

6.4.4 Probability of Image Recovery 130

6.5 Collective Wisdom Indexing for Advanced Image Semantics . . . 131

6.6 Index Convergence 132

6.6.1 Case A 132

6.6.2 Case B 132

6.7 Conclusions 133

References 134

7 Suited Support for Distributed Web Intelligence Cooperative Work . 137
 Dominique Decouchant, Sonia Mendoza, and José Rodríguez

7.1 Introduction to Web Cooperative Work 138

7.1.1 Some Efforts at Protocol Level 139

7.1.2 Requirements for a Suited Web Cooperative Infrastructure . 141

7.2 What Architecture for What Groupware? 142

- 7.2.1 Classification Model 142
- 7.2.2 Centralized Architecture 145
- 7.2.3 Fully Distributed Architecture 148
- 7.2.4 Hybrid Architecture 150
- 7.2.5 Coordination Architectures 153
- 7.3 The PIÑAS Web Infrastructure 161
 - 7.3.1 PIÑAS Naming Space 163
 - 7.3.2 Web Document Replication 166
 - 7.3.3 Multi-Site User Work Organization 168
 - 7.3.4 Establishing the Cooperation 172
 - 7.3.5 Disconnected and Nomadic Cooperative Work 176
- 7.4 Conclusions 181
- References 182

8 Web Services and Software Agents for Tailorable Groupware Design 185

Nader Cheaib, Samir Otmame, and Malik Mallem

- 8.1 Introduction 185
- 8.2 Motivating Scenario 187
 - 8.2.1 Tailorability and the Need of a New Architecture 187
- 8.3 Tailorability Approaches 188
 - 8.3.1 Activity Theory 189
 - 8.3.2 Component-Based Architecture 189
 - 8.3.3 Service-Oriented Architecture (SOA) 191
- 8.4 Background 192
 - 8.4.1 Ellis’s 3C Model 192
 - 8.4.2 Web Services and the World Wide Web 192
 - 8.4.3 Software Agents 193
 - 8.4.4 JADE Platform 193
 - 8.4.5 Related Work—Web Services and Agents’ Integration 194
 - 8.4.6 Purpose of Integration 195
- 8.5 The *UD*³ Theoretical Model 197
 - 8.5.1 Description of the Functional Core (FC) 198
 - 8.5.2 FC Decomposition 199
 - 8.5.3 SOA Environment 199
 - 8.5.4 JADE Agents’ Environment 200
 - 8.5.5 Universal Directory for Description and Discovery 201
 - 8.5.6 Dynamic Discovery and the Semantic Web 201
 - 8.5.7 FC Implementation 202
- 8.6 Case Study—Oce@nyd 204
 - 8.6.1 Physical Layer 205
 - 8.6.2 Shared FC 206
 - 8.6.3 Invocation of an External Web Service by an Agent 206
 - 8.6.4 Properties and Discussion 207
- 8.7 Conclusion 208
- References 208

