

Lecture Notes in Computer Science
Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

2397

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Tokyo

Giorgio Ghelli Gösta Grahné (Eds.)

Database Programming Languages

8th International Workshop, DBPL 2001
Frascati, Italy, September 8-10, 2001
Revised Papers



Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Giorgio Ghelli
Univesità di Pisa, Dipartimento di Informatica
Corso Italia 40, 56125 Pisa, Italy
E-mail: ghelli@di.unipi.it

Gösta Grahne
Concordia University, Department of Computer Science
1455 De Maisonneuve Blvd. West
Montréal, Québec, Canada H3G 1M8
E-mail: grahne@cs.concordia.ca

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Database programming languages : 8th international workshop ; revised papers / DBPL 2001, Frascati, Italy, September 8 - 10, 2001. Gösta Grahne ; Giorgio Ghelli (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2002
(Lecture notes in computer science ; 2397)
ISBN 3-540-44080-1

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

ISSN 0302-9743

ISBN 3-540-44080-1 Springer-Verlag 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-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York,
a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2002
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Olgun Computergrafik
Printed on acid-free paper SPIN: 10873560 06/3142 5 4 3 2 1 0

Preface

The papers in this volume represent the technical program of the 8th Biennial Workshop on Data Bases and Programming Languages (DBPL 2001), that was held during September 8–10, 2001, in Frascati, located on the beautiful hills surrounding Rome, in an area favored by the ancient Roman patricians who built their summer residences there. DBPL 2001 continued the tradition of excellence initiated by its predecessors in Roscoff, Finistère (1987), Salishan, Oregon (1989), Nafplion, Argolida (1991), Manhattan, New York (1993), Gubbio, Umbria (1995), Estes Park, Colorado (1997), and Kinloch Rannoch, Scotland (1999).

Databases grew out of a separation between physical and logical data, thus enabling high-level query languages. Database query languages have evolved in expressive power and structural capabilities. Programming languages have seen a development from assembly languages to high-level declarative paradigms. Thus the two areas approach each other as they mature. Earlier successful cross-fertilizations between the fields include the combination of relational theory, type theory and object-oriented languages, resulting in object-oriented databases, object-relational databases and persistent programming languages. The combination of database logic programming and constraint programming produced deductive and constraint databases. Recently, with the emergence of semi-structured data models, there is a renewed synergy between databases and programming languages, in particular in the design of languages to manipulate XML data.

The DBPL 2001 Program Co-Chairs were Giorgio Ghelli (Pisa) and Gösta Grahne (Montréal). The Program Committee Members were Catriel Beerli (Jerusalem), Diego Calvanese (Rome), Richard Connor (Glasgow), Alon Halevy (Seattle), Leonid Libkin (Toronto), Gianni Mecca (Potenza), Frank Neven (Limburg), Benjamin Pierce (Philadelphia), Chris Ramming (Menlo Park), Jérôme Siméon (Murray Hill), Victor Vianu (San Diego), and Philip Wadler (Basking Ridge). We warmly thank the Program Committee members for devoting their time and effort generously to DBPL 2001.

In response to the call for papers, the Program Committee received 36 submissions. The submission process and the Committee's extensive deliberations were carried out using the Confman Software (<http://confman.unik.no/confman>). Technical support during this process was generously provided by Jianfei Zhu from Concordia University. The Program Committee accepted 18 papers for presentation at the workshop. The papers were grouped into technical sessions on Semistructured Data, OLAP and Datamining, XML, Spatial Databases, Systems, Schema Integration, Index Concurrency, User Languages, and Rules. Here we find both traditional DBPL topics, as well as the new thrust of Semistructured Data and XML, a theme that made its first appearance in the DBPL series in 1997.

DBPL 2001 also featured an invited talk by Tony Bonner (Toronto) on “Languages for workflows and e-commerce,” and an invited tutorial by Dan Suciu (Seattle) on “Type-checking in semi-structured databases,” in addition to the technical program. Dr. Bonner’s talk gave a survey of the development of declarative languages for the transaction-oriented applications of workflows and e-commerce. These applications represent new functionalities required for high-level database languages. Dr. Suciu’s invited tutorial gave a background in type checking in programming languages, and it explained very recent advances in type-checking XML and semi-structured databases.

The main *raison d’être* for DBPL has always been to overcome the so-called impedance mismatch between databases and programming languages. Although the term impedance mismatch was coined it’s already some time since, as we heard Bill Gates emphasize during his keynote talk at ACM SIGMOD/PODS in 1998, it still remains a major challenge for the Database and Programming Language communities. We believe that the research carried out by these communities has created a body of knowledge that can be tapped in facing the challenge of providing the technology for high-level expressive and yet optimizable database programming languages in the context of XML and semi-structured data.

DBPL 2001 was partially hosted and sponsored by the local establishment of the European Space Agency ESA (<http://www.esa.int>), the European Space Research Institute ESRIN (<http://www.esa.int/esrin>), and by the Esprit Working Group AppSem (Applied Semantics, <http://www.md.chalmers.se/Cs/Research/Semantics/APPSEM/>). On behalf of the DBPL community we gratefully acknowledge their valuable support. Many thanks are also due to Ettore Ricciardi of the National Research Council of Pisa, for the impeccable local arrangements.

During the workshop, only mimeographed informal proceedings were distributed. The papers in this volume were revised by the authors based on the comments from the refereeing stage and ensuing discussions during the workshop, and were subjected to a final acceptance by the Program Committee. We are very grateful to Paolo Manghi, from the University of Pisa, who functioned as Production Editor for the final compilation of the material.

