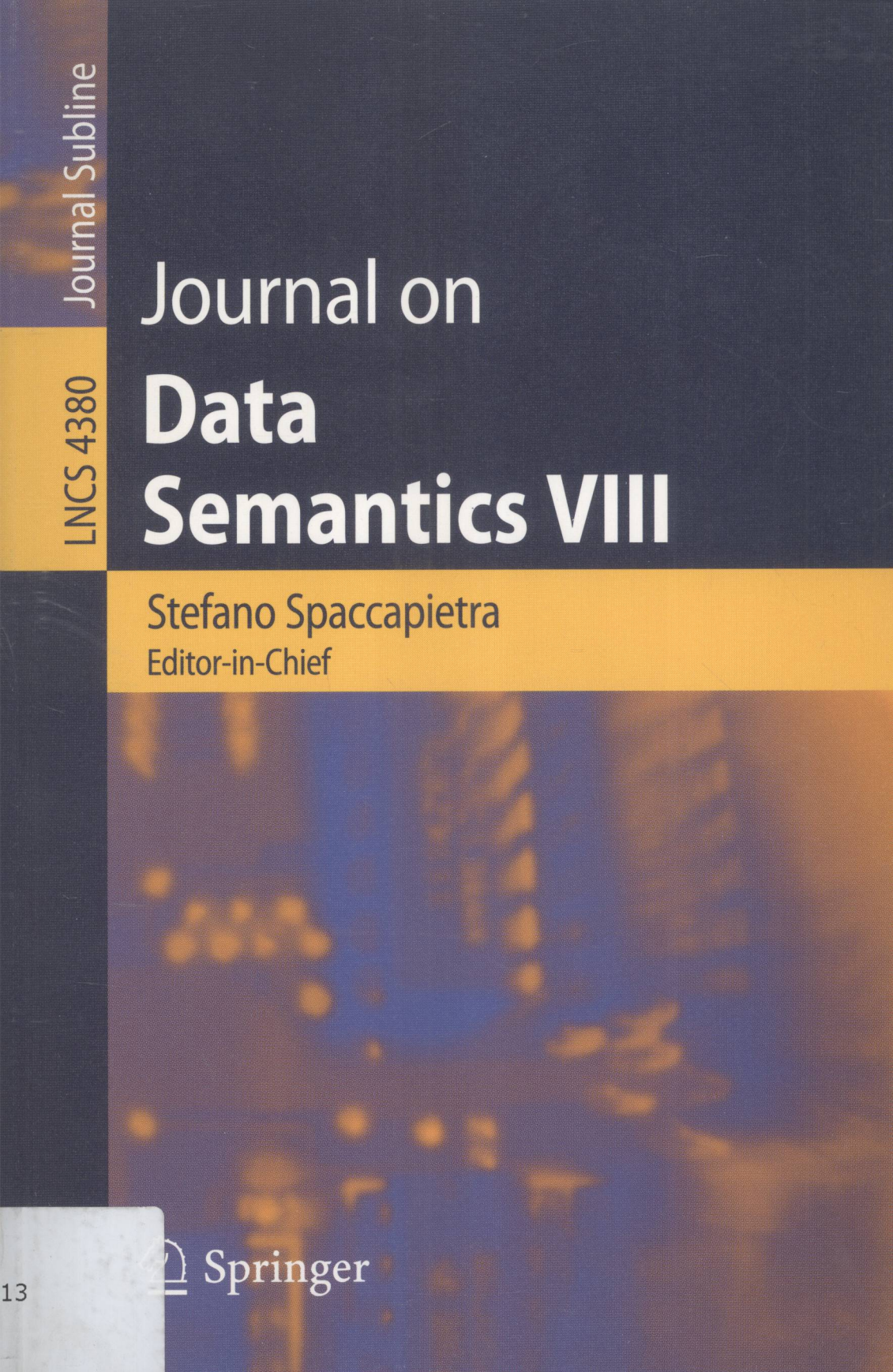
Journal Subline

LNCS 4380

Journal on Data Semantics VIII

Stefano Spaccapietra
Editor-in-Chief



Springer

13

TP311.13
J86.7

Stefano Spaccapietra Paolo Atzeni
François Fages Mohand-Saïd Hacid
Michael Kifer John Mylopoulos
Barbara Pernici Pavel Shvaiko
Juan Trujillo Ilya Zaihrayeu (Eds.)

Journal on Data Semantics VIII

Springer

Volume Editors

Stefano Spaccapietra
EPFL, Lausanne, Switzerland, E-mail: stefano.spaccapietra@epfl.ch

Paolo Atzeni
Università Roma Tre, Italy, E-mail: atzeni@dia.uniroma3.it

François Fages
INRIA Rocquencourt, France, E-mail: Francois.Fages@inria.fr

Mohand-Saïd Hacid
Université Lyon 1, France, E-mail: mshacid@bat710.univ-lyon1.fr

Michael Kifer
State University of New York at Stony Brook, USA, E-mail: kifer@cs.sunysb.edu

John Mylopoulos
University of Toronto, Canada/University of Trento, Italy, E-mail: jm@cs.toronto.edu

Barbara Pernici
Politecnico di Milano, Italy, E-mail: barbara.pernici@polimi.it

Pavel Shvaiko
Ilya Zaihrayeu
University of Trento, Italy, E-mail: {pavel,ilya}@dit.unitn.it

