LIPIcs – Leibniz International Proceedings in Informatics

LIPIcs is a series of high-quality conference proceedings across all fields in informatics. LIPIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

Editorial Board

- Susanne Albers (TU München)
- Chris Hankin (Imperial College London)
- Deepak Kapur (University of New Mexico)
- Michael Mitzenmacher (Harvard University)
- Madhavan Mukund (Chennai Mathematical Institute)
- Catuscia Palamidessi (INRIA)
- Wolfgang Thomas (RWTH Aachen)
- Pascal Weil (Chair, CNRS and University Bordeaux)
- Reinhard Wilhelm (Saarland University)

ISSN 1868-8969

http://www.dagstuhl.de/lipics
Contents

Preface vii
ICDT 2015 Test of Time Award ix
Organization xi
External Reviewers xiii
List of Authors xv

Invited Talks

The Confounding Problem of Private Data Release
Graham Cormode ................................................................. 1

Using Locality for Efficient Query Evaluation in Various Computation Models
Nicole Schweikardt ............................................................. 13

Large-Scale Similarity Joins With Guarantees
Rasmus Pagh ................................................................. 15

Awards Session

A Declarative Framework for Linking Entities
Douglas Burdick, Ronald Fagin, Phokion G. Kolaitis, Lucian Popa, and
Wang-Chiew Tan ................................................................. 25

Asymptotic Determinacy of Path Queries using Union-of-Paths Views
Nadime Francis ................................................................. 44

(Regular Paper)

Games for Active XML Revisited
Martin Schuster and Thomas Schwentick ........................................ 60

Query Evaluation

Answering Conjunctive Queries with Inequalities
Paraschos Koutris, Tova Milo, Sudeepa Roy, and Dan Suciu ................. 76

SQL’s Three-Valued Logic and Certain Answers
Leonid Libkin ................................................................. 94

A Trichotomy in the Complexity of Counting Answers to Conjunctive Queries
Hubie Chen and Stefan Mengel ............................................. 110

Data Examples and Learning

Learning Tree Patterns from Example Graphs
Sara Cohen and Yaacov Y. Weiss ............................................. 127

18th International Conference on Database Theory (ICDT’15).
Editors: Marcelo Arenas and Martín Ugarte
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characterizing XML Twig Queries with Examples</td>
<td>Slawek Staworko and Piotr Wieczorek</td>
<td>144</td>
</tr>
<tr>
<td>The Product Homomorphism Problem and Applications</td>
<td>Balder ten Cate and Victor Dalmau</td>
<td>161</td>
</tr>
<tr>
<td><strong>Graph Databases and Semantic Web</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular Queries on Graph Databases</td>
<td>Juan L. Reutter, Miguel Romero, and Moshe Y. Vardi</td>
<td>177</td>
</tr>
<tr>
<td>Complexity and Expressiveness of ShEx for RDF</td>
<td>Slawek Staworko, Iovka Boneva, Jose E. Labra Gayo, Samuel Hym, Eric G. Prud’hommeaux, and Harold Solbrig</td>
<td>195</td>
</tr>
<tr>
<td>CONSTRUCT Queries in SPARQL</td>
<td>Egor V. Kostylev, Juan L. Reutter, and Martín Ugarte</td>
<td>212</td>
</tr>
<tr>
<td>Separability by Short Subsequences and Subwords</td>
<td>Piotr Hofman and Wim Martens</td>
<td>230</td>
</tr>
<tr>
<td><strong>Algorithms and Workflows</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process-Centric Views of Data-Driven Business Artifacts</td>
<td>Adrien Koutsos and Victor Vianu</td>
<td>247</td>
</tr>
<tr>
<td>On The I/O Complexity of Dynamic Distinct Counting</td>
<td>Xiaocheng Hu, Yufei Tao, Yi Yang, Shengyu Zhang, and Shuigeng Zhou</td>
<td>265</td>
</tr>
<tr>
<td>Shared-Constraint Range Reporting</td>
<td>Sudip Biswas, Manish Patil, Rahul Shah, and Sharma V. Thankachan</td>
<td>277</td>
</tr>
<tr>
<td><strong>Distributed Query Processing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal Broadcasting Strategies for Conjunctive Queries over Distributed Data</td>
<td>Bas Ketsman and Frank Neven</td>
<td>291</td>
</tr>
<tr>
<td>Datalog Queries Distributing over Components</td>
<td>Tom Ameloot, Bas Ketsman, Frank Neven, and Daniel Zinn</td>
<td>308</td>
</tr>
<tr>
<td>Distributed Streaming with Finite Memory</td>
<td>Frank Neven, Nicole Schweikardt, Frédéric Servais, and Tony Tan</td>
<td>324</td>
</tr>
<tr>
<td><strong>Consistency and Repairs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From Causes for Database Queries to Repairs and Model-Based Diagnosis and Back</td>
<td>Babak Salimi and Leopoldo Bertossi</td>
<td>342</td>
</tr>
<tr>
<td>On the Relationship between Consistent Query Answering and Constraint Satisfaction Problems</td>
<td>Carsten Lutz and Frank Wolter</td>
<td>363</td>
</tr>
<tr>
<td>On the Data Complexity of Consistent Query Answering over Graph Databases</td>
<td>Pablo Barceló and Gaëlle Fontaine</td>
<td>380</td>
</tr>
</tbody>
</table>
Preface

The 18th International Conference on Database Theory (ICDT 2015) was held in Brussels, Belgium, March 23–27, 2015. Originally biennial, the ICDT conference has been held annually and jointly with EDBT (“Extending Database Technology”) since 2009.

