34th Annual Symposium on Combinatorial Pattern Matching

CPM 2023, June 26–28, 2023, Marne-la-Vallée, France

Edited by

Laurent Bulteau

Zsuzsanna Lipták

Editors

Laurent Bulteau
LIGM, CNRS, Université Gustave Eiffel, Marne-la-vallée, France
laurent.bulteau@univ-eiffel.fr

Zsuzsanna Lipták
University of Verona, Italy
zsuzsanna.liptak@univr.it

ACM Classification 2012
Theory of computation → Design and analysis of algorithms; Theory of computation → Pattern matching; Theory of computation → Data compression; Mathematics of computing → Discrete mathematics; Mathematics of computing → Combinatorics on words; Mathematics of computing → Combinatoric problems; Applied computing → Computational biology

ISBN 978-3-95977-276-1

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/978-3-95977-276-1.

Publication date
June, 2023

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.CPM.2023.0

ISBN 978-3-95977-276-1 **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 (*Chair*, Reykjavik University, IS and Gran Sasso Science Institute, IT)
- Christel Baier (TU Dresden, DE)
- Mikolaj Bojanczyk (University of Warsaw, PL)
- 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 (Institute of Mathematical Sciences, Chennai, IN)
- Anca Muscholl (University of Bordeaux, FR)
- Chih-Hao Luke Ong (University of Oxford, GB and Nanyang Technological University, SG)
- 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)

ISSN 1868-8969

https://www.dagstuhl.de/lipics

To all men and women whose passion are algorithms, data structures, and combinatorics on strings.

Contents

Preface
Laurent Bulteau and Zsuzsanna Lipták 0:ix

Program Committee
..... 0:xi

External Reviewers
..... 0:xiii

List of Authors
..... 0:xv–0:xvi

Regular Papers

Trie-Compressed Adaptive Set Intersection
Diego Arroyuelo and Juan Pablo Castillo 1:1–1:19

Approximation Algorithms for the Longest Run Subsequence Problem
Yuichi Asahiro, Hiroshi Eto, Mingyang Gong, Jesper Jansson, Guohui Lin, Eiji Miyano, Hirotaka Ono, and Shunichi Tanaka 2:1–2:12

Optimal LZ-End Parsing Is Hard
Hideo Bannai, Mitsuru Funakoshi, Kazuhiro Kurita, Yuto Nakashima, Kazuhisa Seto, and Takeaki Uno 3:1–3:11

Sliding Window String Indexing in Streams
Philip Bille, Johannes Fischer, Inge Li Gørtz, Max Rishøj Pedersen, and Tord Joakim Stordalen 4:1–4:18

Faster Algorithms for Computing the Hairpin Completion Distance and Minimum Ancestor
Itai Boneh, Dvir Fried, Adrian Miclăuş, and Alexandru Popa 5:1–5:18

On Distances Between Words with Parameters
Pierre Bourhis, Aaron Boussidan, and Philippe Gambette 6:1–6:23

Parameterized Algorithms for String Matching to DAGs: Funnels and Beyond
Manuel Cáceres 7:1–7:19

Optimal Near-Linear Space Heaviest Induced Ancestors
Panagiotis Charalampopoulos, Bartłomiej Dudek, Paweł Gawrychowski, and Karol Pokorski 8:1–8:18

From Bit-Parallelism to Quantum String Matching for Labelled Graphs
Massimo Equi, Arianne Meijer-van de Griend, and Veli Mäkinen 9:1–9:20

On the Impact of Morphisms on BWT-Runs
Gabriele Fici, Giuseppe Romana, Marinella Sciortino, and Cristian Urbina 10:1–10:18

34th Annual Symposium on Combinatorial Pattern Matching (CPM 2023).
Editors: Laurent Bulteau and Zsuzsanna Lipták
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Comparing Elastic-Degenerate Strings: Algorithms, Lower Bounds, and Applications
Esteban Gabory, Moses Njagi Mwaniki, Nadia Pisanti, Solon P. Pissis, Jakub Radoszewski, Michelle Sweering, and Wiktor Zuba 11:1–11:20

Compressed Indexing for Consecutive Occurrences
Paweł Gawrychowski, Garance Gourdel, Tatiana Starikovskaya, and Teresa Anna Steiner 12:1–12:22

Order-Preserving Squares in Strings
Paweł Gawrychowski, Samah Ghazawi, and Gad M. Landau 13:1–13:19

MUL-Tree Pruning for Consistency and Compatibility
Christopher Hampson, Daniel J. Harvey, Costas S. Iliopoulos, Jesper Jansson, Zara Lim, and Wing-Kin Sung 14:1–14:18

Linear-Time Computation of Cyclic Roots and Cyclic Covers of a String
Costas S. Iliopoulos, Tomasz Kociumaka, Jakub Radoszewski, Wojciech Rytter, Tomasz Waleń, and Wiktor Zuba 15:1–15:15

Faster Prefix-Sorting Algorithms for Deterministic Finite Automata
Sung-Hwan Kim, Francisco Olivares, and Nicola Prezza 16:1–16:16

