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)
- Christel Baier (TU Dresden)
- Mikolaj Bojanczyk (University of Warsaw)
- Roberto Di Cosmo (INRIA and University Paris Diderot)
- Javier Esparza (TU München)
- Meena Mahajan (Institute of Mathematical Sciences)
- Dieter van Melkebeek (University of Wisconsin-Madison)
- Anca Muscholl (University Bordeaux)
- Luke Ong (University of Oxford)
- Catuscia Palamidessi (INRIA)
- Thomas Schwentick (TU Dortmund)
- Raimund Seidel (Saarland University and Schloss Dagstuhl – Leibniz-Zentrum für Informatik)

ISSN 1868-8969

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

Preface
Carsten Lutz and Jean Christoph Jung ........................................... 0:vii

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

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

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

ICDT 2020 Test of Time Award ................................................... 0:xv

Invited Talks

Facets of Probabilistic Databases
Benny Kimelfeld ............................................................................. 1:1–1:1

What Makes a Variant of Query Determinacy (Un)Decidable?
Jerzy Marcinkowski ......................................................................... 2:1–2:20

Current Challenges in Graph Databases
Juan L. Reutter ................................................................................... 3:1–3:1

Regular Papers

Executable First-Order Queries in the Logic of Information Flows
Heba Aamer, Bart Bogaerts, Dimitri Sarinx, Eugenia Ternovska, and
Jan Van den Bussche ............................................................ 4:1–4:14

A Dichotomy for Homomorphism-Closed Queries on Probabilistic Graphs
Antoine Amarilli and İsmail İlkan Ceylan ............................. 5:1–5:20

On the Expressiveness of LARA: A Unified Language for Linear and Relational
Algebra
Pablo Barceló, Nelson Higuera, Jorge Pérez, and Bernardo Subercaseaux ......... 6:1–6:20

Random Sampling and Size Estimation Over Cyclic Joins
Yu Chen and Ke Yi .............................................................. 7:1–7:18

Weight Annotation in Information Extraction
Johannes Doleschal, Benny Kimelfeld, Wim Martens, and Liat Peterfreund ...... 8:1–8:18

Containment of UC2RPQ: The Hard and Easy Cases
Diego Figueira .............................................................................. 9:1–9:18

On Equivalence and Cores for Incomplete Databases in Open and Closed Worlds
Henrik Forssell, Evgeny Kharchamov, and Eegenij Thorstensen .......................... 10:1–10:21

23rd International Conference on Database Theory (ICDT 2020).
Editors: Carsten Lutz and Jean Christoph Jung
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
Contents

Dynamic Complexity of Document Spanners
Dominik D. Freydenberger and Sam M. Thompson ......................... 11:1–11:21

When Can Matrix Query Languages Discern Matrices?
Floris Geerts .......................................................... 12:1–12:18

Distribution Constraints: The Chase for Distributed Data

Towards Streaming Evaluation of Queries with Correlation in Complex Event Processing
Alejandro Grez and Cristian Riveros ...................................... 14:1–14:17

On the Expressiveness of Languages for Complex Event Recognition
Alejandro Grez, Cristian Riveros, Martín Ugarte, and Stijn Vansummeren ...... 15:1–15:17

Infinite Probabilistic Databases
Martin Grohe and Peter Lindner ............................................. 16:1–16:20

Coordination-Free Byzantine Replication with Minimal Communication Costs
Jelle Hellings and Mohammad Sadoghi ..................................... 17:1–17:20

Integrity Constraints Revisited: From Exact to Approximate Implication
Batya Kenig and Dan Suciu ................................................ 18:1–18:20

Datalog with Negation and Monotonicity
Bas Ketsman and Christoph Koch ......................................... 19:1–19:18

The Shapley Value of Tuples in Query Answering
Ester Livshits, Leopoldo Bertossi, Benny Kimelfeld, and Moshe Sebag .......... 20:1–20:19

Optimal Joins Using Compact Data Structures
Gonzalo Navarro, Juan L. Reutter, and Javiel Rojas-Ledesma ............... 21:1–21:21

The Space Complexity of Inner Product Filters
Rasmus Pagh and Johan Sivertsen ......................................... 22:1–22:14

A Family of Centrality Measures for Graph Data Based on Subgraphs
Cristian Riveros and Jorge Salas ......................................... 23:1–23:18

Reverse Prevention Sampling for Misinformation Mitigation in Social Networks
Michael Simpson, Venkatesh Srinivasan, and Alex Thomo ................... 24:1–24:18

A Simple Parallel Algorithm for Natural Joins on Binary Relations
Yufei Tao ................................................................. 25:1–25:18
Preface

The 23rd International Conference on Database Theory (ICDT 2020) was held in Copenhagen, Denmark, from March 30 to April 2, 2020. The Program Committee has selected 22 research papers out of 69 submissions for publication at the conference. It has further decided to give the best paper award to *A Dichotomy for Homomorphism-Closed Queries on Probabilistic Graphs* by Antoine Amarilli and İsmail İlkan Ceylan. We congratulate the winners! Apart from the 22 regular papers, these proceedings include abstracts for the invited (shared) EDBT/ICDT keynotes by Benny Kimelfeld (Technion, Israel) and by Juan L. Reutter (PUC Chile) and the invited paper associated with the ICDT invited talk by Jerzy Marcinkowski (University of Wrocław, Poland).

