

Database Theory – ICDT 2010

13th International Conference
on Database Theory
Lausanne, Switzerland, March 23–25, 2010
Proceedings

Editor:
Luc Segoufin (INRIA & ENS Chacan, France)

Database Theory – ICDT 2010
Proceedings of the 13th International Conference on Database Theory
Lausanne, Switzerland, March 23–25, 2010

Editor:
Luc Segoufin

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

ACM COPYRIGHT NOTICE. Copyright © 2009 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Publications Dept., ACM, Inc., fax +1 (212) 869-0481, or permissions@acm.org.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, +1-978-750-8400, +1-978-750-4470 (fax).

ACM ISBN: 978-1-60558-947-3

Table of Contents

Foreword	v
Program Committee Members and External Reviewers	vii

Invited Papers

Increasing Representational Power and Scaling Reasoning in Probabilistic Databases <i>Amol Deshpande</i>	1
From Polynomial Time Queries to Graph Structure Theory <i>Martin Grohe</i>	2
Informative Labeling Schemes <i>Pierre Fraigniaud</i>	3

Research Sessions

Optimization, Indexing, and Transactions

Foundations of SPARQL Query Optimization <i>Michael Schmidt, Michael Meier, and Georg Lausen</i>	4–33
Efficient Processing of 3-Sided Range Queries with Probabilistic Guarantees <i>Alexis Kaporis, Apostolos Papadopoulos, Spyros Sioutas, Kostas Tsichlas, and Konstantinos Tsakalidis</i>	34–43
A Theoretical Study of "Snapshot Isolation" <i>Ragnar Normann and Lene T. Østby</i>	44–49

Probabilistic Databases

Aggregate Queries for Discrete and Continuous Probabilistic XML <i>Serge Abiteboul, T.-H. Hubert Chan, Evgeny Kharlamov, Werner Nutt, and Pierre Senellart</i>	50–61
Querying Parse Trees of Stochastic Context-Free Grammars <i>Sara Cohen and Benny Kimelfeld</i>	62–75
Probabilistic Data Exchange <i>Ronald Fagin, Benny Kimelfeld, and Phokion Kolaitis</i>	76–88

Views, Inclusion, and Equivalence

Synthesizing View Definitions from Data <i>Anish Das Sarma, Aditya Parameswaran, Hector Garcia-Molina, and Jennifer Widom</i>	89–103
K-Jump Strategy for Preserving Privacy in Micro-Data Disclosure <i>Lingyu Wang, Lei Zhang, and Wen Ming Liu</i>	104–115

Bag Equivalence of XPath Queries <i>Sara Cohen and Yaacov Y. Weiss</i>	116–128
Data Exchange 1	
Composition with Target Constraints <i>Marcelo Arenas, Ronald Fagin, and Alan Nash</i>	129–142
Answering Non-Monotonic Queries in Relational Data Exchange <i>André Hernich</i>	143–154
On the Tradeoff between Mapping and Querying Power in XML Data Exchange <i>Shunichi Amano, Claire David, Leonid Libkin, and Filip Murlak</i>	155–164
Logic and XPath	
The Complexity of Rooted Phylogeny Problems <i>Manuel Bodirsky and Jens K. Mueller</i>	165–173
A Greedy Algorithm for Constructing a Low-Width Generalized Hypertree Decomposition <i>Kaoru Katayama, Tatsuro Okawara, and Yuka Ito</i>	174–181
Data Exchange 2	
Static Analysis of Schema-Mappings Ensuring Oblivious Termination <i>Bruno Marnette and Floris Geerts</i>	182–194
Mapping Polymorphism <i>Ryan Wisnesky, Mauricio Hernandez, and Lucian Popa</i>	195–207
Composing Local-As-View Mappings: Closure and Applications <i>Patricia C. Arocena, Ariel Fuxman, and Renée J. Miller</i>	208–217
Session 7	
Data Correspondence, Exchange and Repair <i>Gösta Grahne and Adrian Onet</i>	218–229
Forward-XPath and extended register automata on data-trees <i>Diego Figueira</i>	230–240
On the Aggregation Problem for Synthesized Web Services <i>Ting Deng, Wenfei Fan, Leonid Libkin, and Yinghui Wu</i>	241–250
Author Index	251

Foreword

The papers in this volume were presented at the 13th International Conference on Database Theory (ICDT'10), held in Lausanne, Switzerland, March 23–25, 2010. Starting last year, ICDT is being held jointly with the EDBT (Extending Database Technology) conference. EDBT took place on March 24–26, 2010.