Juan Trujillo
University of Alicante, Spain, E-mail: jtrujillo@dlsi.ua.es

CR Subject Classification (1998): H.2, H.3, I.2, H.4, C.2

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 1861-2032
ISBN-10 3-540-70663-1 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-70663-2 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11980711 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 4380

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK
Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA
Josef Kittler
University of Surrey, Guildford, UK
Jon M. Kleinberg
Cornell University, Ithaca, NY, USA
Friedemann Mattern
ETH Zurich, Switzerland
John C. Mitchell
Stanford University, CA, USA
Moni Naor
Weizmann Institute of Science, Rehovot, Israel
Oscar Nierstrasz
University of Bern, Switzerland
C. Pandu Rangan
Indian Institute of Technology, Madras, India
Bernhard Steffen
University of Dortmund, Germany
Madhu Sudan
Massachusetts Institute of Technology, MA, USA
Demetri Terzopoulos
University of California, Los Angeles, CA, USA
Doug Tygar
University of California, Berkeley, CA, USA
Moshe Y. Vardi
Rice University, Houston, TX, USA
Gerhard Weikum
Max-Planck Institute of Computer Science, Saarbruecken, Germany

The LNCS Journal on Data Semantics

Computerized information handling has changed its focus from centralized data management systems to decentralized data exchange facilities. Modern distribution channels, such as high-speed Internet networks and wireless communication infrastructure, provide reliable technical support for data distribution and data access, materializing the new, popular idea that data may be available to anybody, anywhere, anytime. However, providing huge amounts of data on request often turns into a counterproductive service, making the data useless because of poor relevance or inappropriate level of detail. Semantic knowledge is the essential missing piece that allows the delivery of information that matches user requirements. Semantic agreement, in particular, is essential to meaningful data exchange.

Semantic issues have long been open issues in data and knowledge management. However, the boom in semantically poor technologies, such as the Web and XML, has boosted renewed interest in semantics. Conferences on the Semantic Web, for instance, attract big crowds of participants, while ontologies on their own have become a hot and popular topic in the database and artificial intelligence communities.

Springer's LNCS *Journal on Data Semantics* aims at providing a highly visible dissemination channel for most remarkable work that in one way or another addresses research and development on issues related to the semantics of data. The target domain ranges from theories supporting the formal definition of semantic content to innovative domain-specific application of semantic knowledge. This publication channel should be of the highest interest to researchers and advanced practitioners working on the Semantic Web, interoperability, mobile information services, data warehousing, knowledge representation and reasoning, conceptual database modeling, ontologies, and artificial intelligence.

Topics of relevance to this journal include:

- Semantic interoperability, semantic mediators
- Ontologies
- Ontology, schema and data integration, reconciliation and alignment
- Multiple representations, alternative representations
- Knowledge representation and reasoning
- Conceptualization and representation
- Multi-model and multi-paradigm approaches
- Mappings, transformations, reverse engineering
- Metadata
- Conceptual data modeling
- Integrity description and handling
- Evolution and change
- Web semantics and semi-structured data

此为试读,需要完整PDF请访问: www.ertongbook.com

- Semantic caching
- Data warehousing and semantic data mining
- Spatial, temporal, multimedia and multimodal semantics
- Semantics in data visualization
- Semantic services for mobile users
- Supporting tools
- Applications of semantic-driven approaches

These topics are to be understood as specifically related to semantic issues. Contributions submitted to the journal and dealing with semantics of data will be considered even if they are not from the topics in the list.

While the physical appearance of the journal issues is like the books from the well-known Springer LNCS series, the mode of operation is that of a journal. Contributions can be freely submitted by authors and are reviewed by the Editorial Board. Contributions may also be invited, and nevertheless carefully reviewed, as in the case for issues that contain extended versions of best papers from major conferences addressing data semantics issues. Special issues, focusing on a specific topic, are coordinated by guest editors once the proposal for a special issue is accepted by the Editorial Board. Finally, it is also possible that a journal issue be devoted to a single text.

The journal published its first volume in 2003 (LNCS 2800). That initial volume, as well as volumes II (LNCS 3360), V (LNCS 3870), this volume, VIII, and the next volume, IX, represent the annual occurrence of a special issue devoted to publication of selected extended versions of best conference papers from the previous year's conferences. Volumes III and VI were special issues on a dedicated topic. Volume III (LNCS 3534), coordinated by guest editor Esteban Zimányi, addressed Semantic-Based Geographical Information Systems, while volume VI (LNCS 4090), coordinated by guest editors Karl Aberer and Philippe Cudre-Mauroux, addressed Emergent Semantics. Volumes IV and VI were "normal" volumes, built from spontaneous submissions on any of the topics of interest to the journal.

The Editorial Board comprises an Editor-in-Chief (with overall responsibility), a Co-editor-in-Chief, and several members. The Editor-in-Chief has a four-year mandate. Members of the board have a three-year mandate. Mandates are renewable, and new members may be elected anytime.

We are happy to welcome you to our readership and authorship, and hope we will share this privileged contact for a long time.