The 9th Biennial Workshop on Data Bases and Programming Languages (DBPL 2003) will be held in Berlin in September 2003, and will be Co-Chaired by Georg Lausen (Freiburg) and Dan Suciu (Seattle).

Organization

DBPL 2001 was organized by the Dipartimento di Informatica, University of Pisa (Italy), and the Department of Computer Science, Concordia University (Canada).

Conference Chairs

Giorgio Ghelli (University of Pisa, Italy)
Gösta Grahne (Concordia University, Canada)

Program Committee

Catriel Beeri (Jerusalem)	Frank Neven (Limburg)
Diego Calvanese (Rome)	Benjamin Pierce (Philadelphia)
Richard Connor (Glasgow)	Chris Ramming (Menlo Park)
Alon Halevy (Seattle)	Jérôme Siméon (Murray Hill)
Leonid Libkin (Toronto)	Victor Vianu (San Diego)
Gianni Mecca (Potenza)	Philip Wadler (Basking Ridge)

External Referees

Marcelo Arenas	Giuseppe De Giacomo	Alberto Mendelzon
Enrico Franconi	Lee Chin Soon	Pratik Mukhopadhyay
Tony Bonner	Val Tannen	Luigi Dragone
Zack Ives	Dan Suciu	Matthew Fuchs
Alfons Kemper	Peter Buneman	Sanjeev Khanna
Paolo Atzeni	Arnaud Sahuguet	Haruo Hasoya
Alin Deutsch	Peter Fankhauser	Vladimir Gapeyev
Renzo Orsini	Wang-Chiew Tan	Francesco Donini
Gottfried Vossen	Daniele Nardi	Francesco Scarcello
Riccardo Rosati	Paolo Merialdo	Giovanbattista Ianni
Riccardo Torlone	Nils Klarlund	Peter Sewell
Luca Cabibbo	Dana Florescu	Giuseppe Santucci
Rick Snodgrass	Keishi Tajima	
Mary Fernandez	Danilo Montesi	

Proceedings Production Editor

Paolo Manghi (University of Pisa, Italy)

VIII Organization

Confman Webmaster

Jianfei Zhu (Concordia University, Canada)

Local Arrangements

Ettore Ricciardi (National Research Council of Pisa, Italy)

Sponsoring Institutions

European Space Agency (ESA)

European Space Research Institute (ESRIN)

Esprit Working Group AppSem

Table of Contents

Invited Contribution

Typechecking for Semistructured Data	1
<i>Dan Suciu (University of Washington)</i>	

Semistructured Data

Optimization Properties for Classes of Conjunctive Regular Path Queries	21
<i>Alin Deutsch, Val Tannen (University of Pennsylvania)</i>	
View-Based Query Answering and Query Containment over Semistructured Data	40
<i>Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini (Università di Roma), Moshe Y. Vardi (Rice University)</i>	
Model-Checking Based Data Retrieval	62
<i>Agostino Dovier (Università di Udine), Elisa Quintarelli (Politecnico di Milano)</i>	

OLAP and Data Mining

A Temporal Query Language for OLAP: Implementation and a Case Study	78
<i>Alejandro A. Vaisman (Universidad de Buenos Aires), Alberto O. Mendelzon (University of Toronto)</i>	
Attribute Metadata for Relational OLAP and Data Mining	97
<i>T.H. Merrett (McGill University)</i>	
On Monotone Data Mining Languages	119
<i>Toon Calders (University of Antwerp), Jef Wijsen (University of Mons-Hainaut)</i>	

XML

Reasoning about Keys for XML	133
<i>Peter Buneman, Susan Davidson (University of Pennsylvania), Wenfei Fan (Bell Laboratories), Carmem Hara (Universidade Federal do Parana), Wang-Chiew Tan (University of Pennsylvania)</i>	
TAX: A Tree Algebra for XML	149
<i>H.V. Jagadish (University of Michigan), Laks V.S. Lakshmanan (University of British Columbia), Divesh Srivastava (AT&T Labs-Research), Keith Thompson (University of Michigan)</i>	

A Rule-Based Querying and Updating Language for XML 165
Wolfgang May (Universität Freiburg)

Spatial Databases

Linear Approximation of Semi-algebraic Spatial Databases
Using Transitive Closure Logic, in Arbitrary Dimension 182
Floris Geerts (University of Limburg)

A Theory of Spatio-temporal Database Queries 198
Floris Geerts, Sofie Haesevoets, Bart Kuijpers (University of Limburg)

Systems, Schema Integration, Index Concurrency

An Application-Specific Database 213
*Kathleen Fisher, Colin Goodall, Karin Högstedt, Anne Rogers
(AT&T Labs)*

A Model Theory for Generic Schema Management 228
Suad Alagić, Philip A. Bernstein (Microsoft Research)

View Serializable Updates of Concurrent Index Structures 247
*Stavros Cosmadakis (University of Patras), Kleoni Ioannidou
(University of Toronto), Stergios Stergiou (University of Athens)*

User Languages

SQL4X: A Flexible Query Language for XML and Relational Databases . . 263
Sara Cohen, Yaron Kanza, Yehoshua Sagiv (The Hebrew University)

ERX-QL: Querying an Entity-Relationship DB
to Obtain XML Documents 281
Giuseppe Psaila (Università degli Studi di Bergamo)

Rules

Optimising Active Database Rules by Partial Evaluation
and Abstract Interpretation 300
*James Bailey (University of Melbourne),
Alexandra Poulouvasilis (University of London),
Simon Courtenage (University of Westminster)*

Simulation of Advanced Transaction Models Using GOLOG 318
Iluju Kiringa (University of Toronto)

Author Index 343