The proceedings of ICDT 2015 include a paper by Graham Cormode (University of Warwick) based on the keynote address by him, a paper by Rasmus Pagh (IT University of Copenhagen) based on the keynote address by him, an overview of an invited lecture by Nicole Schweikardt (Humboldt University of Berlin), a laudation concerning the ICDT 2015 Test of Time Award, and 22 research papers that were selected by the Program Committee from 50 submissions.

Out of the 22 accepted papers, the Program Committee selected the paper *A Declarative Framework for Linking Entities* by Douglas Burdick, Ronald Fagin, Phokion G. Kolaitis, Lucian Popa and Wang-Chiew Tan for the ICDT 2015 Best Paper Award. Furthermore, the Program Committee selected the paper *Asymptotic Determinacy of Path Queries using Union-of-Paths Views* by Nadime Francis for the ICDT 2015 Best Student Paper Award. The ICDT 2015 Test of Time Award is given to the paper *Efficient Computation of Frequent and Top-k Elements in Data Streams* by Ahmed Metwally, Divyakant Agrawal, and Amr El Abbadi, which originally appeared in the proceedings of ICDT 2005. Warmest congratulations to the authors of these award winning papers!

I thank all authors who submitted papers to ICDT 2015. I would also like to thank all members of the Program Committee, and the external reviewers, for the enormous amount of work they have done. The Program Committee did not meet in person, but carried out extensive discussions during the electronic PC meeting. I thank Andrei Voronkov for his EasyChair system, which made it easy to manage and coordinate the discussion.

I thank the ICDT Council members for their help in selecting the Program Committee and their advice on issues of policy during the conference. Special thanks also go to Martin Ugarte, the Proceedings Chair of EDBT/ICDT 2015. I thank many colleagues involved in the organisation of the conference for fruitful collaboration, in particular, Floris Geerts (EDBT/ICDT 2015 Conference Chair).

Marcelo Arenas
January 2015
ICDT 2015 Test of Time Award

In 2013, the International Conference on Database Theory (ICDT) began awarding the ICDT test-of-time award, with the goal of recognising one paper, or a small number of papers, presented at ICDT a decade earlier that have best met the “test of time”. In 2015, the award will recognise a paper from the ICDT 2005 proceedings that has had the most impact in terms of research, methodology, conceptual contribution, or transfer to practise over the past decade. The award will be presented during the EDBT/ICDT 2015 Joint Conference, March 23–27, 2015 in Brussels, Belgium.

The committee consisting of Serge Abiteboul, Sudeepa Roy, and chaired by Leonid Libkin (2005 PC co-chair) has chosen the following recipient of the 2015 ICDT Test of Time Award:

“Efficient Computation of Frequent and Top-k Elements in Data Streams”

by Ahmed Metwally, Divyakant Agrawal, and Amr El Abbadi

The paper studies the problem of finding items which occur most frequently in a data stream. This is a basic algorithmic problem of great practical importance. The paper proposed a SpaceSaving algorithm which was shown to both provide theoretical guarantees and to perform significantly better than others in practical scenarios. Since its publication, the paper has made impact in both algorithms research and practical implementations of streaming algorithms. Several implementations of the algorithm have been made available, and they are used both in industry and as benchmarks to compare against other streaming algorithms. The paper has been highly cited: many papers have made use of the data structure, either directly, or to solve new problems. The algorithm itself is easy to motivate and state, and consequently it is taught in a number of algorithms courses.
Organization

Conference Chair
Floris Geerts (U. of Antwerp)

Program Chair
Marcelo Arenas (PUC Chile)

Program Committee
Marcelo Arenas (PUC Chile)
Pankaj Agarwal (Duke U.)
Angela Bonifati (U. of Lille 1 & Inria)
Edith Cohen (Microsoft Research)
Giuseppe De Giacomo (Sapienza U. di Roma)
Daniel Deutch (Tel Aviv U.)
Gaelle Fontaine (U. of Chile)
Todd Green (LogicBlox & UC Davis)
Sebastian Maneth (U. of Edinburgh)
Filip Murlak (U. of Warsaw)
S Muthukrishnan (Rutgers U.)
Reinhard Pichler (Vienna U. of Technology)
Christopher Re (Stanford U.)
Cristian Riveros (PUC Chile)
Sudeepa Roy (U. of Washington)
Cristina Sirangelo (LSV, ENS-Cachan)
Yufei Tao (Chinese U. of Hong Kong)
Balder Ten Cate (LogicBlox & UC Santa Cruz)
Jan Van Den Bussche (Hasselt U.)
Stijn Vansummeren (U. Libre de Bruxelles)
Victor Vianu (UC San Diego)
David Woodruff (IBM Almaden)

Proceedings Chair
Martín Ugarte (PUC Chile)
External Reviewers

Foto Afrati  Filip Mazowiecki
Peter Alvaro  Andrew McGregor
Tom Ameloot  Marco Montali
Guillaume Bagan  Jelani Nelson
Pablo Barceló  Frank Neven
Manuel Bodirsky  Jorge Pérez
Iovka Boneva  Andreas Pieris
Pierre Bourhis  Chung Keung Poon
Paolo Ciaccia  Miguel Romero
Radu Ciucanu  Riccardo Rosati
Claire David  Emanuel Sallinger
Alin Deutsch  Vadim Savenkov
Esther Ezra  Evgeny Sherkhonov
Wenfei Fan  Mantas Simkus
George H. L. Fletcher  Sebastian Skritek
Olivier Gauwin  Wang-Chiew Tan
Sam Haney  Srikanta Tirthapura
Piotr Hofman  Sophie Tison
Hossein Jowhari  Domagoj Vrgoc
Phokion G. Kolaitis  Johannes Wallner
Paraschos Koutris  Adam Witkowski
Kasper Larsen  Peter Wood
Domenico Lembo  Zhihui Wu
Maurizio Lenzerini  You Wu
Jerry Li  Xiaokui Xiao
Leonid Libkin  Ke Yi
Katja Losemann  Marc Zeitoun
Wim Martens  Thomas Zeume
Dániel Marx  Wuzhou Zhang
## List of Authors

<table>
<thead>
<tr>
<th>Authors</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom Ameloot</td>
<td>Manish Patil</td>
</tr>
<tr>
<td>Pablo Barceló</td>
<td>Lucian Popa</td>
</tr>
<tr>
<td>Leopoldo Bertossi</td>
<td>Eric Prud’Hommeaux</td>
</tr>
<tr>
<td>Sudip Biswas</td>
<td>Juan L. Reutter</td>
</tr>
<tr>
<td>Iovka Boneva</td>
<td>Miguel Romero</td>
</tr>
<tr>
<td>Douglas Burdick</td>
<td>Sudeepa Roy</td>
</tr>
<tr>
<td>Balder ten Cate</td>
<td>Babak Salimi</td>
</tr>
<tr>
<td>Hubie Chen</td>
<td>Martin Schuster</td>
</tr>
<tr>
<td>Sara Cohen</td>
<td>Nicole Schweikard</td>
</tr>
<tr>
<td>Graham Cormode</td>
<td>Thomas Schwentick</td>
</tr>
<tr>
<td>Victor Dalmau</td>
<td>Frédéric Servais</td>
</tr>
<tr>
<td>Ronald Fagin</td>
<td>Rahul Shah</td>
</tr>
<tr>
<td>Gaëlle Fontaine</td>
<td>Harold Solbrig</td>
</tr>
<tr>
<td>Nadime Francis</td>
<td>Shawek Staworko</td>
</tr>
<tr>
<td>Piotr Hofman</td>
<td>Dan Suciu</td>
</tr>
<tr>
<td>Xiaocheng Hu</td>
<td>Tony Tan</td>
</tr>
<tr>
<td>Samuel Hym</td>
<td>Wang-Chiew Tan</td>
</tr>
<tr>
<td>Bas Ketsman</td>
<td>Yufei Tao</td>
</tr>
<tr>
<td>Phokion G. Kolaitis</td>
<td>Sharma Thankachan</td>
</tr>
<tr>
<td>Egor V. Kostylev</td>
<td>Martin Ugarte</td>
</tr>
<tr>
<td>Paraschos Koutris</td>
<td>Moshe Y. Vardi</td>
</tr>
<tr>
<td>Adrien Koutsos</td>
<td>Victor Vianu</td>
</tr>
<tr>
<td>Jose Emilio Labra Gayo</td>
<td>Yaacov Y. Weiss</td>
</tr>
<tr>
<td>Leonid Libkin</td>
<td>Piotr Wieczorek</td>
</tr>
<tr>
<td>Carsten Lutz</td>
<td>Frank Wolter</td>
</tr>
<tr>
<td>Wim Martens</td>
<td>Yi Yang</td>
</tr>
<tr>
<td>Stefan Mengel</td>
<td>Shengyu Zhang</td>
</tr>
<tr>
<td>Tova Milo</td>
<td>Shuigeng Zhou</td>
</tr>
<tr>
<td>Frank Neven</td>
<td>Daniel Zinn</td>
</tr>
<tr>
<td>Rasmus Pagh</td>
<td></td>
</tr>
</tbody>
</table>