Stefano Spaccapietra
Editor-in-Chief
http://lbdwww.epfl.ch/e/Springer/

JoDS Volume VIII

To foster the dissemination of the best ideas and results, the *Journal on Data Semantics* (JoDS) pursues a policy that includes annually publishing extended versions of the best papers from selected conferences whose scope encompasses or intersects with the scope of the journal.

This initiative is motivated by the difference in goals between conferences and journals. Conferences usually have a faster turnaround and a focused audience, but they have to enforce space limitation and a fixed time frame, with no chances for improving a paper by producing multiple versions. In contrast, journals offer more space, room for debate and refinement, and are usually considered the real archival venue.

Therefore, the publication of an extended version of a conference paper is a much appreciated opportunity for researchers to widely disseminate a significantly improved presentation of their work, where they can develop the appropriate motivations, reasoning, results and comparative analysis. Moreover, by gathering best papers from various conferences, JoDS special issues provide a unique opportunity for researchers to find in a single publication every year the best of ongoing research in the field of data semantics.

For this issue, papers from the following six 2005 international conferences were invited:

- The Seventh International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2005),
- The 3rd International Workshop on Principles and Practice of Semantic Web Reasoning, (PPSWR 2005),
- The 1st International Workshop on Contexts and Ontologies, Theory, Practice and Applications (C&O 2005), joint event with the 25th National Conference on Artificial Intelligence (AAAI 2005),
- The 2nd International Workshop on Peer-to-Peer Knowledge Management (P2PKM 2005), joint event with the 2nd Annual International Conference on Mobile and Ubiquitous systems (Mobiquitous 2005)
- The 13th International Conference on Cooperative Information Systems, (CoopIS 2005) and the International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE 2005), which both took place October 31 to November 4, 2005, Agia Napa, Cyprus.

In addition, this issue includes one paper from the 23rd International Conference on Conceptual Modeling (ER 2004), which was accepted but could not be included in JoDS V because of late delivery of the final version.

Papers from these conferences were invited based on their quality, relevance and significance, and the viability of extending their results. Extended versions prepared by authors were subject to the traditional two-round scholarly review process, and the authors were required to respond to all concerns expressed by the reviewers before papers were accepted.

The paper by Velcin and Ganascia, originating from ER 2004, considers a model dealing with sparse data sets and describes a theoretical framework for inducing knowledge out of them. The general framework relies on a lattice structure, and it is illustrated within two formalisms: the attribute-value formalism and Sowa's conceptual graphs. The induction engine is based on a non-supervised algorithm called default clustering, which uses the concept of stereotype and a new notion of default subsumption, inspired by the default logic theory.

The selection of CoopIS 2005 best papers eventually resulted in the acceptance of three papers. The paper "Semantic Matching: Algorithms and Implementation" by Giunchiglia, Yatskevich and Shvaiko provides a framework that views *match* as an operator that takes two graph-like structures and produces a mapping between the nodes of the graphs that correspond semantically to each other. The authors introduce model-based techniques at the structure level.

The paper "Semantic-Guided Clustering of Heterogeneous XML Schemas" by De Meo, Quattrone, Terracina and Ursino investigates a semantic-based approach for clustering heterogeneous XML schemas. The proposed approach makes use of the semantics of the underlying schemas by capturing the interschema properties among concepts of the schemas. The experimental analysis shows that the approach is scalable.

The paper "A Formal Framework for Adaptive Access Control Models" by Rinderle and Reichert proposes a framework that is suitable for handling evolution of organizational models and related access rules. The paper introduces a set of well-defined operators for defining and capturing changes in organizational models. In addition, the framework allows adaptation of access rules when the model changes.

DaWak 2005 contributed its best paper, "Processing Sequential Patterns in Relational Databases" by Shang and Sattler. The paper proposes an efficient SQL-based algorithm to mine sequential patterns in relational database systems. Authors start by saying that traditionally data mining techniques have been applied on flat files instead of on databases due to the low performance and high cost associated in implementing data mining with SQL on relational databases. Authors claim that it is possible to achieve a reasonable performance by implementing association rule mining and sequential pattern mining with carefully tuned SQL formulations. To this extent, authors depart from inefficient a-priori methods and propose an efficient SQL-based algorithm, called Prospad (PROjection Sequential PAttern Discovery), to mine sequential patterns in relational database systems. Prospad adopts the divide-and-conquer strategy and projects the sequence table into a set of frequent item-related projected tables. Experimental results show that the Prospad algorithm can get higher performance than k-way joins based on a-priori approaches, especially on large and dense datasets, although it has severe limitations in performance compared to in-memory PrefixSpan algorithms.

Two extended articles were selected from PPSWR 2005. The paper "A Tool for Evaluating Ontology Alignment Strategies," by Lambrix and Tan, addresses the important issue of aligning different ont logies, so that multiple sources of information can be exploited altogether. The paper describes a framework for the comparative evaluation of ontology alignment strategies and their combinations, and reports on the performance of an implementation of this framework. The test cases

used for the evaluation are composed of five biomedical ontologies. A detailed example shows the use of two matchers in combination.

The paper "SomeRDFS in the Semantic Web," by Adjiman, Goasdoué and Rousset, envisions the Semantic Web as a huge peer data management system, where data on the Web are annotated by ontologies networked together by mappings. The paper describes the SomeRDFS peer data management system architecture, its data model, query language, and query answering algorithm based on query rewriting techniques that are formally justified in the paper.

