

27th International Conference on Database Theory

ICDT 2024, March 25–28, 2024, Paestum, Italy

Edited by

Graham Cormode

Michael Shekelyan



Editors

Graham Cormode 

University of Warwick, UK
g.cormode@warwick.ac.uk

Michael Shekelyan 

Queen Mary University of London, UK
m.shekelyan@qmul.ac.uk

ACM Classification 2012

Information systems → Data management systems; Information systems → Structured Query Language; Information systems → Relational database query languages; Information systems → Information extraction; Information systems → Join algorithms; Information systems → Query languages; Information systems → Relational database model; Theory of computation; Theory of computation → Database theory; Theory of computation → Logic; Theory of computation → Logic and databases; Theory of computation → Database query languages (principles); Theory of computation → Database query processing and optimization (theory); Theory of computation → Data structures design and analysis; Theory of computation → Graph algorithms analysis; Theory of computation → Streaming, sublinear and near linear time algorithms; Theory of computation → Database constraints theory; Theory of computation → Rewrite systems; Theory of computation → Regular languages; Theory of computation → Grammars and context-free languages; Theory of computation → Computational Geometry; Theory of computation → Regular languages; Theory of computation → Semantics and reasoning; Theory of computation → Automated reasoning; Theory of computation → Complexity theory and logic; Theory of computation → Data provenance

ISBN 978-3-95977-312-6

Published online and open access by

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <https://www.dagstuhl.de/dagpub/978-3-95977-312-6>.

Publication date

March, 2024

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <https://portal.dnb.de>.

License

This work is licensed under a Creative Commons Attribution 4.0 International license (CC-BY 4.0):

<https://creativecommons.org/licenses/by/4.0/legalcode>.



In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.

The copyright is retained by the corresponding authors.

Digital Object Identifier: 10.4230/LIPIcs.ICDT.2024.0

ISBN 978-3-95977-312-6

ISSN 1868-8969

<https://www.dagstuhl.de/lipics>

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 (Reykjavik University, IS and Gran Sasso Science Institute, IT)
- Christel Baier (TU Dresden, DE)
- Roberto Di Cosmo (Inria and Université de Paris, FR)
- Faith Ellen (University of Toronto, CA)
- Javier Esparza (TU München, DE)
- Daniel Král' (Masaryk University, Brno, CZ)
- Meena Mahajan (*Chair*, Institute of Mathematical Sciences, Chennai, IN)
- Anca Muscholl (University of Bordeaux, FR)
- Chih-Hao Luke Ong (University of Oxford, GB)
- Phillip Rogaway (University of California, Davis, US)
- Eva Rotenberg (Technical University of Denmark, Lyngby, DK)
- Raimund Seidel (Universität des Saarlandes, Saarbrücken, DE and Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Wadern, DE)
- Pierre Senellart (ENS, Université PSL, Paris, FR)

ISSN 1868-8969

<https://www.dagstuhl.de/lipics>

■ Contents

Preface	
<i>Graham Cormode and Michael Shekelyan</i>	0:vii
The ICDT 2024 Test of Time Award	
.....	0:ix
Organization	
.....	0:xi
External Reviewers	
.....	0:xiii
Contributing Authors	
.....	0:xv

Invited Talks


Natural Language Data Interfaces: A Data Access Odyssey	
<i>Georgia Koutrika</i>	1:1–1:22
How Database Theory Helps Teach Relational Queries in Database Education	
<i>Sudeepa Roy, Amir Gilad, Yihao Hu, Hanze Meng, Zhengjie Miao,</i> <i>Kristin Stephens-Martinez, and Jun Yang</i>	2:1–2:9
Rule-Based Ontologies: From Semantics to Syntax	
<i>Andreas Pieris</i>	3:1–3:1

Regular Papers

Direct Access for Answers to Conjunctive Queries with Aggregation	
<i>Idan Eldar, Nofar Carmeli, and Benny Kimelfeld</i>	4:1–4:20
Communication Cost of Joins over Federated Data	
<i>Tamara Cucumides and Juan Reutter</i>	5:1–5:19
Range Entropy Queries and Partitioning	
<i>Sanjay Krishnan and Stavros Sintos</i>	6:1–6:21
Skyline Operators for Document Spanners	
<i>Antoine Amarilli, Benny Kimelfeld, Sébastien Labbé, and Stefan Mengel</i>	7:1–7:18
When Do Homomorphism Counts Help in Query Algorithms?	
<i>Balder ten Cate, Victor Dalmau, Phokion G. Kolaitis, and Wei-Lin Wu</i>	8:1–8:20
Approximating Single-Source Personalized PageRank with Absolute Error Guarantees	
<i>Zhewei Wei, Ji-Rong Wen, and Mingji Yang</i>	9:1–9:19
Right-Adjoints for Datalog Programs	
<i>Balder ten Cate, Víctor Dalmau, and Jakub Opršal</i>	10:1–10:20

27th International Conference on Database Theory (ICDT 2024).

