Foreword

This book constitutes the first volume of the first journal in the new LNCS Journal Subline, the Journal on Data Semantics. Publishing a journal in a book series might come as a surprise to customers, readers, and librarians, thus we would like to provide some background information and our motivation for introducing this new LNCS subline.

As a consequence of the very tight interaction between the Lecture Notes in Computer Science series and the international computer science research and development community, we receive quite a few proposals for new archive journals. From the successful launch of workshops or conferences and publication of their proceedings in the LNCS series, it might seem like a natural step to approach the publisher about launching a journal once this specific field has gained a certain level of maturity and stability. Each year we receive about a dozen such proposals and even more informal inquiries.

Like other publishers, it has been our experience that launching a new journal and making it a long-term success is a hard job nowadays, due to a generally difficult market situation, and library budget restrictions in particular. Because many of the proceedings in LNCS, and especially many of the LNCS postproceedings, apply the same strict reviewing and selection criteria as established journals, we started discussing with proposers of new journals the alternative of devoting a few volumes in LNCS to their field, instead of going through the painful Sisyphean adventure of establishing a new journal on its own. The advantages are obvious: the worldwide circulation and the international visibility of LNCS are as high as that of an average established journal, and the page price in LNCS is factors lower. Furthermore, such a journal-substitute volume would be available in LNCS Online and would thus be embedded into a comprehensive and popular digital library, in the direct neighborhood of many related conference papers.

As you will see from the frontmatter pages that follow, the LNCS Journal on Data Semantics has an Editor-in-Chief, who holds the overall mid-term editorial responsibility, and an Editorial Board consisting of excellent researchers in the area, who will devote their expertise to thoroughly reviewing the papers submitted for evaluation. In certain cases, like the present one, comparable to a special issue of a journal, the Editor-in-Chief may co-opt guest editors to jointly review the candidate papers for the volume. The Journal on Data Semantics can be subscribed to independently of the LNCS mother series, institutionally as well as individually, and, in contrast to a journal, each volume of the Journal on Data Semantics can be purchased individually at an affordable price.

Given that we have a number of pending proposals, several other LNCS journals might follow soon. We are convinced that our approach is a timely service to the international computer science research community, adding one more facet to the Lecture Notes in Computer Science series, as the most comprehensive publication platform for this community.

September 2003   Alfred Hofmann,
                                                        Springer-Verlag

LNCS Journal on Data Semantics Preface

Congratulations! What you have in your hands is the first issue of this new LNCS Journal on Data Semantics. We hope you will find much information of interest to you in the papers contained in this issue, and that you are already looking forward to the next one.

We felt that producing a new journal on this theme was most appropriate at this time, considering the evolution of computer science and practice. Computerized information handling has recently 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 infrastructures, provide reliable technical support for data distribution and data access, realizing the new popular idea that data may be available to anybody, anywhere, anytime. However, providing huge amounts of data on request often turns out to be a counterproductive service, making the data useless because of poor relevance or an 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 topics in data and knowledge management. However, the boom in semantically poor technologies, such as the Web and XML, has excited renewed interest in semantics. For instance, conferences on the Semantic Web attract crowds of participants, while ontologies on their own has become a hot topic in the database and artificial intelligence communities.

This new journal aims to provide a highly visible dissemination channel for remarkable work that in one way or another addresses research and development on issues related to data semantics. The target domain ranges from theories supporting the formal definition of semantic content to innovative domain-specific applications of semantic knowledge. We expect such a publication channel to be of 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
- Multimodel and multiparadigm approaches
- Mappings, transformations, reverse engineering
- Metadata

- Conceptual data modeling
- Integrity description and handling
- Evolution and change
- Web semantics and semistructured data
- 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 dealing with the semantics of data may be considered even if they are not covered by the topics in the list.