Encoding Hard String Problems with Answer Set Programming
Dominik Köppl 17:1–17:21

On the Complexity of Parameterized Local Search for the Maximum Parsimony Problem
Christian Komusiewicz, Simone Linz, Nils Morawietz, and Jannik Schestag 18:1–18:18

String Factorization via Prefix Free Families
Matan Kraus, Moshe Lewenstein, Alexandru Popa, Ely Porat, and Yonathan Sadia 19:1–19:10

Improving the Sensitivity of MinHash Through Hash-Value Analysis
Gregory Kucherov and Steven Skiena 20:1–20:12

Suffix-Prefix Queries on a Dictionary
Grigorios Loukides, Solon P. Pissis, Sharma V. Thankachan, and Wiktor Zuba .. 21:1–21:20

Merging Sorted Lists of Similar Strings
Gene Myers 22:1–22:15

PalFM-Index: FM-Index for Palindrome Pattern Matching
Shinya Nagashita and Tomohiro I 23:1–23:15

Computing MEMs on Repetitive Text Collections
Gonzalo Navarro 24:1–24:17

L-Systems for Measuring Repetitiveness
Gonzalo Navarro and Cristian Urbina 25:1–25:17

MONI Can Find $\boldsymbol{k}$-MEMs
Igor Tatarnikov, Ardavan Shahrabi Farahani, Sana Kashgouli, and Travis Gagie .. 26:1–26:14

Preface

The Annual Symposium on Combinatorial Pattern Matching (CPM) has by now over 30 years of tradition and is considered to be the leading conference for the community working on Stringology. The objective of the annual CPM meetings is to provide an international forum for research in combinatorial pattern matching and related applications such as computational biology, data compression and data mining, coding, information retrieval, natural language processing, and pattern recognition.

This volume contains the papers presented at the 34th Annual Symposium on Combinatorial Pattern Matching (CPM 2023) held on June 26–28, 2023 in Marne-la-Vallée, France. The conference program includes 26 contributed papers and three invited talks, by

- Olgica Milenkovic (University of Illinois Urbana-Champaign, USA),
- Tatiana Starikovskaya (École Normale Supérieure, France), and
- Virginia Vassilevska Williams (Massachusetts Institute of Technology, USA).

For the fifth time, CPM includes the "Highlights of CPM" special session, for presenting the highlights of recent developments in combinatorial pattern matching. In this fifth edition we selected as highlight papers "Separating words and trace reconstruction", by Zachary Chase, presented at STOC 2021, and "Approximating Dynamic Time Warping Distance Between Run-Length Encoded Strings", by Zoe Xi and William Kuszmaul, presented at ESA 2022. The conference was preceded by a one-day student summer school taught by Karel Břinda (Inria/IRISA Rennes) and Panagiotis Charalampopoulos (Birkbeck, University of London), and organized at the École Normale Supérieure in Paris by Gabriel Bathie, Paweł Gawrychowski, Garance Gourdel and Tatiana Starikovskaya.

The contributed papers were selected out of 44 submissions, corresponding to an acceptance ratio of 59%. Each submission received at least three reviews. We thank the members of the Program Committee and all the additional external subreviewers, who are listed below, for their hard, invaluable, and collaborative effort that resulted in an excellent scientific program.

The Annual Symposium on Combinatorial Pattern Matching started in 1990, and has since then taken place every year. Previous CPM meetings were held in Paris, London (UK), Tucson, Padova, Asilomar, Helsinki, Laguna Beach, Aarhus, Piscataway, Warwick, Montreal, Jerusalem, Fukuoka, Morelia, Istanbul, Jeju Island, Barcelona, London (Ontario, Canada), Pisa, Lille, New York, Palermo, Helsinki, Bad Herrenalb, Moscow, Ischia, Tel Aviv, Warsaw, Qingdao, Pisa, Copenhagen (on-line), Wrocław and Prague. From 1992 to the 2015 meeting, all proceedings were published in the LNCS (Lecture Notes in Computer Science) series. Since 2016, the CPM proceedings have appeared in the LIPIcs (Leibniz International Proceedings in Informatics) series, as volume 54 (CPM 2016), 78 (CPM 2017), 105 (CPM 2018), 128 (CPM 2019), 161 (CPM 2020), 191 (CPM 2021) and 223 (CPM 2022). The entire submission and review process was carried out using the EasyChair conference system.

We thank the CPM Steering Committee for their support and advice.