A committee formed by Frank Neven, Andreas Pieris, and Jorge Pérez has decided to give the Test of Time Award for ICDT 2020 to the ICDT 2010 paper *Foundations of SPARQL query optimizations* by Michael Schmidt, Michael Meier, and Georg Lausen. We congratulate also the winners of this award!

We would like to thank all people who contributed to the success of ICDT 2020, including the authors of all submitted papers, keynote and invited talk speakers, and, of course, all members of the Program Committee as well as the external reviewers, for the very substantial work that they have invested over the two submission cycles of ICDT 2020. Their commitment and sagacity were crucial to ensure that the final program of the conference satisfies the highest standards. We would also like to thank the ICDT Council members for their support on a wide variety of matters, the local organizers of the EDBT/ICDT 2020 conference, led by General Chairs Yongluan Zhou and Marcos Antonio Vaz Salles, for the great job they did in organizing the conference and co-located events. Finally, we wish to acknowledge Dagstuhl Publishing for their support with the publication of the proceedings in the LIPIcs (Leibniz International Proceedings in Informatics) series.

Carsten Lutz and Jean Christoph Jung
March 2020
Organization

General Chairs
Yongluan Zhou (University of Copenhagen)
Marcos Antonio Vaz Salles (University of Copenhagen)

Program Chair
Carsten Lutz (University of Bremen)

Proceedings Chair
Jean Christoph Jung (University of Bremen)

Program Committee
Marcelo Arenas (PUC, Santiago de Chile)
Michael Benedikt (University of Oxford)
Christoph Berkholz (HU Berlin)
Angela Bonifati (University Lyon 1)
Pierre Bourhis (University of Lille)
James Cheney (University of Edinburgh)
Graham Cormode (University of Warwick)
Victor Dalmau (UPF, Barcelona)
Claire David (University Paris-Est)
Floris Geerts (University of Antwerp)
Bas Ketsman (EPFL, Lausanne)
Daniel Kifer (Penn State University)
Leonid Libkin (University of Edinburgh)
Sebastian Maneth (University of Bremen)
Filip Murlak (University of Warsaw)
Reinhard Pichler (TU Vienna)
Andreas Pieris (University of Edinburgh)
Sebastian Rudolph (TU Dresden)
Thomas Schwentick (University of Dortmund)
Uri Stemmer (Ben-Gurion University, Negev)
Domagoj Vrgoc (PUC, Santiago de Chile)
Frank Wolter (University of Liverpool)
External Reviewers

Diego Arroyuelo
Jan Van den Bussche
Nadine Francis
Alan Fekete
Victor Marsault
Stefan Mengel
Matthias Niewerth
Liat Peterfreund
Ling Ren
Juan L. Reutter
Cristian Riveros
Dimitris Sacharidis
Oskar Skibski
Rossano Venturini
Nils Vortmeier
Marcin Waniecki
Contributing Authors

Heba Aamer Batya Kenig Thomas Schwentick
Antoine Amarilli Bas Ketsman Moshe Sebag
Pablo Barceló Evgeny Kharlamov Michael Simpson
Leopoldo Bertossi Benny Kimelfeld Johan Sivertsen
Bart Bogaerts Christoph Koch Venkatesh Srinivasan
Jan Van den Bussche Peter Lindner Bernardo Subercaseux
Yu Chen Ester Livshits Dan Suciu
Johannes Doleschal Wim Martens Dimitri Surinx
Diego Figueira Gonzalo Navarro Yufei Tao
Henrik Forssell Frank Neven Evgenia Ternovska
Dominik D. Freydenberger Rasmus Pagh Alex Thomo
Gaetano Geck Jorge Pérez Samuel M. Thompson
Floris Geerts Liat Peterfreund Evgenij Thorstensen
Alejandro Grez Juan L. Reutter Martin Ugarte
Martin Grohe Cristian Riveros Stijn Vansummeren
Jelle Hellings Javiel Rojas-Ledesma Ke Yi
Nelson Higuera Mohammad Sadoghi
Ismail Ilkan Ceylan Jorge Salas
ICDT 2020 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 a decade earlier that have best met the “test of time”. In 2020, the award recognizes a paper from the ICDT 2010 proceedings that has had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice over the past decade. The award was presented during the EDBT/ICDT 2020 Joint Conference, March 30-April 2, 2020, in Copenhagen, Denmark.

The 2020 ToT Committee consists of Frank Neven (chair), Andreas Pieris and Jorge Pérez. After careful consideration and soliciting external assessments, the committee has chosen the following recipient of the 2020 ICDT Test of Time Award:

*Foundations of SPARQL query optimization*

Michael Schmidt, Michael Meier, Georg Lausen

This paper is one of the stepping stones that placed Semantic Web query languages on the radar of Database Theory. The paper focuses on SPARQL, the standard language for querying the graph-based model underlying Semantic Web data. It presents an elegant complexity analysis of SPARQL pinpointing the impact of every single operator of the language. It also derives an impressive set of optimization rules highlighting the similarities as well as the important differences between SPARQL and more classical languages such as relational algebra and SQL.

The paper has had a substantial impact counting more than 300 citations. It has influenced the theoretical development of SPARQL and its extensions, the design and construction of Benchmarks for comparing implementations, and also the now ubiquitous research on knowledge-graph data and queries.

Frank Neven Andreas Pieris Jorge Pérez
Hasselt University University of Edinburgh Universidad de Chile

*The ICDT Test-of-Time Award Committee for 2020*