Editors: Graham Cormode and Michael Shekelyan

 Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

On the Convergence Rate of Linear Datalog ^o over Stable Semirings <i>Sungjin Im, Benjamin Moseley, Hung Ngo, and Kirk Pruhs</i>	11:1–11:20
Enumeration and Updates for Conjunctive Linear Algebra Queries Through Expressibility <i>Thomas Muñoz Serrano, Cristian Riveros, and Stijn Vansummeren</i>	12:1–12:20
Direct Access for Conjunctive Queries with Negations <i>Florent Capelli and Oliver Irwin</i>	13:1–13:20
The Importance of Parameters in Database Queries <i>Martin Grohe, Benny Kimelfeld, Peter Lindner, and Christoph Standke</i>	14:1–14:17
Conjunctive Queries on Probabilistic Graphs: The Limits of Approximability <i>Antoine Amarilli, Timothy van Bremen, and Kuldeep S. Meel</i>	15:1–15:20
Optimally Rewriting Formulas and Database Queries: A Confluence of Term Rewriting, Structural Decomposition, and Complexity <i>Hubie Chen and Stefan Mengel</i>	16:1–16:17
Containment of Regular Path Queries Under Path Constraints <i>Sylvain Salvati and Sophie Tison</i>	17:1–17:19
Computing Data Distribution from Query Selectivities <i>Pankaj K. Agarwal, Rahul Raychaudhury, Stavros Sintos, and Jun Yang</i>	18:1–18:20
Information Inequality Problem over Set Functions <i>Miika Hannula</i>	19:1–19:20
Conditional Independence on Semiring Relations <i>Miika Hannula</i>	20:1–20:20
Subgraph Enumeration in Optimal I/O Complexity <i>Shiyuan Deng and Yufei Tao</i>	21:1–21:20
Evaluating Graph Queries Using Semantic Treewidth <i>Cristina Feier, Tomasz Gogacz, and Filip Murlak</i>	22:1–22:20
Join Sampling Under Acyclic Degree Constraints and (Cyclic) Subgraph Sampling <i>Ru Wang and Yufei Tao</i>	23:1–23:20
Finding Smallest Witnesses for Conjunctive Queries <i>Xiao Hu and Stavros Sintos</i>	24:1–24:20
Ranked Enumeration for MSO on Trees via Knowledge Compilation <i>Antoine Amarilli, Pierre Bourhis, Florent Capelli, and Mikaël Monet</i>	25:1–25:18

■ Preface

The 27th International Conference on Database Theory (ICDT 2024) was held in Paestum, Italy, from March 25th to March 28th, 2024. The Program Committee has selected 22 research papers out of 74 submissions for publication at the conference.

The PC has further decided to give the best paper award to:

Finding Smallest Witnesses for Conjunctive Queries
Xiao Hu and Stavros Sintos

The ICDT 2024 best newcomer award goes to:

Direct Access for Conjunctive Queries with Negation
Florent Capelli and Oliver Irwin

We congratulate the winners!

Apart from the 22 regular papers, these proceedings include papers accompanying the invited (shared) EDBT/ICDT keynotes by Sudeepa Roy (Duke University) and the ICDT invited talk by Andreas Pieris (University of Edinburgh and University of Cyprus).

A committee formed by Nofar Carmeli, Reinhard Pichler and Nicole Schweikardt has decided to give the Test of Time Award for ICDT 2024 to the ICDT 2014 paper:

Leapfrog Triejoin: A Simple, Worst-Case Optimal Join Algorithm
Todd L. Veldhuizen

We would like to thank all people who contributed to the success of ICDT 2024, 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 2024. Their effort and wisdom were critical 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 program chairs of EDBT, Letizia Tanca and Qiong Luo, for their assistance; and the local organizers of the EDBT/ICDT 2024 conference, led by General Chairs Giuseppe Polese and Loredana Caruccio, 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.

Graham Cormode and Michael Shekelyan
March 2024

■ The ICDT 2024 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 earlier ICDT conferences that have best met the “test of time”. In 2024, the award recognizes a paper selected from the proceedings of the ICDT 2014 conference that has had the highest impact in terms of research, methodology, conceptual contribution, or transfer to practice over the past decade. The award was presented during the EDBT/ICDT 2024 Joint Conference, March 25–28, 2024 in Paestum, Italy.

