

35th Annual Symposium on Combinatorial Pattern Matching

CPM 2024, June 25–27, 2024, Fukuoka, Japan

Edited by

Shunsuke Inenaga

Simon J. Puglisi



Editors

Shunsuke Inenaga 

Kyushu University, Japan
inenaga.shunsuke.380@m.kyushu-u.ac.jp

Simon J. Puglisi 

University of Helsinki, Finland
simon.puglisi@helsinki.fi

ACM Classification 2012

Theory of computation → Algorithm design techniques; Theory of computation → Data compression; Theory of Computation → Design and analysis of algorithms; Theory of computation → Fixed parameter tractability; Theory of computation → Parameterized complexity and exact algorithms; Theory of computation → Pattern matching; Theory of computation → Problems, reductions and completeness; Theory of computation → Sorting and searching; Theory of computation → Theory and algorithms for application domains; Mathematics of computing → Discrete mathematics; Mathematics of computing → Combinatorics on words; Mathematics of computing → Combinatoric problems; Applied computing → Computational biology

ISBN 978-3-95977-326-3

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/978-3-95977-326-3>.

Publication date

June, 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.CPM.2024.0

ISBN 978-3-95977-326-3

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 Paris Cité, 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 (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)
- Pierre Senellart (ENS, Université PSL, Paris, FR)

ISSN 1868-8969

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

Dedicated to our inspirational friend, colleague, and teacher, Masayuki Takeda,
1964–2022.

■ Contents


Preface	
<i>Shunsuke Inenaga and Simon J. Puglisi</i>	0:ix–0:x
Program Committee	
.....	0:xi
External Reviewers	
.....	0:xiii

Regular Papers

Computing the LCP Array of a Labeled Graph	
<i>Jarno N. Alanko, Davide Cenzato, Nicola Cotumaccio, Sung-Hwan Kim,</i> <i>Giovanni Manzini, and Nicola Prezza</i>	1:1–1:15
Reconstructing General Matching Graphs	
<i>Amihood Amir and Michael Itzhaki</i>	2:1–2:15
Maintaining the Size of LZ77 on Semi-Dynamic Strings	
<i>Hideo Bannai, Panagiotis Charalampopoulos, and Jakub Radoszewski</i>	3:1–3:20
Internal Pattern Matching in Small Space and Applications	
<i>Gabriel Bathie, Panagiotis Charalampopoulos, and Tatiana Starikovskaya</i>	4:1–4:20
Random Wheeler Automata	
<i>Ruben Becker, Davide Cenzato, Sung-Hwan Kim, Bojana Kodric,</i> <i>Riccardo Maso, and Nicola Prezza</i>	5:1–5:15
Connecting de Bruijn Graphs	
<i>Giulia Bernardini, Huiping Chen, Inge Li Gørtz, Christoffer Krogh,</i> <i>Grigorios Loukides, Solon P. Pissis, Leen Stougie, and Michelle Sweering</i>	6:1–6:16
A Class of Heuristics for Reducing the Number of BWT-Runs in the String Ordering Problem	
<i>Gianmarco Bertola, Anthony J. Cox, Veronica Guerrini, and Giovanna Rosone</i> ..	7:1–7:15
Faster Sliding Window String Indexing in Streams	
<i>Philip Bille, Pawel Gawrychowski, Inge Li Gørtz, and Simon R. Tarnow</i>	8:1–8:14
Tight Bounds for Compressing Substring Samples	
<i>Philip Bille, Christian Mikkelsen Fuglsang, and Inge Li Gørtz</i>	9:1–9:14
Searching 2D-Strings for Matching Frames	
<i>Itai Boneh, Dvir Fried, Shay Golan, Matan Kraus, Adrian Miclăuș, and</i> <i>Arseny Shur</i>	10:1–10:19
Hairpin Completion Distance Lower Bound	
<i>Itai Boneh, Dvir Fried, Shay Golan, and Matan Kraus</i>	11:1–11:16
Solving the Minimal Positional Substring Cover Problem in Sublinear Space	
<i>Paola Bonizzoni, Christina Boucher, Davide Cozzi, Travis Gagie, and Yuri Pirola</i>	12:1–12:16

35th Annual Symposium on Combinatorial Pattern Matching (CPM 2024).