While the physical appearance of each journal issue looks like that of the books in the well-known Springer LNCS series, the mode of operation will be that of a journal. That is to say, publication results from contributions freely submitted by authors and reviewed by the Editorial Board. Contributions may also be invited, and nevertheless carefully reviewed, as was the case for this first issue that publishes extended versions of some of the best papers addressing data semantics topics presented at major conferences in 2002. Special issues are foreseen, focusing on specific topics under the responsibility of guest editors. Finally, it is also possible that a journal issue could be devoted to a single text.

The journal is currently aiming at one to three volumes per year.

The Editorial Board comprises one Editor-in-Chief (with overall responsibility) and a board with a number of members. The Editor-in-Chief has a four-year mandate to run the journal. Board members have three-year mandates. Mandates are renewable. More members may be added to the board as appropriate.

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

September 2003   Stefano Spaccapietra
                                                         Editor-in-Chief

LNCS Journal on Data Semantics – Editorial Board

Volume Preface

Conferences provide researchers with the fastest way to disseminate their ideas and results to a select community of other researchers in the same domain. Conferences, however, must enforce limitations in the size of the written contribution as well as in the time allocated for the on-site presentation of the contribution. They also have a limited audience, although some publishers such as Springer have a dissemination scheme that brings conference proceedings to a much wider audience than just the actual participants at the conference.

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.

To foster 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 that of the journal. We are happy that this 2003 issue provides readers with the best papers from three major 2002 conferences that addressed data semantics issues.

The selection comprises some of the most well-known conferences in this domain: the International Conference on Conceptual Modeling (Entity-Relationship (ER) Conference), the International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE), and the International Conference on Cooperative Information Systems (COOPIS). Papers from these conferences were selected based on their quality, relevancy, significance, and the viability of extending their results. All extended papers were subjected to a stringent review process and authors were required to respond to all concerns expressed by the reviewers before papers were accepted.

Three papers were selected from the 21st International Conference on Conceptual Modeling (ER 2002), held during October 7–11 in Tampere, Finland. Its Program Co-chairs, Stefano Spaccapietra, Sal March and Yahiko Kambayashi, made the selections based on recommendations of the reviewers and the presentation of the papers at the conference. Sal March managed the review process for the extended versions. The first paper, "Formal Reasoning Techniques for Goal Models" by Paolo Giorgini, John Mylopoulos, Eleonora Nicchiarelli and Roberto Sebastiani, proposes a set of axioms and algorithms used to relate software requirements, business objectives, and design qualities in information system development projects. The axioms are both qualitative and numerical, enabling the definition of precise semantics of relationships among such goals even when they are not functionally specified.

The second paper, "Attribute-Based Semantic Reconciliation of Multiple Data Sources" by Jeffrey Parsons and Yair Wand, addresses a fundamental problem faced by organizations and individuals as they seek to integrate and make

sense of data contained in independently developed data sources, that is, reconciling their semantics. They use an ontology-based approach, seeking to identify the underlying "properties" of the "things" represented in the different data sources. Such an approach can more effectively support interoperability among data sources when attributes have different structural characteristics.

In the third paper, "Data Quality in Web Information Systems," Barbara Pernici and Monica Scannapieco address the problem of managing and communicating quality characteristics of data published using the "continuously evolving and multi-faceted information medium," commonly known as the World-Wide Web. While Web-based systems are used to quickly communicate vast amounts of information to a wide audience, the quality characteristics of that information are often unclear and change as the Web site is updated. As a result visitors to the Web site are unsure about how to interpret the data. Is it complete? Is it accurate? Is it stable? This paper proposes a model and methodological framework for addressing this problem, and demonstrates its effectiveness using a prototype implementation.

