Preface

The papers in this volume represent the technical program of the 9th Biennial Workshop on Data Bases and Programming Languages (DBPL 2003), which was held on September 6–8, 2003, in Potsdam, Germany. The workshop meets every two years, and is a well-established forum for ideas that lie at the intersection of database and programming language research. DBPL 2003 continued the tradition of excellence initiated by its predecessors in Roscoff, Finistre (1987), Salishan, Oregon (1989), Nafplion, Argolida (1991), Manhattan, New York (1993), Gubbio, Umbria (1995), Estes Park, Colorado (1997), Kinloch Rannoch, Scotland (1999), and Frascati, Rome (2001).

The program committee selected 14 papers out of 22 submissions, and invited two contributions. The 16 talks were presented over three days, in seven sessions.

In the **invited talk** Jennifer Widom presented the paper *CQL: a Language for Continuous Queries over Streams and Relations*, coauthored by Arvind Arasu and Shivnath Babu. While a lot of research has been done recently on query processing over data streams, CQL is virtually the first proposal of a query language on streams that is a strict extension of SQL. The language is structured around a simple yet powerful idea: it has two distinct data types, relations and streams, with well-defined operators for mapping between them. Window specification expressions, such as sliding windows, map streams to relations, while operators such as "insert stream," "delete stream," and "relation stream" map relations to streams by returning, at each moment in time, the newly inserted tuples, the deleted tuples, or a snapshot of the entire relation. The numerous examples in this paper make a convincing case for the power and usefulness of CQL.

The **invited tutorial** was presented by Georg Gottlob and Christoph Koch on *XPath Query Processing*. They described why existing XPath processors run in exponential time in the size of the XPath expression, and showed simple examples of XPath expressions and XML documents on which several popular XPath implementations took unacceptably long time to execute, or even failed to terminate. They described then a class of algorithms based on dynamic programming for evaluating XPath expressions in PTIME, and illustrated several optimization techniques that can be applied to these algorithms.

The three papers of the session **Static Analysis** contain some surprising results on query languages. The first, *Satisfiability of XPath Expressions*, by Jan Hidders, shows that XPath expressions that include parent/ancestor axes, union, intersection, and difference, may not always be satisfiable, i.e., there are expressions that return the empty result on any XML document. The paper analyzes when an XPath expression is satisfiable, and shows that for various combinations of these operators this problem ranges from NP-complete to PTIME. The second paper, *Containment of Relational Queries with Annotation Propagation*, by Wang-Chiew Tan, studies the containment and equivalence problems of relational queries that carry along annotations from source data. The paper shows

that certain relational queries that are equivalent under the traditional semantics, become in-equivalent when one takes into account the annotations they carry. As a consequence, one needs to rethink relational query optimization if the queries are expected to return the annotations in the database. The third paper, *Avoiding Unnecessary Ordering Operations in XPath*, by Jan Hidders and Philippe Michiels, notices that often XQuery optimizers introduce unnecessary sort operators to restore the document order of the elements. The paper describes an elegant algorithm for detecting redundant sort operators, which then an optimizer can remove.

In the session **Transactions** three papers addressed various issues related to the integration of transactional semantics in persistent programming languages. *Consistency of Java Transactions*, by Suad Alagic and Jeremy Logan, proposes a model of Java transactions, based on a subtle interplay of constraints, bounded parametric polymorphism, and orthogonal persistence. The paper *Integrating Database and Programming Languages Constraints*, by Oded Shmueli, Mukund Raghavachari, Vivek Sarkar, Rajesh Bordawekar, and Michael Burke, addresses the problem of data consistency in applications that interact with databases. It describes an architecture that automatically inserts checks in the application that ensure data consistency. The third paper, *A Unifying Semantics for Active Databases Using Non-Markovian Theories of Actions*, by Iluju Kiringa and Ray Reiter, develops a new form of theories for modeling active behavior in databases, such as active databases. Their formalism is based on the situation calculus and non-Markovian control.

The session **Modeling Data and Services** contained three papers focusing on Web services, XML constraints, and persistent objects. In *Modeling Dynamic Web Data*, Philippa Gardner and Sergio Maffeis introduced a calculus for describing the dynamic behavior of Web data. The calculus combines semistructured data with an extension of the π-calculus, and can be used to reason about behavior found, for example, in dynamic Web page programming, applet interaction, and service orchestration. In *Semantics of Objectified XML Constraints*, Suad Alagic and David Briggs developed a model theory for a functional object-oriented data model extended with XML-like types, based on Featherweight Java. Finally, the third paper, *M^2ORM^2: a Model for the Transparent Management of Relationally Persistent Objects*, by Luca Cabibbo and Roberto Procelli, describes a "meet in the middle" approach for mapping object-oriented application objects to relational databases.

The session **Novel Applications of XML and XQuery** contained two papers describing some novel applications of the newly standardized XML query language. In *Using XQuery for Flat-File Based Scientific Datasets* the authors, Xiaogang Li and Gagan Agrawal, showed that XQuery is suitable for scientific applications, which were traditionally solved in FORTRAN. The paper describes an application of XQuery to satellite data processing, and proposes a new class of optimization techniques for XQuery to better support these types of applications. The paper *A Query Algebra for Fragmented XML Stream Data*, by Sujoe Bose, Leonidas Fegaras, David Levine, and Vamsi Chaluvadi, describes an application

of XQuery for broadcast XML stream data processing. The authors propose an extension of the XML data model with *holes* and *fillers*, and show how XQuery can be adapted to deal with them.

Finally, three papers formed the session **XML Processing and Validation**. The first, *Updates and Incremental Validation of XML Documents*, by Beatrice Bouchou and Mirian Halfeld Ferrari Alves, describes a practical algorithm for incrementally checking the validity of an XML file against a DTD. Starting from a valid XML document, the algorithm checks the validity of a document obtained by inserting, deleting, or modifying an element (subtree) of the document. In *Attribute Grammars for Scalable Query Processing on XML Streams*, the authors, Christoph Koch and Stefanie Scherzinger, described a simple formalism based on attribute grammars that can be used to specify transformations on XML streams. In the third paper, *A General Framework for Estimating XML Query Cardinality*, Carlo Sartiani addressed one of the most difficult problems in XML query optimization, i.e., cardinality estimation, and described a comprehensive approach that applies to the entire XQuery language.

DBPL 2003 was hosted and sponsored by the Hasso Plattner Institute for Software Systems Engineering at the University of Potsdam. We gratefully acknowledge this valuable support. We also thank Mathias Weske and his team for the perfect organization and all the local arrangements.

November 2003 Georg Lausen, Dan Suciu

http://www.springer.com/978-3-540-20896-9

Database Programming Languages
9th International Workshop, DBPL 2003, Potsdam, Germany, September 6-8, 2003, Revised Papers
Lausen, G.; Suciu, D. (Eds.)
2004, X, 286 p., Softcover
ISBN: 978-3-540-20896-9