Part III Knowledge, Text, Semantic and Intelligence

9 Toward Distributed Knowledge Discovery on Grid Systems 213
 Nhien An Le Khac, Lamine M. Aouad, and M-Tahar Kechadi

9.1 Introduction 213

9.2 *DDM* Systems 215

 9.2.1 *DDM* Issues 215

 9.2.2 Toward an Efficient *DDM* System on Grid Platform 217

9.3 Architecture of a New *DDM* System 218

 9.3.1 Core Layer 218

 9.3.2 Data Grid Layer 219

9.4 Distributed Algorithms for Mining Large Datasets 220

 9.4.1 Variance-Based Clustering 220

 9.4.2 Distributed Frequency Itemsets Generation 223

9.5 Knowledge Map 227

 9.5.1 From Knowledge Representation to Knowledge Maps 227

 9.5.2 Knowledge Map Layer Structure 229

 9.5.3 Evaluation 232

9.6 Exploitation 234

 9.6.1 Interface 234

 9.6.2 An Example of Exploiting the *DDM* System via
 Knowledge Map 236

9.7 Related Works of *DDM* Frameworks on Grid Platforms 238

 9.7.1 Knowledge Grid 239

 9.7.2 GridMiner 239

 9.7.3 Discovery Net 240

9.8 Conclusion 240

 References 241

10 Metamodel of Ontology Learning from Text 245
 Marek Wisniewski

10.1 Motivation 245

10.2 A Survey of Ontology Learning Methods 249

 10.2.1 Terms 250

 10.2.2 Synonyms 252

 10.2.3 Concepts 253

 10.2.4 Taxonomic Relations 254

 10.2.5 Non-taxonomic Relations 254

 10.2.6 Rules 255

 10.2.7 Ontology Learning Maps 255

 10.2.8 Conclusions 259

10.3 Metamodel 260

 10.3.1 Model Duality 261

 10.3.2 Extraction Models 262

 10.3.3 Reference Architecture 265

- 10.4 Evaluation 267
 - 10.4.1 Corpora 268
 - 10.4.2 Term Extraction Methods 269
 - 10.4.3 Synonym and Concept Extraction Methods 270
 - 10.4.4 Relation Extraction Methods 270
- 10.5 Conclusion 270
 - References 271
- 11 An Analysis of Constructed Categories for Textual Classification
Using Fuzzy Similarity and Agglomerative Hierarchical Methods . . . 277**
 Marcus V.C. Guelpele, Ana Cristina Bicharra Garcia, and Flavia Cristina Bernardini
 - 11.1 Introduction 277
 - 11.2 Fuzzy Similarity 280
 - 11.3 Agglomerative Hierarchical Methods 282
 - 11.3.1 Stars Algorithm 285
 - 11.3.2 Cliques Algorithm 286
 - 11.4 An Approach to Text Categorization—A Proposal 286
 - 11.5 Experiments 289
 - 11.5.1 Hypothesis 290
 - 11.5.2 Decision Rule for the *F*-Test 291
 - 11.5.3 Testing the Null Hypothesis 291
 - 11.5.4 Qualitative Analysis of the Constructed Categories 294
 - 11.6 Conclusion 301
 - References 304
- 12 Emergent XML Mining: Discovering an Efficient Mapping
from XML Instances to Relational Schemas 307**
 Hiroshi Ishikawa
 - 12.1 Introduction 307
 - 12.1.1 Web Mining 308
 - 12.1.2 Emergent XML Mining 309
 - 12.1.3 Previous Works on Emergent XML Mining 310
 - 12.2 Discovery of XML Structures 311
 - 12.3 Concepts and Definitions 313
 - 12.3.1 Concepts 313
 - 12.3.2 Definitions 314
 - 12.4 Our Approach 316
 - 12.4.1 Discovery of Complex Data 316
 - 12.4.2 Division of Tables 317
 - 12.4.3 Resultant Database Schema and Short Path 321
 - 12.5 Evaluation 321
 - 12.5.1 Settings 322
 - 12.5.2 Queries 323
 - 12.5.3 Query Translation 324
 - 12.5.4 Results 325

