Advances in Database Technology – EDBT 2010

13th International Conference
on Extending Database Technology
Lausanne, Switzerland, March 22–26, 2010
Proceedings

Editors:
Ioana Manolescu (INRIA, France)
Stefano Spaccapietra (EPFL, Switzerland)
Jens Teubner (ETH Zurich, Switzerland)
Masaru Kitsuregawa (Tokyo University, Japan)
Alain Léger (Orange – France Telecom R&D, France)
Felix Naumann (Hasso Plattner Institute, Germany)
Anastasia Ailamaki (EPFL, Switzerland)
Fatma Özcan (IBM Almaden Research Center, USA)

Advances in Database Technology – EDBT 2010
Proceedings of the 13th International Conference
on Extending Database Technology
Lausanne, Switzerland, March 22–26, 2010

Editors:
Ioana Manolescu
Stefano Spaccapietra
Jens Teubner
Masaru Kitsuregawa
Alain Léger
Felix Naumann
Anastasia Ailamaki
Fatma Özcan

The Association for Computing Machinery
2 Penn Plaza, Suite 701
New York, NY, 10121-0701

ACM ISBN: 978-1-60558-945-9

Table of Contents

Data Provenance

Probabilistic and Spatial Databases

Query Processing and Optimization 2

Spatial Databases

Technologies for the Web

Industrial Sessions

Transactions and Distribution

New Applications

Data Warehousing and Analytics

Demonstrations

Demonstrations

Tutorials

Foreword

The papers in this volume were presented at the 13th International Conference on Extending Database Technology (EDBT 2010), held in Lausanne, Switzerland, March 22–26, 2010. Starting last year, EDBT is held jointly with the ICDT (International Conference on Database Theory) conference.

This year, the conference research track received a total of 307 submissions, out of which 54 were accepted. This amounts to an acceptance rate of about 18 %, confirming the reputation of the EDBT conference as being quite selective! The program committee consisted of 75 members, leading to 12–13 articles to be reviewed by each PC member. While such load is not per se inconsistent with previous editions of EDBT or other prestigious conferences such as, *e.g.*, SIGMOD or VLDB, some reviewers found this quite a heavy load, and significant care had to be taken to ensure appropriate external reviewers or load re-balancing. The submissions were thoroughly discussed, typically as soon as three reviews were available for the paper, even if this was before the review submission deadline; overall, 444 comments were exchanged through the electronic submission tool (EasyChair). The chair is particularly grateful to the PC members which have taken the time to carefully consider papers and debate each other's arguments. Given the high number of interesting submissions, the PC has worked with the chair to select the best submissions, and did not consider issues such as the overall balance of various database research areas in the conference program. Thus, the proceedings you are holding now reflect the best and most interesting works submitted to our conference, and you can consider the way they split over various topics as a snapshot of the state of our research community as of September 2009.

The conference program is beautifully complemented by the other tracks. Ian Horrocks and Val Tannen have gracefully accepted our invitation to deliver keynote talks at the conference. Alain Léger and Masaru Kitsuregawa have put together an exciting industrial program. Fatma Özcan has chaired the demonstration track, while Felix Naumann has organized the selection of two interesting tutorials. Natassa Ailamaki has helped add to the conference a set of workshops on timely issues. The chairs all join together in acknowledging the careful work of their program committees and external reviewers, and thanking them for their effort.

Standing quite proud of our selection, we hope you will enjoy the program as much as we enjoyed putting it together, and are looking forward to an exciting conference in Lausanne!

Ioana Manolescu (Program Chair)
Stefano Spaccapietra (General Chair)
Jens Teubner (Proceedings Chair)
Masaru Kitsuregawa (Industr. & Appl. Co-Chair)
Alain Léger (Industr. & Appl. Co-Chair)
Felix Naumann (Tutorial Chair)
Anastasia Ailamaki (Workshop Chair)
Fatma Özcan (Demonstrations Chair)

