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
Ke Yi and Zhewei Wei .......................................................... 0:vii

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

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

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

ICDT 2021 Test of Time Award .......................................................... 0:xv

Invited Talks

Explainability Queries for ML Models and its Connections with Data Management Problems
Pablo Barceló ........................................................................ 1:1–1:1

Comparing Apples and Oranges: Fairness and Diversity in Ranking
Julia Stoyanovich ............................................................ 2:1–2:1

Regular Papers

Box Covers and Domain Orderings for Beyond Worst-Case Join Processing
Kaleb Alway, Eric Blais, and Semih Salihoglu ........................................ 3:1–3:23

A Purely Regular Approach to Non-Regular Core Spanners
Markus L. Schmid and Nicole Schweikardt ....................................... 4:1–4:19

Ranked Enumeration of Conjunctive Query Results
Shaleen Deep and Paraschos Koutris ............................................. 5:1–5:19

Towards Optimal Dynamic Indexes for Approximate (and Exact) Triangle Counting
Shangqi Lu and Yufei Tao ....................................................... 6:1–6:23

Grammars for Document Spanners
Liat Peterfreund ................................................................. 7:1–7:18

Input–Output Disjointness for Forward Expressions in the Logic of Information Flows
Heba Aamer and Jan Van den Bussche ............................................. 8:1–8:18

Conjunctive Queries: Unique Characterizations and Exact Learnability
Balder ten Cate and Victor Dalmau ............................................... 9:1–9:24

The Complexity of Aggregates over Extractions by Regular Expressions
Johannes Doleuschal, Noa Bratman, Benny Kimelfeld, and Wim Martens ........ 10:1–10:20
Contents

Answer Counting Under Guarded TGDs
Cristina Feier, Carsten Lutz, and Marcin Przybyłko ............................. 11:1–11:22

Maximum Coverage in the Data Stream Model: Parameterized and Generalized
Andrew McGregor, David Tench, and Hoa T. Vu .............................. 12:1–12:20

Diverse Data Selection under Fairness Constraints

Enumeration Algorithms for Conjunctive Queries with Projection
Shaleen Deep, Xiao Hu, and Paraschos Koutris ................................. 14:1–14:17

The Shapley Value of Inconsistency Measures for Functional Dependencies
Ester Livshits and Benny Kimelfeld .............................................. 15:1–15:19

Database Repairing with Soft Functional Dependencies
Nofar Carmeli, Martin Grohe, Benny Kimelfeld, Ester Livshits, and Muhammad Tibi
................................................................. 16:1–16:17

Uniform Reliability of Self-Join-Free Conjunctive Queries
Antoine Amarilli and Benny Kimelfeld ........................................... 17:1–17:17

Efficient Differentially Private $F_0$ Linear Sketching
Rasmus Pagh and Nina Mesing Stausholm ....................................... 18:1–18:19

Fine-Grained Complexity of Regular Path Queries

Ranked Enumeration of MSO Logic on Words

Approximate Similarity Search Under Edit Distance Using Locality-Sensitive
Hashing
Samuel McCauley ............................................................... 21:1–21:22

Locality-Aware Distribution Schemes
Preface

The 24th International Conference on Database Theory (ICDT 2021) was held in Nicosia, Cyprus, from March 23 to 26, 2021. The Program Committee has selected 20 research papers out of 42 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 20 regular papers, these proceedings include abstracts for the invited (shared) EDBT/ICDT keynotes by Pablo Barceló (Pontificia Universidad Católica de Chile) and by Julia Stoyanovich (New York University).

A committee formed by Yael Amsterdamer, Rasmus Pagh, and Pierre Senellart has decided to give the Test of Time Award for ICDT 2021 to the ICDT 2011 paper Knowledge compilation meets database theory: compiling queries to decision diagrams by Abhay Jha and Dan Suciu. We congratulate also the winners of this award!

We would like to thank all people who contributed to the success of ICDT 2021, 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 2021. 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 2021 conference, led by General Chairs Demetris Zeinalipour and Panos K. Chrysanthis, 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.

Ke Yi and Zhewei Wei
March 2021
Organization

General Chairs
Demetris Zeinalipour (University of Cyprus)
Panos K. Chrysanthis (University of Cyprus and University of Pittsburgh)

Program Chair
Ke Yi (The Hong Kong University of Science and Technology)

Proceedings Chair
Zhewei Wei (Renmin University of China)

Program Committee
Yael Amsterdamer (Bar Ilan University)
Meghyn Bienvenu (CNRS, University of Bordeaux)
Vladimir Braverman (Johns Hopkins University)
Marco Calautti (University of Trento)
Hubie Chen (Birkbeck, University of London)
Sara Cohen (The Hebrew University)
Martin Grohe (RWTH Aachen University)
Benny Kimelfeld (Technion, Israel Institute of Technology)
Paraschos Koutris (University of Wisconsin-Madison)
Domenico Lembo (Sapienza University of Rome)
Stefan Mengel (CNRS, CRIL)
Matthias Niewerth (University of Bayreuth)
Dan Olteanu (University of Oxford)
Rasmus Pagh (IT University of Copenhagen)
Sudeepa Roy (Duke University)
Atri Rudra (University at Buffalo, SUNY)
Francesco Scarcello (DIMES, University of Calabria)
Srikanta Tirthapura (Apple Inc., Iowa State University)
Stijn Vansummeren (Université Libre de Bruxelles)
Jef Wijsen (University of Mons)
External Reviewers

