



HAL
open science

Database and Xml Technologies

Mong-Li Lee, Jeffrey Xu Yu, Zohra Bellahsene, Rainer Unland

► **To cite this version:**

Mong-Li Lee, Jeffrey Xu Yu, Zohra Bellahsene, Rainer Unland. Database and Xml Technologies. Lecture Notes in Mathematics, LNCS (6309), 2010, 978-3-642-15683-0. 10.1007/978-3-642-15684-7. lirmm-00547552

HAL Id: lirmm-00547552

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00547552v1>

Submitted on 16 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Alfred Kobsa

University of California, Irvine, CA, 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

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Mong Li Lee Jeffrey Xu Yu
Zohra Bellahsene Rainer Unland (Eds.)

Database and XML Technologies

7th International
XML Database Symposium, XSym 2010
Singapore, September 17, 2010
Proceedings

 Springer

Volume Editors

Mong Li Lee
National University of Singapore
School of Computing
Singapore 117417, Republic of Singapore
E-mail: leeml@comp.nus.edu.sg

Jeffrey Xu Yu
The Chinese University of Hong Kong
Department of Systems Engineering and Engineering Management
Shatin, N.T., Hong Kong
E-mail: yu@se.cuhk.edu.hk

Zohra Bellahsène
Université Montpellier II
LIRMM UMR 5506 CNRS
34392 Montpellier, France
E-mail: bella@lirmm.fr

Rainer Unland
University of Duisburg-Essen
Institute for Computer Science and Business Information Systems (ICB)
45117 Essen, Germany
E-mail: rainer.unland@icb.uni-due.de

Library of Congress Control Number: 2010933595

CR Subject Classification (1998): H.2, H.3, E.1, H.2.8, H.4, F.2

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

ISSN 0302-9743
ISBN-10 3-642-15683-5 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-15683-0 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.com

© Springer-Verlag Berlin Heidelberg 2010
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper 06/3180

Preface

Since its first edition in 2003, the XML Database Symposium series (XSym) has been a forum for academics, practitioners, users and vendors, allowing all to discuss the use of and synergy between database management systems and XML. The symposia have provided many opportunities for timely discussions on a broad range of topics pertaining to the theory and practice of XML data management and its applications. XSym 2010 continued this XSym tradition with a program consisting of 11 papers and a keynote shared with the 36th International Conference on Very Large Data Bases (VLDB 2010). We received 20 paper submissions, out of which 8 papers were accepted as full papers, and 3 as short papers. Each submitted paper underwent a rigorous and careful review by four referees.

The contributions in these proceedings are a fine sample of the current research in XML query processing, including XPath satisfiability, approximate joins, pattern matching, linear index construction for trees, dynamic labeling, and XQuery update translation based on schema. The papers focus on recent advances in detecting functional dependencies, modeling complex XML twig pattern output, promoting semantics capability of XML keys, and searchable compression of Microsoft office documents. In addition, we include a paper that shares lessons learned from real XML database development.

The organizers would like to express their gratitude to the authors, for submitting their work, and to the Program Committee, for providing very thorough evaluations of the submitted papers and for the discussions that followed under significant time constraints. We also would like to thank the invited keynote speaker, Prof. M. Tamer Özsu, for the challenging and thought-provoking contribution. Finally, we are also grateful to Microsoft and Michael Rys for their generous sponsorship, Andrei Voronkov and other contributors for the EasyChair conference management system, and the local organizers for their efforts in making XSym 2010 a pleasant and successful event. Finally, we would also like to thank Alfred Hofmann and his great team from Springer for their support and cooperation in putting this volume together.

July 2010

Mong Li Lee
Jeffrey Xu Yu
Zohra Bellahsene
Rainer Unland

Organization

Steering Committee

| | |
|------------------|--|
| Zohra Bellahsene | LIRMM-CNRS/University Montpellier 2 (France) |
| Ela Hunt | University of Strathclyde (UK) |
| Michael Rys | Microsoft (USA) |
| Rainer Unland | University of Duisburg-Essen (Germany) |

Program Co-chairs

| | |
|---------------|--|
| Mong Li Lee | National University of Singapore (Singapore) |
| Jeffrey Xu Yu | Chinese University of Hong Kong (China) |