Laurent Bulteau and Zsuzsanna Lipták

34th Annual Symposium on Combinatorial Pattern Matching (CPM 2023).
Editors: Laurent Bulteau and Zsuzsanna Lipták
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Program Committee

Golnaz Badkobeh
Goldsmiths, University of London, UK

Hideo Bannai
Tokyo Medical and Dental University, Japan

Marie-Pierre Béal
University Gustave Eiffel, France

Giulia Bernardini
University of Trieste, Italy

Paola Bonizzoni
University of Milano-Bicocca, Italy

Laurent Bulteau (co-chair)
CNRS – University Gustave Eiffel, France

Johannes Fischer
TU Dortmund, Germany

Travis Gagie
Dalhousie University, Canada

Pawel Gawrychowski
University of Wroclaw, Poland

Jan Holub
Czech Technical University in Prague, Czech Republic

Wing-Kai Hon
National Tsing Hua University, Taiwan

Dominik Kempa
Stony Brook University, USA

Tomasz Kociumaka
Max Planck Institute for Informatics, Germany

Avivit Levy
Shenkar College of Engineering, Design and Art, Israel

Moshe Lewenstein
Bar Ilan University, Israel

Zsuzsanna Lipták (co-chair)
University of Verona, Italy

Yuto Nakashima
Kyushu University, Japan

Gonzalo Navarro
University of Chile, Chile

Enno Ohlebusch
University of Ulm, Germany

Svetlana Puzynina
Saint Petersburg State University, Russia

Jakub Radoszewski
University of Warsaw, Poland

Rajeev Raman
University of Leicester, UK

Giovanna Rosone
University of Pisa, Italy

Leena Salmela
University of Helsinki, Finland

Joe Sawada
University of Guelph, Canada

Marinella Sciortino
University of Palermo, Italy

Florian Sikora
Université Paris-Dauphine, PSL University, France

Teresa Anna Steiner
Technical University of Denmark, Denmark

Sharma V. Thankachan
North Carolina State University, USA

Oren Weimann
University of Haifa, Israel

Petra Wolf
University of Bergen, Norway

Wiktor Zuba
CWI, Netherlands

34th Annual Symposium on Combinatorial Pattern Matching (CPM 2023).
Editors: Laurent Bulteau and Zsuzsanna Lipták

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

34th Annual Symposium on Combinatorial Pattern Matching (CPM 2023).
Editors: Laurent Bulteau and Zsuzsanna Lipták

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

External Reviewers

Emmanuel Arrighi
Nathalie Aubrun
Guillaume Blin
Manuel Cáceres
Giuseppa Castiglione
Julien Clément
Sami Davies
Joel Day
Gianluca Della Vedova
Diego Diaz
Riccardo Dondi
Jonas Ellert
Estéban Gabory
Daniel Gabric
Martin Gebser
Samah Ghazawi
Daniel Gibney
Daniel Gibney
Shay Golan
Garance Gourdel
Roberto Grossi
Veronica Guerrini
Wojciech Janczewski
Mark Jones
Saira Karim
Steven Kelk
Dominik Köppl
Dmitry Kosolobov
Thierry Lecroq
Florin Manea
Nils Morawietz
Nicola Prezza
Narad Rampersad
Raffaella Rizzi
Šimon Schierreich
Markus L. Schmid
Tatiana Starikovskaya
Tord Stordalen
Michelle Sweering
Alexander Tiskin
Jacobo Toran
Rossano Venturini
Hilde Verbeek
Tomasz Walen
Dennis Wong

34th Annual Symposium on Combinatorial Pattern Matching (CPM 2023).
Editors: Laurent Bulteau and Zsuzsanna Lipták

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

List of Authors

Diego Arroyuelo
Yuichi Asahiro
Hideo Bannai
Philip Bille
Itai Boneh
Pierre Bourhis
Aaron Boussidan
Manuel Cáceres
Juan Pablo Castillo
Panagiotis Charalampopoulos
Bartlomiej Dudek
Massimo Equi
Hiroshi Eto
Gabriele Fici
Johannes Fischer
Dvir Fried
Mitsuru Funakoshi
Estéban Gabory
Travis Gagie
Philippe Gambette
Pawel Gawrychowski
Samah Ghazawi
Mingyang Gong
Inge Li Gørtz
Garance Gourdel
Christopher Hampson
Daniel J. Harvey
Tomohiro I
Costas S. Iliopoulos
Jesper Jansson
Sana Kashgouli
Sung-Hwan Kim
Tomasz Kociumaka
Christian Komusiewicz
Dominik Köppl
Matan Kraus
Gregory Kucherov
Kazuhiro Kurita
Gad M. Landau
Moshe Lewenstein
Zara Lim
Guohui Lin
Simone Linz
Grigorios Loukides
Veli Mäkinen
Arianne Meijer van de Griend
Adrian Miclăuș
Eiji Miyano
Nils Morawietz
Moses Njagi Mwaniki
Eugene Myers
Shinya Nagashita
Yuto Nakashima
Gonzalo Navarro
Francisco Olivares
Hirotaka Ono
Max Rishøj Pedersen
Nadia Pisanti
Solon Pissis
Karol Pokorski
Alexandru Popa
Ely Porat

Nicola Prezza

Jakub Radoszewski

Giuseppe Romana

Wojciech Rytter

Yonathan Sadia

Jannik Schestag

Marinella Sciortino

Kazuhisa Seto

Ardavan Shahrabi Farahani

Steven Skiena

Tatiana Starikovskaya

Teresa Anna Steiner

Tord Joakim Stordalen

Wing-Kin Sung

Michelle Sweering

Shunichi Tanaka

Igor Tatarnikov

Sharma V. Thankachan

Takeaki Uno

Cristian Urbina

Tomasz Waleń

Wiktor Zuba