Editors: Shunsuke Inenaga and Simon J. Puglisi

 Leibniz International Proceedings in Informatics

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

Online Context-Free Recognition in OMv Time <i>Bartłomiej Dudek and Paweł Gawrychowski</i>	13:1–13:9
When Is the Normalized Edit Distance over Non-Uniform Weights a Metric? <i>Dana Fisman and Ilay Tzarfat</i>	14:1–14:17
Efficient Construction of Long Orientable Sequences <i>Daniel Gabrić and Joe Sawada</i>	15:1–15:12
Exploiting New Properties of String Net Frequency for Efficient Computation <i>Peaker Guo, Patrick Eades, Anthony Wirth, and Justin Zobel</i>	16:1–16:16
Closing the Gap: Minimum Space Optimal Time Distance Labeling Scheme for Interval Graphs <i>Meng He and Kaiyu Wu</i>	17:1–17:18
Algorithms for Galois Words: Detection, Factorization, and Rotation <i>Diptarama Hendrian, Dominik Köppl, Ryo Yoshinaka, and Ayumi Shinohara</i>	18:1–18:16
Simplified Tight Bounds for Monotone Minimal Perfect Hashing <i>Dmitry Kosolobov</i>	19:1–19:13
Construction of Sparse Suffix Trees and LCE Indexes in Optimal Time and Space <i>Dmitry Kosolobov and Nikita Sivukhin</i>	20:1–20:18
BAT-LZ out of hell <i>Zsuzsanna Lipták, Francesco Masillo, and Gonzalo Navarro</i>	21:1–21:17
Subsequences with Generalised Gap Constraints: Upper and Lower Complexity Bounds <i>Florin Manea, Jonas Richardsen, and Markus L. Schmid</i>	22:1–22:17
The <i>Rational</i> Construction of a Wheeler DFA <i>Giovanni Manzini, Alberto Policriti, Nicola Prezza, and Brian Ricciardi</i>	23:1–23:15
Shortest Cover After Edit <i>Kazuki Mitani, Takuya Mieno, Kazuhisa Seto, and Takashi Horiyama</i>	24:1–24:15
Walking on Words <i>Ian Pratt-Hartmann</i>	25:1–25:17
A Data Structure for the Maximum-Sum Segment Problem with Offsets <i>Yoshifumi Sakai</i>	26:1–26:15
Finding Diverse Strings and Longest Common Subsequences in a Graph <i>Yuto Shida, Giulia Punzi, Yasuaki Kobayashi, Takeaki Uno, and Hiroki Arimura</i> .	27:1–27:19
Minimizing the Minimizers via Alphabet Reordering <i>Hilde Verbeek, Lorraine A.K. Ayad, Grigorios Loukides, and Solon P. Pissis</i>	28:1–28:13

■ 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 image processing (i.e. 2D strings).

This volume contains the papers presented at the 35th Annual Symposium on Combinatorial Pattern Matching (CPM 2024) held on June 25–27, 2024 in Fukuoka, Japan. The conference program includes 28 contributed papers and three invited talks, by

- Martin Farach-Colton (New York University, USA),
- Zsuzsanna Lipták (University of Verona, Italy), and
- Tetsuo Shibuya (University of Tokyo, Japan).

For the sixth time, CPM includes the “Highlights of CPM” special session, for presenting the highlights of recent developments in combinatorial pattern matching. In this sixth installment we selected as highlight papers “Gapped String Indexing in Subquadratic Space and Sublinear Query Time”, by Philip Bille, presented at STACS 2024, and “Optimal-Time Queries on BWT-Runs Compressed Indexes” and “An Optimal-Time RLBWT Construction in BWT-Runs Bounded Space”, by Yasuo Tabei / Takaaki Nishimoto, presented at ICALP 2021 and ICALP 2022, respectively.

The contributed papers for CPM 2024 were selected out of 48 submissions, corresponding to an acceptance ratio of 58%. 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. We also thank the CPM Steering Committee for their support and advice. We thank the CPM 2024 Organising Committee chair Yuto Nakashima and all the other members for their excellent and hard work that made this conference a wonderful one.

The CPM 2024 conference was co-located with a StringMasters workshop held on June 24 and June 28, 2024 in Fukuoka. The StringMasters workshop was organized by Dominik Köppl. Preceding these events in Fukuoka, a two-day summer school was held at the University of Electro-Communications, in Tokyo, which provided lectures by Jesper Jansson (Kyoto University) and Philip Wellnitz (National Institute of Informatics). The summer school was organized by Paweł Gawrychowski, Hideo Bannai, and Takuya Meino.