Antoine Amarill
Mohammad Javad Amiri
Alexandr Andoni
Marcelo Arenas
Anton Belyy
Andrea Calí
Nofar Carmeli
Shaleen Deep
Cibele Freire
Dominik D. Freydenberger
Filippo Furfaro
Gianluigi Greco
Montserrat Hermo
Xiao Hu
Raj Jayaram
Zhengjie Miao
Cristian Molinaro
Frank Neven
Milos Nikolic
Francesco Parisi
Tina Popp
Andrea Pugliese
Juan L. Reutter
Cristian Riveros
Domenico Saccà
Uri Stemmer
Philip Wellnitz
Samson Zhou
## Contributing Authors

<table>
<thead>
<tr>
<th>Heba Aamer</th>
<th>Alejandro Grez</th>
<th>Jeffrey Naughton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaleb Alway</td>
<td>Martin Grohe</td>
<td>Rasmus Pagh</td>
</tr>
<tr>
<td>Antoine Amarilli</td>
<td>Xiao Hu</td>
<td>Liat Peterfreund</td>
</tr>
<tr>
<td>Eric Blais</td>
<td>Louis Jachiet</td>
<td>Marcin Przybyłko</td>
</tr>
<tr>
<td>Pierre Bourhis</td>
<td>Benny Kimelfeld</td>
<td>Cristian Riveros</td>
</tr>
<tr>
<td>Noa Bratman</td>
<td>Paraschos Koutris</td>
<td>Semih Salihoglu</td>
</tr>
<tr>
<td>Jan Van den Bussche</td>
<td>Ester Livshits</td>
<td>Markus L. Schmid</td>
</tr>
<tr>
<td>Nofar Carmeli</td>
<td>Shangqi Lu</td>
<td>Nicole Schweikardt</td>
</tr>
<tr>
<td>Katrin Casel</td>
<td>Carsten Lutz</td>
<td>Nina Mosing Stausholm</td>
</tr>
<tr>
<td>Balder ten Cate</td>
<td>Wim Martens</td>
<td>Bruhathi Sundaramurthy</td>
</tr>
<tr>
<td>Victor Dalmau</td>
<td>Samuel McCauley</td>
<td>Yufei Tao</td>
</tr>
<tr>
<td>Shaleen Deep</td>
<td>Andrew McGregor</td>
<td>David Tench</td>
</tr>
<tr>
<td>Johannes Doleschal</td>
<td>Alexandra Meliou</td>
<td>Muhammad Tibi</td>
</tr>
<tr>
<td>Cristina Feier</td>
<td>Zafeiria Moumoulidou</td>
<td>Hoa T. Vu</td>
</tr>
</tbody>
</table>

24th International Conference on Database Theory (ICDT 2021).
Editors: Ke Yi and Zhewei Wei
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
ICDT 2021 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 2021, the award recognizes a paper from the ICDT 2011 proceedings that has had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice over the past decade. The award is to be presented during the EDBT/ICDT 2021 Joint Conference, March 23–26, 2021 in Nicosia, Cyprus.

The ICDT 2021 Test of Time Award committee consists of Yael Amsterdamer (Chair), Rasmus Pagh, and Pierre Senellart. After careful consideration and soliciting external assessments, the committee has chosen the following recipient of the 2021 ICDT Test of Time Award:

Knowledge compilation meets database theory: compiling queries to decision diagrams
Abhay Jha and Dan Suciu

There are two main approaches to computing the probability of a query result over probabilistic databases: the extensional approach exploits the structure of the query for efficient evaluation for some classes of queries; the intensional approach first tractably computes a representation of the lineage of the query and then attempts to compute the probability of this Boolean function. This paper shows that a number of cases known to be tractable in the extensional method lead to tractability in the intensional method because lineages can be produced in specific tractable formalisms (such as OBDDs, FBDDs, d-DNNFs) which are well-studied target compilation classes in knowledge compilation, and for which weighted model counting is tractable. The paper leaves open the major question of whether all tractable cases can be explained in the same manner.

With their work, Jha and Suciu established a strong connection between the fields of knowledge compilation and probabilistic databases, which was both foundational and entirely original. This has sparked research in and across different areas: in database theory in the form of further refinements of the results and progress towards the resolution of the question left open; in database systems by demonstrating that the intensional approach and the use of knowledge compilation techniques are viable for probabilistic query evaluation; and in knowledge compilation by further motivating and reviving interest for the study of the weighted variant of the model counting problem.

Yael Amsterdamer  Rasmus Pagh  Pierre Senellart
Bar-Ilan University  University of Copenhagen  ENS, PSL University

The ICDT Test-of-Time Award Committee for 2021

24th International Conference on Database Theory (ICDT 2021).
Editors: Ke Yi and Zhewei Wei
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany