21st International Conference on Database Theory

ICDT 2018, March 26–29, 2018, Vienna, Austria

Edited by

Benny Kimelfeld
Yael Amsterdamer
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
- Luca Aceto (Chair, Gran Sasso Science Institute and Reykjavik University)
- Susanne Albers (TU München)
- Chris Hankin (Imperial College London)
- Deepak Kapur (University of New Mexico)
- Michael Mitzenmacher (Harvard University)
- Madhavan Mukund (Chennai Mathematical Institute)
- Anca Muscholl (University Bordeaux)
- Catuscia Palamidessi (INRIA)
- Raimund Seidel (Saarland University and Schloss Dagstuhl – Leibniz-Zentrum für Informatik)
- Thomas Schwentick (TU Dortmund)
- Reinhard Wilhelm (Saarland University)

ISSN 1868-8969

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

Preface
Benny Kimelfeld and Yael Amsterdamer ................................................................. 0:vii

Organization ........................................................................................................... 0:ix

External Reviewers ................................................................................................. 0:xi

Contributing Authors ............................................................................................. 0:xiii

ICDT 2018 Test of Time Award
Pablo Barceló, Richard B. Hull and Victor Vianu .................................................... 0:xv

Invited Talks

Fine-grained Algorithms and Complexity
Virginia Vassilevska Williams ................................................................. 1:1–1:1

Join Algorithms: From External Memory to the BSP
Ke Yi .......................................................... 2:1–2:1

An Update on Dynamic Complexity Theory
Thomas Zeume .......................................................... 3:1–3:1

Regular Papers

Rewriting Guarded Existential Rules into Small Datalog Programs
Shqiponja Ahmetaj, Magdalena Ortiz, and Mantas Šimkus ................................. 4:1–4:24

Enumeration on Trees under Relabelings
Antoine Amarilli, Pierre Bourhis, and Stefan Mengel ........................................... 5:1–5:18

Connecting Width and Structure in Knowledge Compilation
Antoine Amarilli, Mikaël Monet, and Pierre Senellart ........................................... 6:1–6:17

A More General Theory of Static Approximations for Conjunctive Queries
Pablo Barceló, Miguel Romero, and Thomas Zeume ............................................ 7:1–7:22

Answering UCQs under Updates and in the Presence of Integrity Constraints
Christoph Berkholtz, Jens Keppeler, and Nicole Schweikardt ............................ 8:1–8:19

Expressivity and Complexity of MongoDB Queries
Elena Botoeva, Diego Calvanese, Benjamin Cogrel and Guohui Xiao ............... 9:1–9:23

On the Expressive Power of Query Languages for Matrices
Robert Brijder, Floris Geerts, Jan Van den Bussche, and Timmy Weerwag ...... 10:1–10:17

Enumeration Complexity of Conjunctive Queries with Functional Dependencies
Nofar Carmeli and Markus Kröll .............................................................. 11:1–11:17
Preserving Constraints with the Stable Chase
David Carral, Markus Krötzsch, Maximilian Marx, Ana Ozaki, and Sebastian Rudolph ............................................................... 12:1–12:19

Fast Sketch-based Recovery of Correlation Outliers
Graham Cormode and Jacques Dark ............................................. 13:1–13:18

Satisfiability for SCULPT-Schemas for CSV-Like Data
Johannes Doleschal, Wim Martens, Frank Neven, and Adam Witkowski 14:1–14:19

Querying the Unary Negation Fragment with Regular Path Expressions
Jean Christoph Jung, Carsten Lutz, Mauricio Martel, and Thomas Schneider .... 15:1–15:18

Covers of Query Results
Ahmet Kara and Dan Olteanu ................................................... 16:1–16:22

Distribution Policies for Datalog
Bas Ketsman, Ays Alibayhouthi, and Paraschos Koutris 17:1–17:22

Parallel-Correctness and Transferability for Conjunctive Queries under Bag Semantics
Bas Ketsman, Frank Neven, and Brecht Vandevoort .............................. 18:1–18:16

Evaluation and Enumeration Problems for Regular Path Queries
Wim Martens and Tina Trautner ................................................ 19:1–19:21

Massively Parallel Entity Matching with Linear Classification in Low Dimensional Space
Yufei Tao ........................................................................ 20:1–20:19
The 21st International Conference on Database Theory (ICDT 2018) was held in Vienna, Austria, on March 26-29, 2018. Originally biennial, the ICDT conference has been held annually and jointly with the conference on Extending Database Technology (EDBT) since 2009. The proceedings of ICDT 2018 includes an overview of the keynote talks by Virginia Vassilevska Williams (MIT EECS, MA, USA), by Ke Yi (Hong Kong University of Science and Technology, Hong Kong, China), and by Thomas Zeume (TU Dortmund University, Germany), as well as 17 research papers that were selected by the Program Committee.

Out of the 17 accepted papers, the Program Committee selected the paper “Evaluation and Enumeration Problems for Regular Path Queries” by Wim Martens and Tina Trautner (University of Bayreuth, Germany) for the ICDT 2018 Best Paper Award.

We wish to thank the many contributors to ICDT 2018: the authors of the submitted papers, the external reviewers, the keynote speakers, the EDBT/ICDT 2018 local organization officers, and the conference General Chair Reinhard Pichler. We wish to especially thank the members of the ICDT 2018 Program Committee, who invested considerable time and effort providing thorough paper reviews and conducting careful discussions. We also wish to thank ICDT Council Chair Wim Martens and ICDT 2017 Program Committee Chair Michael Benedikt for their guidance and support, as well as ICDT 2017 Proceedings Chair Giorgio Orsi for useful advice on the publicity and production of the proceedings. Last, we wish to acknowledge Michael Wagner and Marc Herbstritt from Dagstuhl Publishing for their support with the publication of the proceedings in the LIPIcs (Leibniz International Proceedings in Informatics) Series.