In response to the Call for Papers, 55 submissions were received by the submission deadline of September 6, 2009. All were submitted electronically through EasyChair. EasyChair was also used for the Program Committee deliberations, which were held completely electronically. The Program Committee selected 20 papers for presentation. The paper "Composition with Target Constraints" by Marcelo Arenas, Ronald Fagin and Alan Nash was selected for the ICDT Best Paper Award. The papers "Answering Non-Monotonic Queries in Relational Data Exchange" by André Hernich, and "Forward-XPath and extended register automata on data-trees" by Diego Figueira were selected for the ICDT Best Student Paper Award.

In addition, there were three ICDT keynote speakers: Pierre Fraigniaud, Martin Grohe, and Amol Deshpande. Furthermore, Ian Horrocks and Val Tannen were keynote speakers for EDBT.

The submissions were not formally refereed and many of the papers are reports of ongoing research. It is expected that most will appear in a more polished and complete form in scientific journals.

Alan Nash, one of the recipients of this year's ICDT Best Paper Award, died in a tragic bicycle accident on November 15, 2009. Alan received a dual PhD in Computer Science and Mathematics from U.C. San Diego in 2006, after which he worked at the IBM Almaden Research Center. Alan was not only a brilliant researcher but also a fascinating and warm person who will be greatly missed by our community.

This proceedings are dedicated to his memory.

Luc Segoufin
INRIA & ENS Chacan, France
ICDT 2010 Program Committee Chair

Program Committee Members and External Reviewers

ICDT Program Committee

Henrik Bjorklund (Umeå University)	Maurizio Lenzerini (Sapienza U. di Roma)
Toon Calders (TU Eindhoven)	Sebastian Maneth (NICTA and UNSW)
Venkatesan Chakaravarthy (IBM India)	Wim Martens (TU Dortmund)
Edith Cohen (ATT)	Dan Olteanu (U. of Oxford)
Sara Cohen (Hebrew University of Jerusalem)	Rina Panigrahy (Microsoft)
Alin Deutsch (UC San Diego)	Reinhard Pichler (TU Vienna)
Sudipto Guha (UPenn)	Alkis Polyzotis (UC Santa Cruz)
Claudio Gutierrez (U. Chile at Santiago)	Luc Segoufin (chair, INRIA & ENS Cachan)
Benny Kimelfeld (IBM Almaden)	Cristina Sirangelo (ENS Cachan)
Solmaz Kolahi (British Columbia)	
Bart Kuijpers (U. Hasselt)	

External Reviewers

Dmitri Akatov	Szabolc Mikulas
Lyublena Antova	Anca Muscholl
Marcelo Arenas	Frank Neven
Pablo Barcelo	Vinayaka Pandit
Michael Benedikt	Fabio Patrizi
Jan Van den Bussche	Jorge Perez
Andrea Cali	Juan Reutter
Diego Calvanese	Yogish Sabharwal
Frank Drewes	Anish Das Sarma
Dries Van Dyck	Vadim Savenkov
Ingo Feinerer	Nicole Schweikardt
Alan Fekete	Marc Tommasi
Rajasekar Krishnamurthy	Riccardo Torlone
Ranko Lazic	Alejandro Vaisman
Nuno Lopes	

Provenance for Database Transformations

Pierre Fraigniaud
CNRS and University of Paris 7, France
pierre.fraigniaud@liafa.jussieu.fr

ABSTRACT

Network representations play an important role in many domains of computer science, ranging from data structures and graph algorithms, to parallel and distributed computing, and communication networks. Traditional network representations are usually global in nature. That is, in order to retrieve useful information, one must access a global data structure representing the entire network, even if the desired information is solely local, pertaining to only a few nodes. In contrast, the notion of informative labeling schemes suggests the use of a local representation of the network. The principle is to associate a label with each node, selected in a way that enables to infer information about any two nodes directly from their labels, without using any additional sources of information. Hence in essence, this method bases the entire representation on the set of labels alone. Obviously, labels of unrestricted size can be used to encode any desired information, including in particular the entire graph structure. The focus is thus on informative labeling

schemes which use labels as short as possible.

This talk will introduce the notion of informative labeling scheme to the audience, and will survey some of the important results achieved in this context. In particular, we will focus on the design of compact adjacency-, ancestry-, and distance-labeling schemes for trees. These schemes find applications in various contexts, including the design of small universal graphs, and the design of small universal posets. We will actually specifically emphasize the importance of ancestry- labeling scheme for the design of compact such schemes finds applications in XML search engines. In this context, even small improvements in the label size are important, and we will survey the most recent results in this domain.

The interested person is referred to the following paper, that includes pointers to many of the most important references on informative labeling schemes: Pierre Fraigniaud and Amos Korman, Compact Ancestry Labeling Schemes for XML Trees, in ACM-SIAM Symposium on Discrete Algorithms (SODA), 2010.