Also in our selection of conferences is the newest one – as far as we know – specifically addressing data semantics. ODBASE 2002 (Ontologies, Databases, and Applications of Semantics) was held during October 29–31 in Irvine, USA, with Karl Aberer (in charge of the ODBASE selection for this issue) and Ying Liu as Program Co-chairs. The first selected paper is "Reasoning About Anonymous Resources and Meta Statements on the Semantic Web," in which Guizhen Yang and Michael Kifer address the deficiencies of current ontology languages. They provide within F-logic a precise semantics for anonymous resources and reification in F-logic. The language has a clear semantics and a proof theory, and has been implemented in the Flora-2 system. Thus this approach rectifies a number of drawbacks of the current proposal for the RDF model theory.

In "IF-Map: An Ontology-Mapping Method Based on Information-Flow Theory," Yannis Kalfoglou and Marco Schorlemmer propose the use of the information-flow-based method, which is founded on the concept of logic infomorphisms, for modeling and automating the mapping of ontologies in a distributed environment. Thus they support the automated alignment of ontologies by automatically generating mappings between a reference and various local ontologies.

York Sure, Juergen Angele and Steffen Staab present in "OntoEdit: Multifaceted Inferencing for Ontology Engineering" their work on an ontology editor that supports methodology-based ontology construction and takes advantage of inferencing capabilities. OntoEdit focuses on user-friendly support for the three main steps in ontology development, namely requirements specification, refinement and evaluation.

Finally, the third selected conference, COOPIS 2002 (Cooperative Information Systems), October 30–November 1, collocated with ODBASE, included semantics in interoperability environments as one of its major focuses. Calton Pu and Stefano Spaccapietra co-chaired the Program Committee for COOPIS 2002. The three selected papers, reviewed under the responsibility of Stefano Spac-

capietra, demonstrate the conference focus on interoperability, distribution and cooperation.

"Distributed Description Logics: Assimilating Information from Peer Sources," by Alex Borgida and Luciano Serafini, addresses issues of managing complex mappings between objects in autonomous databases. These mappings play an essential role in building global-level information services on top of existing data sets, be it in a federation or peer-to-peer perspective. The authors materialize their analysis by proposing an extension to the description logics formalism.

In "On Using Conceptual Data Modeling for Ontology Engineering," Mustapha Jarrar, Jan Demey and Robert Meersman advocate that conceptual data modeling formalisms are well suited for building ontologies. Yet they have to be complemented with a conceptual mark-up language that supports run-time exploration and the use of conceptual schemas. A second requirement, in the authors' view, is the separation of application-dependent information from the real ontological (hence application-independent) knowledge. Both requirements are covered by the proposed ontology engineering tool.

Last, but not least, "The DaQuinCIS Broker: Querying Data and Their Quality in Cooperative Information Systems," the paper by Massimo Mecella, Monica Scannapieco, Antonino Virgillito, Roberto Baldoni, Tiziana Catarci and Carlo Batini, complements the ER paper on data quality by presenting the framework they have developed for data quality management in cooperative systems. In particular, the broker component is discussed.

We believe that the set of papers in this issue addresses some of the hottest current issues in data semantics and also presents some of the outstanding running projects in this domain. We hope the readers will find material for enhancing their own research, and that they will possibly contribute to future issues of this journal.

We want to acknowledge the following colleagues who reviewed the submitted extended papers and contributed significantly with their detailed comments to improving further the quality of the published versions. Our warmest thanks to Christoph Bussler, Tiziana Catarci, Robert M. Colomb, Stefan Decker, Lois Delcambre, David W. Embley, Renato Fileto, Frank van Harmelen, Vipul Kashyap, Manolis Koubarakis, Tok Wang Ling, Bertram Ludaescher, Michele Missikoff, John Mylopoulos, Moira C. Norrie, Christine Parent, Tom Redman, Klaus-Dieter Schewe, Federica Schiappelli, Keng L. Siau, Arne Sølvberg, Rudi Studer, Vijayan Sugumaran, Xin Tan, Mohan Tanniru, Ramesh Venkataraman, Yair Wand, Richard Wang and Baowei Wei.

September 2003

Karl Aberer,
Salvatore T. March,
Stefano Spaccapietra