The selection from C&O 2005 resulted in three extended papers being accepted for JoDS.

The paper "Putting Things in Context: A Topological Approach to Mapping Contexts to Ontologies" by Segev and Gal, provides a framework that defines the relationship between contexts and ontologies by using topological structures. This work has been motivated by the needs of the eGovernment domain. In this approach ontologies are viewed as the result of a manual effort to model a domain, while contexts are automatically generated models. The uncertainty, which usually exists in automatic context extraction, is managed through the definition of distance among contexts and a ranking of ontology concepts with respect to a given context. The approach has been implemented and evaluated on two real-world data sets: Reuters news reports and RSS news headlines.

The paper "Context Dependency Management in Ontology Engineering: A Formal Approach," by De Leenheer, de Moor, and Meersman, introduces a framework that uses lexical knowledge to manage context dependences in ontology engineering tasks. The formalization of the context dependency management is built on top of the DOGMA ontology-engineering framework. The proposed approach is validated by a case study of inter-organizational competency ontology engineering.

The paper "Encoding Classifications into Lightweight Ontologies" by Giunchiglia, Marchese, and Zaihrayeu provides a theory of how to translate standard classifications, such as DMoz, into formal classifications, namely, graph structures where labels are written in a propositional concept language. Formal classifications turn out to be a form of lightweight ontologies. This allows reducing essential tasks on classifications, such as document classification and query answering, to reasoning about subsumption.

While the paper selected from P2PKM is still in the review process at this time, four extended versions of ODBASE 2005 papers were granted acceptance. The paper "Creating Ontologies for Content Representation - The OntoSeed Suite," by Paslaru Bontas Simperl and Schlangen, proposes a natural language-based technique to help ontology engineers decide which concepts to model in any particular domain. Unlike other NLP-based techniques, this approach does not require in-depth linguistic expertise. Instead it relies on the Web for collecting the documents against which to compare domain-specific texts.

"Metadata Management in a Multiversion Data Warehouse," a paper by Wrembel and Bebel, deals with the problem of evolving data warehouses in the presence of changes in the schema of the underlying data sources. The paper proposes a solution, called multiversion data warehouse, which maintains extensive metadata about the external data sources.

The paper by Kensche, Quix, Chatti, and Jarke, "GeRoMe: A Generic Role-Based Metamodel for Model Management," proposes a generic mechanism for describing data models. This approach assigns multiple roles to model elements and permits accurate description of these elements using only a small number of roles and metaclasses. This contrasts favorably with metamodel languages that are based exclusively on metaclasses, since such languages may require an exponentially large number of metaclasses.

The paper on "Security Ontology for Annotating Resources," by Kim, Luo, and Kang, proposes an ontology for describing security requirements of Web services. The goal of such an ontology is to enable Web service discovery that meets a client's security requirements, such as protocols, objectives, and credentials. The proposed ontology is more comprehensive and detailed than other similar ontologies.

Because of size limitations, only eight of the 14 above-described papers appear in this volume. The other six papers will appear in the next volume, JoDS IX, scheduled for Spring 2007.

C&O 2005 Co-chair
Pavel Shvaiko, University of Trento, Italy

CoopIS 2005 PC Co-chairs
Mohand-Saïd Hacid, University Lyon 1, France
John Mylopoulos, University of Trento, Italy and University of Toronto, Canada
Barbara Pernici, Politecnico Milano, Italy

DAWAK 2005 Co-chair
Juan Trujillo, University of Alicante, Spain

ER 2004 Co-chair
Paolo Atzeni

ODBASE 2005 Co-chairs
Michael Kifer, State University of New York at Stony Brook, USA
Stefano Spaccapietra, EPFL, Switzerland

PPSWR 2005 Co-chair
François Fages, INRIA Rocquencourt, France

P2PKM Co-chair
Ilya Zaihrayeu, University of Trento, Italy

Organization

Reviewers

We would like to express our gratitude to the following colleagues who helped in the review process by contributing detailed reviews of the submitted papers and provided invaluable feedback to the authors and editors alike:

Roberta Benassi, Università degli Studi di Modena e Reggio Emilia, Italy
Alex Borgida, Rutgers University, USA
Mohand Boughanem, Université Paul Sabatier, France
Paolo Bouquet, University of Trento, Italy
François Bry, University of Munich, Germany
Barbara Carminati, Università degli Studi dell'Insubria, Italy
Yen-Liang Chen, National Central University, Taiwan
Sunil Chenoi, University of Twente, The Netherlands
Paul D. Clough, University of Sheffield, UK
Emmanuel Coquery, Université Lyon 1, France
Grit Denker, SRI, USA
Johann Eder, University of Vienna, Austria
Jerome Euzenat, INRIA Rhône-Alpes, France
Ling Feng, University of Twente, The Netherlands
Enrico Franconi, Free University of Bozen-Bolzano, Italy
Norbert Fuchs, University of Zurich, Switzerland
Avigdor Gal, Technion, Israel
Anna Goy, University of Torino, Italy
Andreas Herzig, Université Paul Sabatier, France
Pascal Hitzler, University of Karlsruhe, Germany
Carlos Hurtado, University of Chile, Chile
Valérie Issarny, INRIA Rocquencourt, France
Yaron Kanza, University of Toronto, Canada
Vipul Kashyap, Clinical Informatics R&D, Partners HealthCare System, USA
Zoubida Kedad, Université de Versailles St-Quentin en Yvelines, Frances
Akhil Kumar, Penn State University, USA
Ruben Lara, Universidad Autonoma de Madrid, Spain
Alain Leger, France Telecom R&D, France
Jan Maluszynski, Linköping University, Sweden
Massimo Marchiori, MIT Cambridge, USA
Sergei Melnik, Microsoft Research, USA
Saikat Mukherjee, Siemens Research, USA
Roberto Navigli, University of Rome, Italy
Natalya Noy, Stanford University, USA
Hans-Jürgen Ohlbach, University of Munich, Germany