Benny Kimelfeld and Yael Amsterdamer

March 2018
Organization

Conference Chair
Vim Martens (University of Bayreuth, Germany)

Program Chair
Benny Kimelfeld (Technion - Israel Institute of Technology, Israel)

Publication and Proceedings Chair
Yael Amsterdamer (Bar-Ilan University, Israel)

Program Committee Members
Antoine Amarilli (Institut Télécom; Télécom ParisTech; CNRS LTCI, France)
Yael Amsterdamer (Bar-Ilan University, Israel)
Pablo Barcelo (DCC, Universidad de Chile, Chile)
Keren Censor-Hillel (Technion - Israel Institute of Technology, Israel)
Claire David (Universite Paris-Est Marne-la-Vallee, France)
Martin Grohe (RWTH Aachen, Germany)
Aidan Hogan (DCC, Universidad de Chile, Chile)
Benny Kimelfeld (Technion - Israel Institute of Technology, Israel)
Roman Kontchakov (Birkbeck, University of London, UK)
Ilya Mironov (Google, US)
Filip Murlak (University of Warsaw, Poland)
Jelani Nelson (Harvard University, US)
Matthias Niewerth (University of Bayreuth, Germany)
Rasmus Pagh (IT University of Copenhagen, Denmark)
Sudeepa Roy (Duke University, US)
Dan Suciu (University of Washington, US)
Yufei Tao (Chinese University of Hong Kong, Hong Kong)
Jan Van den Bussche (Hasselt University; Transnational University of Limburg, Belgium)
Stijn Vansummeren (Université Libre de Bruxelles, Belgium)
Victor Vianu (UC San Diego, US)
David P. Woodruff (IBM Almaden, US)
External Reviewers

Foto Afrati
Arvind Arasu
Daniel Deutch
Alin Deutsch
Nadime Francis
Dominik D. Freydenberger
Herman Haverkort
Dylan Hutchison
Neil Immerman
Silvio Lattanzi
Yi Li
Leonid Libkin
Andrew McGregor
Hung Q. Ngo
Thomas Pellissier Tanon
Sylvain Schmitz
Luc Segoufin
Francesco Silvestri
Xiaorui Sun
Martin Ugarte
Contributing Authors

Shqiponja Ahmetaj
Aws Albarghouthi
Antoine Amarilli
Pablo Barcelo
Christoph Berkholz
Elena Botoceva
Pierre Bourhis
Robert Brijder
Diego Calvanese
Nofar Carmeli
David Carral
Benjamin Cogrel
Graham Cormode
Jacques Dark
Johannes Doleschal
Floris Geerts
Jean Christoph Jung
Ahmet Kara
Jens Keppeler
Bas Ketsman
Paraschos Koutris
Markus Kröll
Markus Krötzsch
Carsten Lutz
Mauricio Martel

Wim Martens
Maximilian Marx
Stefan Mengel
Mikaël Monet
Frank Neven
Dan Olteanu
Magdalena Ortiz
Ana Ozaki
Miguel Romero
Sebastian Rudolph
Thomas Schneider
Nicole Schweikardt
Pierre Senellart
Mantas Simkus
Yufei Tao
Tina Trautner
Jan Van den Bussche
Brecht Vandevoort
Virginia Vassilevska Williams
Timmy Weerwag
Adam Witkowski
Guohui Xiao
Ke Yi
Thomas Zeume
ICDT 2018 Test of Time Award

In 2013, the International Conference on Database Theory (ICDT) began awarding the ICDT Test of Time (ToT) award, with the goal of recognizing one paper, or a small number of papers, presented at ICDT at least a decade earlier, that have best met the “test of time”. The ICDT ToT award for 2018 was presented during the EDBT/ICDT 2018 Joint Conference, March 26–29, 2018 in Vienna, Austria. The ICDT 2018 ToT Award Committee was charged with selecting the paper(s) from the ICDT 1999 and 2001 proceedings that have had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice over the past decade. After careful consideration, the committee selected the following papers as joint award winners for 2018:

Kevin S. Beyer, Jonathan Goldstein, Raghu Ramakrishnan, Uri Shaft:
When Is "Nearest Neighbor" Meaningful?

This much cited, thought-provoking paper challenged the conventional wisdom in high-dimensional similarity search by highlighting a surprising phenomenon: the distance to a point’s nearest neighbor is almost the same as the distance to its farthest neighbor, in high dimensions. This opened up new fundamental questions about the utility of indexes for high-dimensional similarity search and pointed out a major weakness in the methodology of empirical evaluations of prior works.

Peter Buneman, Sanjeev Khanna, Wang Chiew Tan:
Why and Where: A Characterization of Data Provenance

A seminal contribution to the foundations of data provenance, this article considered for the first time the 'where' and 'why' flavors of provenance in a unified, application independent framework, using a general-purpose semi-structured data model. 'Why' provenance refers to the portion of the database that justifies the presence of an item in the answer to a query, while 'where' provenance identifies actual locations in the database from which the answer values were extracted. This influential and highly cited paper opened the way to an important research area focusing on various semantic and computational aspects of data provenance, still very active today.

Pablo Barcelo Richard B. Hull Victor Vianu
Univ. of Chile IBM T.J. Watson Research Center UC San Diego

The ICDT Test-of-Time Award Committee for 2018