Program Committee Members

Research

Amr El Abbadi
Gustavo Alonso
Denilson Barbosa
Klemens Boehm
Angela Bonifati
Philippe Bonnet
Luc Bouganim
Loreto Bravo
Luca Cabibbo
Diego Calvanese
Alessandro Campi
Malu Castellanos
Dario Colazzo
Emiran Curtmola
AnHai Doan
Anne Doucet
Cedric DuMouza
Piero Fraternali
Christine Froidevaux
Irini Fundulaki
Ariel Fuxman
Helena Galhardas
Floris Geerts
Aristides Gionis
Lukasz Golab
Daniela Grigori
Torsten Grust
Vagelis Hristidis
Wynne Hsu
Ihab Ilyas
Yaron Kanza
Alfons Kemper
Eamonn Keogh
Martin Kersten
Alexandros Labrinidis
Philippe Lamarre
Paul Larson
Bertram Ludascher
Stefan Manegold
Ioana Manolescu (chair)
Amélie Marian
Volker Markl
Giansalvatore Mecca
Paolo Missier
Mirella Moro
Wolfgang Nejdl
Raymond Ng
Matthias Nicola
Boris Novikov
Nicola Onose
Michalis Petropoulos
Neoklis Polyzotis
Jaroslav Pokorny
Ravishankar Ramamurthy
Philippe Rigaux
Tore Risch
Domenico Sacca
Mohand Said-Hacid
Pierangela Samarati
Thomas Seidl
Kyuseok Shim
Altigran Da Silva
Divesh Srivastava
Dan Suciu
Jens Teubner
Peter Triantafillou
Patrick Valduriez
Maurice Van Keulen
Vassilis Vassalos
Michalis Vazirgiannis
Jef Wijsen
Haruo Yokota
Masatoshi Yoshikawa
Jeffrey Yu Xu
Aoying Zhou

Industrial & Applications

Soren Auer
Francois Bry
Orri Erling
Fabien Gandon
German Herrero
Ruben Lara Hernandez
Masaru Kitsuregawa (chair)
Alain Léger (chair)
Wen-Syan Li
Anne Monceaux
Miyuki Nakano
Massimo Paolucci
Dimitris Plexousakis
Steffen Saab
Carsten Saathoff
Tomas Vitvar
Alexander Wahler
Haofen Wang
Kun-Lung Wu
Yiannis Kompatsiaris

Demonstrations

Laurent Amsaleg
Christof Bornhoevd
Bjorn Tor Jonsson
Hakan Hacigumus
Ohad Greenshpan
Kristen LeFevre
Qiong Luo
Fatma Özcan (chair)
Stelios Paparizos
Evaggelia Pitoura
Pinar Senkul
Alkis Simitsis

David Simmen
Nesime Tatbul
Zografoula Vagena
Yannis Velegrakis
Stratis Viglas
Cong Yu

Provenance for Database Transformations

Val Tannen
University of Pennsylvania, USA
val@cis.upenn.edu

ABSTRACT

Database transformations (queries, views, mappings) take apart, filter, and recombine source data in order to populate warehouses, materialize views, and provide inputs to analysis tools. As they do so, applications often need to track the relationship between parts and pieces of the sources and parts and pieces of the transformations' output. This relationship is what we call database provenance.

This talk presents an approach to database provenance that is based on two observations. First, provenance is a kind of annotation, and we can develop a general approach to annotation propagation that also covers other applications, for example to uncertainty and access control. In fact, provenance turns out to be the most general kind of such annotation, in a precise and practically useful sense. Second, the propagation of annotation through a broad class of transformations relies on just two operations: one when annotations are jointly used and one when they are used alternatively. This leads to annotations forming a specific algebraic structure, a commutative semiring.

The semiring approach works for annotating tuples, field values and attributes in standard relations, in nested relations (complex values), and for annotating nodes in (unordered) XML. It works for transformations expressed in the positive fragment of relational algebra, nested relational calculus, unordered XQuery, as well as for Datalog, GLAV schema mappings, and tgd constraints. Specific semirings correspond to earlier approaches to provenance, while others correspond to forms of uncertainty, trust, cost, and access control.

This is joint work with J.N. Foster, T.J. Green, Z. Ives, and G. Karvounarakis, done in part within the frameworks of the Orchestra and pPOD projects.

EDBT 2010, March 22–26, 2010, Lausanne, Switzerland.

Scalable Ontology-Based Information Systems

Ian Horrocks
University of Oxford, UK
ian.horrocks@comlab.ox.ac.uk

ABSTRACT

Ontologies and ontology based systems are becoming increasingly important in meeting the demand for more powerful and flexible information systems. Requirements for such systems include the need to deal with incomplete and semi-structured information, to integrate information from heterogeneous sources, to employ richer and more flexible schemas, and for query answers to reflect both knowledge and data. Provision of such enhanced capabilities must, however, be in addition to, and not instead of, the well-established features of existing database systems, in particular their robust scalability. Achieving this is, of course, extremely challenging. In this talk I will present some recent research efforts that tackle this problem, including investigations of tractable fragments, new algorithmic techniques, new optimisations and the exploitation of relational database technology.

EDBT 2010, March 22–26, 2010, Lausanne, Switzerland.