International Program Committee

| | |
|--------------------|--|
| Bernd Amann | Université Paris 6 (France) |
| Veronique Benzaken | Université Paris-Sud (France) |
| Sourav S. Bhowmick | Nanyang Technological University (Singapore) |
| Stéphane Bressan | National University of Singapore (Singapore) |
| Chee Yong Chan | National University of Singapore (Singapore) |
| Yi Chen | Arizona State University (USA) |
| Minos Garofalakis | Technical University of Crete (Greece) |
| Giorgio Ghelli | Università di Pisa (Italy) |
| Torsten Grust | Universität Tübingen (Germany) |
| Giovanna Guerrini | Università di Genova (Italy) |
| H.V. Jagadish | University of Michigan (USA) |
| Yaron Kanza | Technion Israel Institute of Technology (Israel) |
| Raghav Kaushik | Microsoft Research (USA) |
| Jiaheng Lu | Renmin University of China (China) |
| Murali Mani | Worcester Polytechnic Institute (USA) |
| Peter McBrien | Imperial College - London (UK) |
| Atsuyuki Morishima | University of Tsukuba (Japan) |
| Tadeusz Pankowski | Poznan University of Technology (Poland) |
| Prakash Ramanan | Wichita State University (USA) |
| Pierre Senellart | Télécom ParisTech (France) |
| Jerome Simeon | IBM Research (USA) |
| Martin Theobald | Max-Planck-Institut für Informatik (Germany) |
| Wee Hyong Tok | Microsoft (China) |

VIII Organization

Vasilis Vassalos

Yuqing Wu

Ni Yuan

Xiaofang Zhou

Athens University of Economics and Business (Greece)

Indiana University (USA)

IBM (China)

University of Queensland (Australia)

External Reviewers

Pantelis Aravogliadis

Federico Cavalieri

Vassilis Christophides

Pierre Genevhès

Françoise Gire

Mirian Halfeld-Ferrari.

Table of Contents

Keynote Address

| | |
|--|---|
| Distributed XML Query Processing (Extended Abstract) | 1 |
| <i>M. Tamer Özsu and Patrick Kling</i> | |

XML Query Processing

| | |
|---|----|
| Approximate Joins for XML Using g-String | 3 |
| <i>Fei Li, Hongzhi Wang, Cheng Zhang, Liang Hao, Jianzhong Li, and Hong Gao</i> | |
| Linear Computation of the Maximum Simultaneous Forward and Backward Bisimulation for Node-Labeled Trees | 18 |
| <i>Nils Grimsmo, Truls Amundsen Bjørklund, and Magnus Lie Hetland</i> | |
| Extending the Tractability Results on XPath Satisfiability with Sibling Axes | 33 |
| <i>Yasunori Ishihara, Shogo Shimizu, and Toru Fujiwara</i> | |
| Extending XQuery with a Pattern Matching Facility | 48 |
| <i>Peter M. Fischer, Aayush Garg, and Kyumars Sheykh Esmaili</i> | |

XML Update and Applications

| | |
|---|-----|
| A Schema-Based Translation of XQuery Updates | 58 |
| <i>Leonidas Fegaras</i> | |
| EBSL: Supporting Deleted Node Label Reuse in XML | 73 |
| <i>Martin F. O'Connor and Mark Roantree</i> | |
| Lessons Learned from DB2 pureXML Applications: A Practitioner's Perspective | 88 |
| <i>Matthias Nicola</i> | |
| Searchable Compression of Office Documents by XML Schema Subtraction | 103 |
| <i>Stefan Böttcher, Rita Hartel, and Christian Messinger</i> | |

XML Modeling

| | |
|---|-----|
| Fast Detection of Functional Dependencies in XML Data | 113 |
| <i>Hang Shi, Toshiyuki Amagasa, and Hiroyuki Kitagawa</i> | |

| | |
|---|-----|
| TP+Output: Modeling Complex Output Information in XML Twig Pattern Query | 128 |
| <i>Huayu Wu, Tok Wang Ling, and Gillian Dobbie</i> | |
| Promoting the Semantic Capability of XML Keys | 144 |
| <i>Flavio Ferrarotti, Sven Hartmann, Sebastian Link, and Jing Wang</i> | |
| Author Index | 155 |