此为试读,需要完整PDF请访问: www.ertongbook.com

Antoni Olive, Universitat Politècnica de Catalunya, Spain
Massimo Paolucci, University of Genova, Italy
Torben Bach Pedersen, Aalborg University, Denmark
Jean-Marc Petit, INSA de Lyon, France
Dimitris Plexousakis, University of Crete, Greece
Axel Polleres, Universidad Rey Juan Carlos, Spain
David Robertson, The University of Edinburgh, UK
Marie-Christine Rousset, IMAG, France
Peter Pater-Schneider, Bell Labs, Murray Hill, USA
Marta Sabou, The Open University, UK
Vasile-Marian Scuturici, INSA de Lyon, France
Manuel Serrano, University of Castilla La Mancha, Spain
Sylvain Soliman, INRIA Rocquencourt, France
Heiner Stuckenschmidt, University of Mannheim, Germany
York Sure, University of Karlsruhe, Germany
Thodoros Topaloglou, University of Toronto, Canada
Farouk Toumani, Université Blaise Pascal, France
Athena Vakali, Aristotle University of Thessaloniki, Greece
Panos Vassiliadis, University of Ioannina, Greece
Holger Wache, Vrije Universiteit Amsterdam, The Netherlands
Howard Williams, Heriot-Watt University, Edinburgh, UK
Xing Xie, Microsoft Research, USA
Esteban Zimányi, Université Libre de Bruxelles, Belgium

JoDS Editorial Board

Co-Editors-in-Chief Lois Delcambre, Portland State University, USA
Stefano Spaccapietra, EPFL, Switzerland

Members of the Board
Carlo Batini, Università di Milano Bicocca, Italy
Alex Borgida, Rutgers University, USA
Shawn Bowers, University of California Davis, USA
Tiziana Catarci, Università di Roma La Sapienza, Italy
David W. Embley, Brigham Young University, USA
Jerome Euzenat, INRIA Alpes, France
Dieter Fensel, University of Innsbruck, Austria
Nicola Guarino, National Research Council, Italy
Jean-Luc Hainaut, FUNDP Namur, Belgium
Ian Horrocks, University of Manchester, UK
Arantza Illarramendi, Universidad del País Vasco, Spain
Larry Kerschberg, George Mason University, USA
Michael Kifer, State University of New York at Stony Brook, USA
Tok Wang Ling, National University of Singapore, Singapore

Shamkant B. Navathe, Georgia Institute of Technology, USA
Antoni Olivé, Universitat Politècnica de Catalunya, Spain
José Palazzo M. de Oliveira, Universidade Federal do Rio Grande do Sul, Brazil
Christine Parent, Université de Lausanne, Switzerland
John Roddick, Flinders University, Australia
Klaus-Dieter Schewe, Massey University, New Zealand
Heiner Stuckenschmidt, University of Mannheim, Germany
Katsumi Tanaka, University of Kyoto, Japan
Yair Wand, University of British Columbia, Canada
Eric Yu, University of Toronto, Canada
Esteban Zimányi, Université Libre de Bruxelles (ULB), Belgium

Lecture Notes in Computer Science

For information about Vols. 1–4281

please contact your bookseller or Springer

Vol. 4384: T. Washio, K. Satoh, H. Takeda, A. Inokuchi (Eds.), New Frontiers in Artificial Intelligence. IX, 401 pages. 2007. (Sublibrary LNAI).

Vol. 4380: S. Spaccapietra, P. Atzeni, F. Fages, M.-S. Hacid, M. Kifer, J. Mylopoulos, B. Pernici, P. Shvaiko, J. Trujillo, I. Zaihrayeu (Eds.), Journal on Data Semantics VIII. XV, 219 pages. 2007.

Vol. 4377: M. Abe (Ed.), Topics in Cryptology – CT-RSA 2007. XI, 403 pages. 2006.

Vol. 4373: K. Langendoen, T. Voigt (Eds.), Wireless Sensor Networks. XIII, 358 pages. 2006.

Vol. 4371: K. Inoue, K. Satoh, F. Toni (Eds.), Computational Logic in Multi-Agent Systems. X, 315 pages. 2007. (Sublibrary LNAI).

Vol. 4369: M. Umeda, A. Wolf, O. Bartenstein, U. Geske, D. Seipel, O. Takata (Eds.), Declarative Programming for Knowledge Management. X, 229 pages. 2006. (Sublibrary LNAI).