The 2024 ToT Award Committee consists of Nofar Carmeli, Reinhard Pichler, and Nicole Schweikardt (chair). After careful consideration and soliciting external assessments, the committee has chosen the following contribution for the 2024 ICDT Test-of-Time Award:

Leapfrog Triejoin: A Simple, Worst-Case Optimal Join Algorithm
Todd L. Veldhuizen

This paper introduced the worst-case optimal join algorithm called *Leapfrog Triejoin* (LFTJ). What singles out LFTJ from previously published worst-case optimal join algorithms is that it is very intuitive, simple to describe and easy to implement. It is based on backtracking search, and in contrast to many previously used join algorithms it follows the variable-at-a-time paradigm rather than the classical relation-at-a-time paradigm. Variants of LFTJ have found their way into open-source and commercial database systems, including LogicBlox, Umbra, Kùzu, RelationalAI, FreeJoin, EmptyHeaded, FBENCH, etc.

Apart from its impact on systems, the paper also contains an entirely new technique for proving LFTJ’s worst-case optimality. While earlier proofs for other worst-case optimal join algorithms used a sophisticated entropy-based machinery, Todd L. Veldhuizen presents a self-contained and intuitively simple proof technique. This proof shows that for any database instance on which LFTJ runs in n steps, there exists another database instance with the same relation sizes, where the query answer has size n . This implies that the running time is bounded by the query size on some instance, and therefore it is bounded by the general query size upper bound. Due to its optimality, elegance, and simplicity, the algorithm as well as the proof method today are taught in university courses on the principles of database systems.

In the last decade, LFTJ has greatly influenced theoretical as well as practical database research. For all these reasons, Todd L. Veldhuizen’s ICDT 2014 paper is one of the great gems of ICDT.

Nofar Carmeli
Inria and Univ. Montpellier

Reinhard Pichler
TU Wien

Nicole Schweikardt
HU Berlin

The ICDT Test-of-Time Award Committee for 2024



■ Organization

General Chairs

Giuseppe Polese, University of Salerno

Loredana Caruccio, University of Salerno

Markus Schmid, Humboldt-Universität

Yufei Tao, CUHK

Jef Wijsen, University of Mons

Program Chair

Graham Cormode, University of Warwick

Proceedings Chair

Michael Shekelyan, Queen Mary University of London

Program Committee

Mahmoud Abo Khamis, Relational.AI

Antoine Amarilli, Telecom Paris

Yael Amsterdamer, Bar Ilan University

Pablo Barceló, PUC Chile

Vladimir Braverman, Rice University

Marco Calautti, Università degli Studi di Milano

Hubie Chen, King's College London

Cristina Feier, University of Warsaw

Diego Figueira, Université de Bordeaux

Dominik Freydenberger, Loughborough University

Martin Grohe, RWTH Aachen University

Daniel Kifer, Penn State University

Benny Kimelfeld, Technion

Paraschos Koutris, University of Wisconsin-Madison

Stefan Mengel, CNRS

Frank Neven, Hasselt University

Matthias Niewerth, Bayreuth University

Jeff M. Phillips, University of Utah

Reinhard Pichler, TU Wien

Cristian Riveros, PUC Chile

Francesco Scarcello, Università della Calabria

27th International Conference on Database Theory (ICDT 2024).

Editors: Graham Cormode and Michael Shekelyan



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

■ External Reviewers

Artur Czumaj

Cibele Freire

Cinzia Marté

Davide Mario Longo

Florent Capelli

Hangdong Zhao

Harry Vinnall-Smeeth

Jelle Hellings

Ke Yi

Marcelo Arenas

Nofar Carmeli

Peyman Afshani

Rana Shahout

Rémi Morvan

Sam Thompson

Samson Zhou

Sandeep Silwal

Sudeepa Roy

Xuan Wu

Yuxin Tang

■ Contributing Authors

Pankaj K. Agarwal  (18)

Department of Computer Science, Duke University,
Durham, NC, USA

Antoine Amarilli  (7, 15, 25)

LTCI, Télécom Paris, Institut Polytechnique de Paris, France

Pierre Bourhis  (25)

Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France

Florent Capelli  (13, 25)

Univ. Artois, CNRS, UMR 8188, Centre de Recherche en Informatique de Lens (CRIL), F-62300 Lens, France

Nofar Carmeli  (4)

Inria, LIRMM, Univ Montpellier, CNRS, France

Hubie Chen (16)

Department of Informatics, King's College London, UK

Tamara Cucumides (5)