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, Barcelona, London (Ontario, Canada), Pisa, Lille, New York, Palermo, Helsinki, Bad Herrenalb, Moscow, Ischia, Tel Aviv, Warsaw, Qingdao, Pisa, Copenhagen (on-line), Wrocław, Prague, and Marne-la-Vallée. 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), 223 (CPM 2022), and 259 (CPM 2023). The entire submission and review process was carried out using the EasyChair conference system.

35th Annual Symposium on Combinatorial Pattern Matching (CPM 2024).

Editors: Shunsuke Inenaga and Simon J. Puglisi



Leibniz International Proceedings in Informatics

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

This proceedings of CPM 2024 is dedicated to the memory of Masayuki Takeda, who passed away in December 2022, in Fukuoka. As a professor at the Department of Informatics, Kyushu University, Masayuki was one of the pioneers of the field of pattern matching and compressed string processing in Japan. He was a passionate researcher, and published over 200 articles, including 34 papers in the CPM series of conferences – more than all but one other author. Masayuki also co-chaired CPM 2002 in Fukuoka with late Alberto Apostolico. Masayuki was a great teacher, and he supervised a number of students at the undergraduate, master, and PhD levels, many of whom have gone on to have successful research careers. Implicitly, and explicitly, he is the reason the CPM conference returns to Fukuoka this year.

Shunsuke Inenaga and Simon J. Puglisi
CPM 2024 Program Committee Chairs

■ Program Committee

- Golnaz Badkobeh, City, University of London
- Giulia Bernardini, University of Trieste
- Philip Bille, Technical University of Denmark
- Manuel Cáceres, University of Helsinki
- Panagiotis Charalampopoulos, Birkbeck, University of London
- Gabriele Fici, University of Palermo
- Daniel Gibney, University of Texas at Dallas
- Veronica Guerrini, University of Pisa
- Shunsuke Inenaga, Kyushu University (Chair)
- Jesper Jansson, Kyoto University
- Artur Jeż, University of Wrocław
- Dominik Köppl, University of Yamanashi
- Gregory Kucherov, Gustave Eiffel University
- Susana Ladra, University of A Coruna
- Avivit Levy, Shenkar College
- Gonzalo Navarro, University of Chile
- Nicola Prezza, Ca' Foscari University of Venice
- Simon J. Puglisi, University of Helsinki (Chair)
- Giulia Punzi, NII
- Marinella Sciortino, Università di Palermo
- Kana Shimizu, Waseda University

■ List of External Reviewers

Adrián Gómez Brandón
Alejandro Pacheco
Alessandro Berti
Alexandru I. Tomescu
Arnab Ganguly
B. Riva Shalom
Bartłomiej Dudek
Binhai Zhu
Bojana Kodric
Carlo Tosoni
Chiaki Sakama
Chiara Epifanio
Conrado Martínez
Cristian Urbina
Cyril Gavoille
Daniel Puttini
Davide Rucci
Diego Arroyuelo
Diego Seco
Diptarama Hendrian
Djamal Belazzougui
Dominika Draesslerová
Duncan Adamson
Dustin Cobas
Eitan Kondratovsky
Elena Biagi
Estéban Gabory
Felipe A. Louza
Florin Manea
Francisco Olivares
Gabriel Bathie
Giovanna Rosone
Giuseppa Castiglione
Hideo Bannai
Hongyu Zheng
Itai Boneh
Jakub Radoszewski
Jamshed Khan
Jonas Ellert
Jouni Sirén
Kevin Ryan
Laurent Bulteau
Louisa Seelbach Benkner
Massimo Equi
Md Helal Hossen
Meng He
Mikołaj Marciniak
Nicola Rizzo
Nicolaos Matsakis
Paniz Abedin
Rahul Shah
Rocco Ascone
Romina Doz
Ruben Becker
Sharma V. Thankachan
Shay Golan
Simon Rumle Tarnow
Solon Pissis
Sung-Hwan Kim
Sven Rahmann
Takuya Mieno
Teresa Anna Steiner
Tomasz Walen
Travis Gagie
Vincent Jugé
Vitaly Aksenov
Wiktor Zuba
William Kuszmaul
Wojciech Janczewski
Yinzhan Xu
Yuto Nakashima