12.6 Schema Change 327

12.7 Conclusions 327

 References 328

13 XML Based Information Systems and Formal Semantics

of Programming Languages 331

Thierry Despeyroux

13.1 Introduction 331

13.2 Syntax and Semantics of Programming Languages 333

 13.2.1 Syntax Versus Semantics 333

 13.2.2 Formal Semantics 335

13.3 XML Files Viewed as Programs 336

 13.3.1 Similarities and Differences 336

 13.3.2 Semantic Constraints in XML Documents 337

13.4 A Specification Language to Define the Semantics of XML Documents and Web Sites 339

 13.4.1 Formalizing Web Sites 339

 13.4.2 Definition of SeXML 340

 13.4.3 Dynamic Semantics 342

13.5 Ontologies as Types 343

13.6 Applications 346

 13.6.1 Verifying a Web Site 346

 13.6.2 Verifying a Document and Inferring New Data 347

 13.6.3 Implementation Notes 349

13.7 Conclusion 349

 References 350

Part IV Applications and Case Studies

14 Modeling and Testing of Web-Based Systems 355

Ana Cavalli, Mounir Lallali, Stephane Maag, Gerardo Morales, and Fatiha Zaidi

14.1 Introduction 355

14.2 Preliminaries 358

 14.2.1 Definitions 358

 14.2.2 Testing Techniques 359

 14.2.3 Related Works 362

14.3 Web-Based System Modeling 364

 14.3.1 Modeling of Web Services 365

 14.3.2 Modeling of Web Applications 371

14.4 Web-Based System Testing 373

 14.4.1 Web Services Testing 373

 14.4.2 Test Generation for Web Applications 378

14.5 Case Study: DotLRN 380

 14.5.1 The DotLRN Framework 380

 14.5.2 Test Generation from the UML Model 381

- 14.5.3 Non-regression Testing for Web Applications 385
- 14.5.4 Alternative Method for the Test Generation of Web Applications 385
- 14.5.5 Migrating Web Applications Functionalities into Web Services 388
- 14.6 Conclusion 389
- Appendix 390
- References 392
- 15 Web-Based Support by Thin-Client Co-browsing 395**
Matthias Niederhausen, Stefan Pietschmann, Tobias Ruch, and Klaus Meißner
- 15.1 Introduction 395
- 15.2 State of the Art in Co-browsing 396
 - 15.2.1 Related Co-browsing Approaches 397
- 15.3 CoCAB: Collaborative Context-Aware Browsing 399
 - 15.3.1 Basic Concepts 399
 - 15.3.2 Architectural Overview 399
 - 15.3.3 Event-Based Synchronization 401
 - 15.3.4 Co-browsing Awareness 402
 - 15.3.5 Context-Awareness 403
- 15.4 Fields of Application 411
 - 15.4.1 Consulting 411
 - 15.4.2 Co-shopping 412
 - 15.4.3 Online Presentation 413
 - 15.4.4 Communities 414
- 15.5 Implementation 414
 - 15.5.1 Co-browsing Server 414
 - 15.5.2 Content Processing and Delivery 417
 - 15.5.3 Context Modeling and Adaptation 418
 - 15.5.4 User Interface and Interaction Customization 420
 - 15.5.5 Client Component Delivery 420
 - 15.5.6 Synchronization Techniques 424
- 15.6 Discussion 424
 - 15.6.1 Security Aspects 424
 - 15.6.2 Adaptation Level 425
 - 15.6.3 Enhanced Web Content 426
- 15.7 Conclusion and Future Work 426
- References 427
- 16 NetPay Micro-Payment Protocols for Three Networks 429**
Xiaoling Dai and John Grundy
- 16.1 Introduction 429
- 16.2 Motivation 430
- 16.3 NetPay Micro-Payment Protocol for E-commerce in Client–Server Networks 433

- 16.3.1 NetPay Transactions 434
- 16.3.2 NetPay Architectures 437
- 16.4 P2P-NetPay for Content Sharing in Peer-to-Peer Networks 437
 - 16.4.1 P2P-NetPay Transactions 439
 - 16.4.2 P2P-NetPay Architecture 439
- 16.5 Mobile-NetPay for Mobile Commerce in Wireless Networks 440
 - 16.5.1 Mobile-NetPay Transactions 442
 - 16.5.2 Mobile-NetPay Architectures 443
- 16.6 Discussion 443
 - 16.6.1 P2P Micro-Payment Systems Comparison 444
 - 16.6.2 Mobile Micro-Payment Systems Comparison 446
- 16.7 Summary 448
 - References 448
- 17 Enforcing Honesty in Fair Exchange Protocols 451**
 - Abdullah M. Alaraj and Malcolm Munro
 - 17.1 Introduction 451
 - 17.2 Review of Literature 452
 - 17.2.1 Protocols that Do Not Involve a TTP 455
 - 17.2.2 Protocols that Involve a TTP 456
 - 17.2.3 Examples of Fair Exchange Protocols 459
 - 17.2.4 Discussions 463
 - 17.3 Enforcing Merchant Honesty (EMH) Protocol 464
 - 17.3.1 Notations 464
 - 17.3.2 Protocol Description 465
 - 17.3.3 Pre-exchange Phase 465
 - 17.3.4 The Exchange Phase 466
 - 17.3.5 After Exchange (Dispute Resolution) 469
 - 17.4 Disputes Analysis 471
 - 17.5 All Possible Cases (Scenarios) of Executing EMH Protocol 474
 - 17.6 Comparisons 475
 - 17.7 Conclusion 477
 - References 478
- Index 481**

Contributors

Abdullah M. Alaraj Department of Information Technology, College of Computer, Qassim University, Qassim, Saudi Arabia, arj@qu.edu.sa

Lamine M. Aouad School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland, lamine.aouad@ucd.ie

Flavia Cristina Bernardini Departamento de Ciência e Tecnologia—RCT, Pólo Universitário de Rio das Ostras—PURO, Universidade Federal Fluminense—UFF, Rua Recife, s/n, Jardim Bela Vista, Rio das Ostras, RJ CEP 28890-000, Brazil, fbernardini@vm.uff.br

Ana Cristina Bicharra Garcia Departamento de Ciência da Computação, Instituto de Computação—IC, Universidade Federal Fluminense—UFF, Rua Passo da Pátria 156, Bloco E, 3º andar, São Domingos, Niterói, RJ CEP 24210-240, Brazil, bicharra@ic.uff.br

M.L. Caliusco CONICET, CIDISI-UTN-FRSF, Lavaise 610, Santa Fe, Argentina, mcaliusc@frsf.utn.edu.ar

Ana Cavalli SAMOVAR CNRS UMR 5157, Telecom & Management SudParis, 9 rue Charles Fourier, 91011 Evry Cedex, France, Ana.Cavalli@it-sudparis.eu

W.S. Chan Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong, wschan@comp.hkbu.edu.hk

Nader Cheaib IBISC CNRS FRE 3190, University of Evry, 91020 Evry Cedex, France, nader.cheaib@ibisc.fr

Xiaoling Dai The University of the South Pacific, Suva, Fiji Islands, dai_s@usp.ac.fj

Dominique Decouchant Laboratoire LIG de Grenoble, Grenoble, France; UAM-Cuajimalpa, México D.F., México, Dominique.Decouchant@imag.fr

Thierry Despeyroux INRIA Paris-Rocquencourt, Domaine de Voluceau, B.P. 105, F-78153 Le Chesnay Cedex, France, thierry.despeyroux@inria.fr

John Grundy The University of Auckland, Auckland, New Zealand,
john-g@cs.auckland.ac.nz

Marcus V.C. Guelpeli Departamento de Ciência da Computação, Instituto de Computação—IC, Universidade Federal Fluminense—UFF, Rua Passo da Pátria 156, Bloco E, 3º andar, São Domingos, Niterói, RJ CEP 24210-240, Brazil,
mguelpe@ic.uff.br

Wolfgang A. Halang Fernuniversität, 58084 Hagen, Germany,
wolfgang.halang@fernuni-hagen.de

Hiroshi Ishikawa Faculty of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka, Japan, ishikawah@acm.org

Anastasia Karanastasi Decision Support Systems Laboratory, Technical University of Crete, Chania, 73100, Greece, natasha@ergasya.tuc.gr

M-Tahar Kechadi School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland, tahar.kechadi@ucd.ie

Nhien An Khac School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland, an.lekhac@ucd.ie

Mounir Lallali SAMOVAR CNRS UMR 5157, Telecom & Management SudParis, 9 rue Charles Fourier, 91011 Evry Cedex, France,
Mounir.Lallali@it-sudparis.eu

C.H.C. Leung Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong, clement@comp.hkbu.edu.hk

J. Liu Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong, jiming@comp.hkbu.edu.hk

Stephane Maag SAMOVAR CNRS UMR 5157, Telecom & Management SudParis, 9 rue Charles Fourier, 91011 Evry Cedex, France,
Stephane.Maag@it-sudparis.eu

Malik Mallem IBISC CNRS FRE 3190, University of Evry, 91020 Evry Cedex, France, malik.mallem@ibisc.fr

Nikolaos Matsatsinis Decision Support Systems Laboratory, Technical University of Crete, Chania, 73100, Greece, nikos@ergasya.tuc.gr

Klaus Meißner Chair of Multimedia Technology, Technische Universität Dresden, Dresden, Germany, kmeiss@inf.tu-dresden.de

Sonia Mendoza Departamento de Computación, CINVESTAV-IPN, México D.F., México, smendoza@cs.cinvestav.mx

A. Milani Department of Mathematics & Computer Science, University of Perugia, Perugia, Italy, milani@unipg.it

Gerardo Morales SAMOVAR CNRS UMR 5157, Telecom & Management
SudParis, 9 rue Charles Fourier, 91011 Evry Cedex, France,
Gerardo.Morales@it-sudparis.eu

Malcolm Munro Department of Computer Science, Durham University, Durham,
UK, malcolm.munro@durham.ac.uk

Matthias Niederhausen Chair of Multimedia Technology, Technische Universität
Dresden, Dresden, Germany, matthias.niederhausen@tu-dresden.de

Samir Otmane IBISC CNRS FRE 3190, University of Evry, 91020 Evry Cedex,
France, samir.otmane@ibisc.fr

Alberto Pepe Center for Embedded Networked Sensing, University of California
at Los Angeles, 3551 Boelter Hall, Los Angeles, CA 90095-1596, USA,
apepe@ucla.edu

Stefan Pietschmann Chair of Multimedia Technology, Technische Universität
Dresden, Dresden, Germany, stefan.pietschmann@tu-dresden.de

José Rodríguez Departamento de Computación, CINVESTAV-IPN, México D.F.,
México, rodriguez@cs.cinvestav.mx

Marko A. Rodriguez T-5, Center for Nonlinear Studies, Los Alamos National
Laboratory, Los Alamos, NM 87545, USA, marko@lanl.gov

Tobias Ruch T-Systems Multimedia Solutions GmbH, Dresden, Germany,
tobias.ruch@t-systems.com

Joshua Shinavier Semantic Network Research Group, Knowledge Reef Systems
Inc., Santa Fe, NM 87501, USA, josh@fortytwo.net

G. Stegmayer CONICET, CIDISI-UTN-FRSF, Lavaise 610, Santa Fe, Argentina,
gstegmayer@santafe-conicet.gov.ar

Xia Wang Fernuniversität, 58084 Hagen, Germany, xia.wang@fernuni-hagen.de

Marek Wisniewski Poznan University of Economics, Al. Niepodleglosci 10,
Poznan, Poland

Fatiha Zaidi LRI, Université Paris-Sud 11, 91405 Orsay Cedex, France,
Fatiha.Zaidi@lri.fr

Yi Zhao Fernuniversität, 58084 Hagen, Germany, yi.zhao@fernuni-hagen.de