University of Antwerp, Belgium

Victor Dalmau  (8)

Universitat Pompeu Fabra, Barcelona, Spain

Víctor Dalmau  (10)

Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

Shiyuan Deng (21)


The Chinese University of Hong Kong, China

Idan Eldar  (4)

Technion – Israel Institute of Technology, Haifa, Israel

Cristina Feier (22)

University of Warsaw, Poland

Amir Gilad  (2)


Hebrew University of Jerusalem, Israel

Tomasz Gogacz (22)


University of Warsaw, Poland

Martin Grohe  (14)


RWTH Aachen University, Germany

Miika Hannula  (19, 20)

University of Helsinki, Finland

Xiao Hu  (24)


University of Waterloo, Canada

Yihao Hu  (2)

Duke University, Durham, NC, USA

Sungjin Im (11)

University of California, Merced, CA, USA

Oliver Irwin  (13)


Université de Lille, CNRS, Inria, UMR 9189 – CRISTAL, F-59000 Lille, France

Benny Kimelfeld  (4, 7, 14)

Technion – Israel Institute of Technology, Haifa, Israel

Phokion G. Kolaitis  (8)

University of California Santa Cruz, CA, USA; IBM Almaden Research Center, San Jose, CA, USA

Georgia Koutrika  (1)


Athena Research Center, Athens, Greece

Sanjay Krishnan  (6)


Department of Computer Science, University of Chicago, IL, USA

Sébastien Labbé (7)

École normale supérieure, Paris, France

Peter Lindner  (14)

École Polytechnique Fédérale de Lausanne, Switzerland

Kuldeep S. Meel  (15)

University of Toronto, Canada

Hanze Meng  (2)

Duke University, Durham, NC, USA

Stefan Mengel (7, 16)

Univ. Artois, CNRS, Centre de Recherche en Informatique de Lens (CRIL), France

Zhengjie Miao  (2)

Simon Fraser University, Burnaby, Canada

Mikaël Monet  (25)

Université de Lille, CNRS, Inria, UMR 9189 – CRISTAL, F-59000 Lille, France

Benjamin Moseley (11)

Carnegie Mellon University, Pittsburgh, PA, USA

Filip Murlak (22)

University of Warsaw, Poland


27th International Conference on Database Theory (ICDT 2024).

Editors: Graham Cormode and Michael Shekelyan





Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Thomas Muñoz Serrano  (12)
UHasselt, Data Science Institute,
Diepenbeek, Belgium

Hung Ngo (11)
RelationalAI, Berkeley, CA, USA


Jakub Opršal  (10)
School of Computer Science,
University of Birmingham, UK


Andreas Pieris  (3)
University of Edinburgh, UK;
University of Cyprus, Nicosia, Cyprus


Kirk Pruhs (11)
University of Pittsburgh, Pittsburgh, PA, USA


Rahul Raychaudhury (18)
Department of Computer Science, Duke
University, Durham, NC, USA


Juan Reutter (5)
PUC Chile & IMFD Chile, Santiago, Chile


Cristian Riveros  (12)
Pontificia Universidad Católica de Chile,
Santiago, Chile; Millennium Institute for
Foundational Research on Data, Santiago, Chile

Sudeepa Roy  (2)
Duke University, Durham, NC, USA;
RelationalAI, Berkeley, CA, USA
(Visiting Scientist)

Sylvain Salvati  (17)
Université de Lille, INRIA, CRISTAL/CNRS
UMR 9189, France


Stavros Sintos  (6, 18, 24)
Department of Computer Science,
University of Illinois at Chicago, IL, USA


Christoph Standke  (14)
RWTH Aachen University, Germany

Kristin Stephens-Martinez  (2)
Duke University, Durham NC, USA

Yufei Tao (21, 23)
The Chinese University of Hong Kong, China


Balder ten Cate  (8, 10)
University of Amsterdam, The Netherlands


Sophie Tison  (17)
Université de Lille, INRIA, CRISTAL/CNRS
UMR 9189, France


Timothy van Bremen  (15)
National University of Singapore, Singapore


Stijn Vansummeren  (12)
UHasselt, Data Science Institute,
Diepenbeek, Belgium


Ru Wang (23)
The Chinese University of Hong Kong, China

Zhewei Wei  (9)
Renmin University of China, Beijing, China

Ji-Rong Wen  (9)
Renmin University of China, Beijing, China

Wei-Lin Wu  (8)
University of California Santa Cruz, CA, USA

Jun Yang  (2, 18)
Duke University, Durham, NC, USA

Mingji Yang  (9)
Renmin University of China, Beijing, China