Vol. 4368: T. Erlebach, C. Kaklamanis (Eds.), Approximation and Online Algorithms. X, 345 pages. 2007.

Vol. 4367: K. De Bosschere, D. Kaeli, P. Stenström, D. Whalley, T. Ungerer (Eds.), High Performance Embedded Architectures and Compilers. XI, 307 pages. 2007.

Vol. 4364: T. Kühne (Ed.), Models in Software Engineering. XI, 332 pages. 2007.

Vol. 4362: J. van Leeuwen, G.F. Italiano, W. van der Hoek, C. Meinel, H. Sack, F. Plášil (Eds.), SOFSEM 2007: Theory and Practice of Computer Science. XXI, 937 pages. 2007.

Vol. 4361: H.J. Hoogeboom, G. Păun, G. Rozenberg, A. Salomaa (Eds.), Membrane Computing. IX, 555 pages. 2006.

Vol. 4360: W. Dubitzky, A. Schuster, P.M.A. Sloot, M. Schroeder, M. Romberg (Eds.), Distributed, High-Performance and Grid Computing in Computational Biology. X, 192 pages. 2007. (Sublibrary LNBI).

Vol. 4357: L. Buttyán, V. Gligor, D. Westhoff (Eds.), Security and Privacy in Ad-Hoc and Sensor Networks. X, 193 pages. 2006.

Vol. 4355: J. Julliand, O. Kouchnarenko (Eds.), B 2007: Formal Specification and Development in B. XIII, 293 pages. 2006.

Vol. 4354: M. Hanus (Ed.), Practical Aspects of Declarative Languages. X, 335 pages. 2006.

Vol. 4353: T. Schwentick, D. Suciu (Eds.), Database Theory – ICDT 2007. XI, 419 pages. 2006.

Vol. 4352: T.-J. Cham, J. Cai, C. Dorai, D. Rajan, T.-S. Chua, L.-T. Chia (Eds.), Advances in Multimedia Modeling, Part II. XVIII, 743 pages. 2006.

Vol. 4351: T.-J. Cham, J. Cai, C. Dorai, D. Rajan, T.-S. Chua, L.-T. Chia (Eds.), Advances in Multimedia Modeling, Part I. XIX, 797 pages. 2006.

Vol. 4349: B. Cook, A. Podelski (Eds.), Verification, Model Checking, and Abstract Interpretation. XI, 395 pages. 2007.

Vol. 4348: S.T. Taft, R.A. Duff, R.L. Brukardt, E. Ploedereder, P. Leroy (Eds.), Ada 2005 Reference Manual. XXII, 765 pages. 2006.

Vol. 4347: J. Lopez (Ed.), Critical Information Infrastructures Security. X, 286 pages. 2006.

Vol. 4345: N. Maglaveras, I. Chouvarda, V. Koutkias, R. Brause (Eds.), Biological and Medical Data Analysis. XIII, 496 pages. 2006. (Sublibrary LNBI).

Vol. 4344: V. Gruhn, F. Oquendo (Eds.), Software Architecture. X, 245 pages. 2006.

Vol. 4342: H. de Swart, E. Orłowska, G. Schmidt, M. Roubens (Eds.), Theory and Applications of Relational Structures as Knowledge Instruments II. X, 373 pages. 2006. (Sublibrary LNAI).

Vol. 4341: P.Q. Nguyen (Ed.), Progress in Cryptology - VIETCRYPT 2006. XI, 385 pages. 2006.

Vol. 4340: R. Prodan, T. Fahringer, Grid Computing. XXIII, 317 pages. 2007.

Vol. 4339: E. Ayguadé, G. Baumgartner, J. Ramanujam, P. Sadayappan (Eds.), Languages and Compilers for Parallel Computing. XI, 476 pages. 2006.

Vol. 4338: P. Kalra, S. Peleg (Eds.), Computer Vision, Graphics and Image Processing. XV, 965 pages. 2006.

Vol. 4337: S. Arun-Kumar, N. Garg (Eds.), FSTTCS 2006: Foundations of Software Technology and Theoretical Computer Science. XIII, 430 pages. 2006.

Vol. 4335: S.A. Brueckner, S. Hassas, M. Jelasity, D. Yamins (Eds.), Engineering Self-Organising Systems. XII, 212 pages. 2007. (Sublibrary LNAI).

Vol. 4334: B. Beckert, R. Hähnle, P.H. Schmitt (Eds.), Verification of Object-Oriented Software. XXIX, 658 pages. 2007. (Sublibrary LNAI).

Vol. 4333: U. Reimer, D. Karagiannis (Eds.), Practical Aspects of Knowledge Management. XII, 338 pages. 2006. (Sublibrary LNAI).

Vol. 4332: A. Bagchi, V. Atluri (Eds.), Information Systems Security. XV, 382 pages. 2006.

Vol. 4331: G. Min, B. Di Martino, L.T. Yang, M. Guo, G. Ruenger (Eds.), Frontiers of High Performance Computing and Networking – ISPA 2006 Workshops. XXXVII, 1141 pages. 2006.

Vol. 4330: M. Guo, L.T. Yang, B. Di Martino, H.P. Zima, J. Dongarra, F. Tang (Eds.), Parallel and Distributed Processing and Applications. XVIII, 953 pages. 2006.

Vol. 4329: R. Barua, T. Lange (Eds.), Progress in Cryptology - INDOCRYPT 2006. X, 454 pages. 2006.

Vol. 4328: D. Penkler, M. Reitenspiess, F. Tam (Eds.), Service Availability. X, 289 pages. 2006.

Vol. 4327: M. Baldoni, U. Endriss (Eds.), Declarative Agent Languages and Technologies IV. VIII, 257 pages. 2006. (Sublibrary LNAI).

Vol. 4326: S. Göbel, R. Malkewitz, I. Iurgel (Eds.), Technologies for Interactive Digital Storytelling and Entertainment. X, 384 pages. 2006.

Vol. 4325: J. Cao, I. Stojmenovic, X. Jia, S.K. Das (Eds.), Mobile Ad-hoc and Sensor Networks. XIX, 887 pages. 2006.

Vol. 4323: G. Doherty, A. Blandford (Eds.), Interactive Systems. XI, 269 pages. 2006.

Vol. 4320: R. Gotzhein, R. Reed (Eds.), System Analysis and Modeling: Language Profiles. X, 229 pages. 2006.

Vol. 4319: L.-W. Chang, W.-N. Lie (Eds.), Advances in Image and Video Technology. XXVI, 1347 pages. 2006.

Vol. 4318: H. Lipmaa, M. Yung, D. Lin (Eds.), Information Security and Cryptology. XI, 305 pages. 2006.

Vol. 4317: S.K. Madria, K.T. Claypool, R. Kannan, P. Uppuluri, M.M. Gore (Eds.), Distributed Computing and Internet Technology. XIX, 466 pages. 2006.

Vol. 4316: M.M. Dalkilic, S. Kim, J. Yang (Eds.), Data Mining and Bioinformatics. VIII, 197 pages. 2006. (Sublibrary LNBI).

Vol. 4313: T. Margaria, B. Steffen (Eds.), Leveraging Applications of Formal Methods. IX, 197 pages. 2006.

Vol. 4312: S. Sugimoto, J. Hunter, A. Rauber, A. Morishima (Eds.), Digital Libraries: Achievements, Challenges and Opportunities. XVIII, 571 pages. 2006.

Vol. 4311: K. Cho, P. Jacquet (Eds.), Technologies for Advanced Heterogeneous Networks II. XI, 253 pages. 2006.

Vol. 4309: P. Inverardi, M. Jazayeri (Eds.), Software Engineering Education in the Modern Age. VIII, 207 pages. 2006.

Vol. 4308: S. Chaudhuri, S.R. Das, H.S. Paul, S. Tirthapura (Eds.), Distributed Computing and Networking. XIX, 608 pages. 2006.

Vol. 4307: P. Ning, S. Qing, N. Li (Eds.), Information and Communications Security. XIV, 558 pages. 2006.

Vol. 4306: Y. Avrithis, Y. Kompatsiaris, S. Staab, N.E. O'Connor (Eds.), Semantic Multimedia. XII, 241 pages. 2006.

Vol. 4305: A.A. Shvartsman (Ed.), Principles of Distributed Systems. XIII, 441 pages. 2006.

Vol. 4304: A. Sattar, B.-H. Kang (Eds.), AI 2006: Advances in Artificial Intelligence. XXVII, 1303 pages. 2006. (Sublibrary LNAI).

Vol. 4303: A. Hoffmann, B.-H. Kang, D. Richards, S. Tsumoto (Eds.), Advances in Knowledge Acquisition and Management. XI, 259 pages. 2006. (Sublibrary LNAI).

Vol. 4302: J. Domingo-Ferrer, L. Franconi (Eds.), Privacy in Statistical Databases. XI, 383 pages. 2006.

Vol. 4301: D. Pointcheval, Y. Mu, K. Chen (Eds.), Cryptology and Network Security. XIII, 381 pages. 2006.

Vol. 4300: Y.Q. Shi (Ed.), Transactions on Data Hiding and Multimedia Security I. IX, 139 pages. 2006.

Vol. 4299: S. Renals, S. Bengio, J.G. Fiscus (Eds.), Machine Learning for Multimodal Interaction. XII, 470 pages. 2006.

Vol. 4297: Y. Robert, M. Parashar, R. Badrinath, V.K. Prasanna (Eds.), High Performance Computing - HiPC 2006. XXIV, 642 pages. 2006.

Vol. 4296: M.S. Rhee, B. Lee (Eds.), Information Security and Cryptology – ICISC 2006. XIII, 358 pages. 2006.

Vol. 4295: J.D. Carswell, T. Tezuka (Eds.), Web and Wireless Geographical Information Systems. XI, 269 pages. 2006.

Vol. 4294: A. Dan, W. Lamersdorf (Eds.), Service-Oriented Computing – ICSOC 2006. XIX, 653 pages. 2006.

Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), MICAI 2006: Advances in Artificial Intelligence. XXVIII, 1232 pages. 2006. (Sublibrary LNAI).

Vol. 4292: G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. Nefian, G. Meenakshisundaram, V. Pascucci, J. Zara, J. Molineros, H. Theisel, T. Malzbender (Eds.), Advances in Visual Computing, Part II. XXXII, 906 pages. 2006.

Vol. 4291: G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. Nefian, G. Meenakshisundaram, V. Pascucci, J. Zara, J. Molineros, H. Theisel, T. Malzbender (Eds.), Advances in Visual Computing, Part I. XXXI, 916 pages. 2006.

Vol. 4290: M. van Steen, M. Henning (Eds.), Middleware 2006. XIII, 425 pages. 2006.

Vol. 4289: M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenič, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svátek, M. van Someren (Eds.), Semantics, Web and Mining. X, 197 pages. 2006. (Sublibrary LNAI).

Vol. 4288: T. Asano (Ed.), Algorithms and Computation. XX, 766 pages. 2006.

Vol. 4287: C. Mao, T. Yokomori (Eds.), DNA Computing. XII, 440 pages. 2006.

Vol. 4286: P.G. Spirakis, M. Mavronicolas, S.C. Kontogiannis (Eds.), Internet and Network Economics. XI, 401 pages. 2006.

Vol. 4285: Y. Matsumoto, R.W. Sproat, K.-F. Wong, M. Zhang (Eds.), Computer Processing of Oriental Languages. XVII, 544 pages. 2006. (Sublibrary LNAI).

Vol. 4284: X. Lai, K. Chen (Eds.), Advances in Cryptology – ASIACRYPT 2006. XIV, 468 pages. 2006.

Vol. 4283: Y.Q. Shi, B. Jeon (Eds.), Digital Watermarking. XII, 474 pages. 2006.

Vol. 4282: Z. Pan, A. Cheok, M. Haller, R.W.H. Lau, H. Saito, R. Liang (Eds.), Advances in Artificial Reality and Tele-Existence. XXIII, 1347 pages. 2006.

¥581.00元

Table of Contents

23rd International Conference on Conceptual Modeling (ER 2004)

Default Clustering with Conceptual Structures 1
Julien Velcin and Jean-Gabriel Ganascia

Workshop on Context and Ontologies: Theory, Practice and Applications (C&O 2005)

Context Dependency Management in Ontology Engineering: A Formal Approach 26
Pieter De Leenheer, Aldo de Moor, and Robert Meersman

Encoding Classifications into Lightweight Ontologies 57
Fausto Giunchiglia, Maurizio Marchese, and Ilya Zaihrayeu

Second International Conference on Ontologies, DataBases and Applications of SEmantics (ODBASE 2005)

GeRoMe: A Generic Role Based Metamodel for Model Management 82
David Kensche, Christoph Quix, Mohamed Amine Chatti, and Matthias Jarke

Metadata Management in a Multiversion Data Warehouse 118
Robert Wrembel and Bartosz Bębel

International Workshop on Principles and Practice of Semantic Web Reasoning (PPSWR 2005)

SomeRDFS in the Semantic Web 158
Philippe Adjiman, François Goasdoué, and Marie-Christine Rousset

A Tool for Evaluating Ontology Alignment Strategies 182
Patrick Lambrix and He Tan

7th International Conference on Data Warehousing and Knowledge Discovery (DAWAK 2005)

Processing Sequential Patterns in Relational Databases 203
Xuequn Shang and Kai-Uwe Sattler

Author Index 219

Default Clustering with Conceptual Structures

J. Velcin and J.-G. Ganascia

LIP6, Université Paris VI
8 rue du Capitaine Scott
75015 Paris
France
{julien.velcin,jean-gabriel.ganascia}@lip6.fr

Abstract. This paper describes a theoretical framework for inducing knowledge from incomplete data sets. The general framework can be used with any formalism based on a lattice structure. It is illustrated within two formalisms: the attribute-value formalism and Sowa's conceptual graphs. The induction engine is based on a non-supervised algorithm called default clustering which uses the concept of stereotype and the new notion of default subsumption, inspired by the default logic theory. A validation using artificial data sets and an application concerning the extraction of stereotypes from newspaper articles are given at the end of the paper.

1 Introduction

This paper presents a model dealing with sparse data sets. Our original goal was to simulate common-sense inductive reasoning. It appears from previous research [1,2,3] that common-sense reasoning is highly related to reasoning from partially described data. The general framework we propose treats such data by following a default reasoning. It can be applied to automatically process heterogeneous data which often fit this property of sparseness. The main considered application deals with newspaper articles translated into a logical formalism. Our goal is to extract characteristic representations called *stereotypes* from these newspapers. The extracted representations can be seen as a way of summarizing the data in a simplified and rough manner.

More precisely, we refer to representations that can be obtained through a task of clustering. Conceptual clustering, a fundamental machine learning task [4], takes a set of object descriptions, or *observations*, as input and creates a classification scheme. Our interest will be more focused on the concepts used to name the classes than on the classes themselves. However, information sparseness is a recurrent problem in clustering. This may be for several reasons: voluntary omissions, human error, broken equipment which causes a loss of data, etc. [5]. The phenomena increases drastically when you consider the information extracted from newspapers because the vocabulary used is often very heterogeneous. Usually, existing algorithms are not adapted when there are many missing values. One solution is to fill the holes, i.e. the unknown values, by analogy with the

S. Spaccapietra et al. (Eds.): Journal on Data Semantics VIII, LNCS 4380, pp. 1–25, 2007.
© Springer-Verlag Berlin Heidelberg 2007

此为试读,需要完整PDF请访问: www.ertongbook.com