

28th Annual Symposium on Combinatorial Pattern Matching

CPM 2017, July 4–6, 2017, Warsaw, Poland

Edited by

Juha Kärkkäinen

Jakub Radoszewski

Wojciech Rytter



Editors

Juha Kärkkäinen
Department of Computer Science

University of Helsinki, Finland
juha.karkkainen@cs.helsinki.fi

Jakub Radoszewski
Faculty of Mathematics,
Informatics and Mechanics
University of Warsaw, Poland
jrad@mimuw.edu.pl

Wojciech Rytter
Faculty of Mathematics,
Informatics and Mechanics
University of Warsaw, Poland
rytter@mimuw.edu.pl

ACM Classification 1998

E.1 Data Structures, E.2 Data Storage Representations, E.4 Coding and Information Theory, F. Theory of Computation, G.2 Discrete Mathematics, H. Information Systems, I.7 Document and Text Processing

ISBN 978-3-95977-039-2

Published online and open access by

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <http://www.dagstuhl.de/dagpub/978-3-95977-039-2>.

Publication date

July, 2017

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 <http://dnb.d-nb.de>.

License

This work is licensed under a Creative Commons Attribution 3.0 Unported license (CC-BY 3.0): <http://creativecommons.org/licenses/by/3.0/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.2017.0

ISBN 978-3-95977-039-2

ISSN 1868-8969

<http://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)
- Susanne Albers (TU München)
- Chris Hankin (Imperial College London)
- Deepak Kapur (University of New Mexico)
- Michael Mitzenmacher (Harvard University)
- Madhavan Mukund (Chennai Mathematical Institute)
- Anca Muscholl (University Bordeaux)
- Catuscia Palamidessi (INRIA)
- Raimund Seidel (Saarland University and Schloss Dagstuhl – Leibniz-Zentrum für Informatik)
- Thomas Schwentick (TU Dortmund)
- Reinhard Wilhelm (Saarland University)

ISSN 1868-8969

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

To all algorithmic stringologists in the world

■ Contents

Preface	
<i>Juha Kärkkäinen, Jakub Radoszewski, and Wojciech Rytter</i>	0:ix

Invited Talks

Wheeler Graphs: Variations on a Theme by Burrows and Wheeler	
<i>Giovanni Manzini</i>	1:1–1:1
Recompression of SLPs	
<i>Artur Jeż</i>	2:1–2:1
Shortest Superstring	
<i>Marcin Mucha</i>	3:1–3:1

Regular Papers

Document Listing on Repetitive Collections with Guaranteed Performance	
<i>Gonzalo Navarro</i>	4:1–4:13
Path Queries on Functions	
<i>Travis Gagie, Meng He, and Gonzalo Navarro</i>	5:1–5:15
Deterministic Indexing for Packed Strings	
<i>Philip Bille, Inge Li Gørtz, and Frederik Rye Skjoldjensen</i>	6:1–6:11
Representing the Suffix Tree with the CDAWG	
<i>Djamal Belazzougui and Fabio Cunial</i>	7:1–7:13
Position Heaps for Parameterized Strings	
<i>Diptarama, Takashi Katsura, Yuhei Otomo, Kazuyuki Narisawa, and Ayumi Shinohara</i>	8:1–8:13
On-Line Pattern Matching on Similar Texts	
<i>Roberto Grossi, Costas S. Iliopoulos, Chang Liu, Nadia Pisanti, Solon P. Pissis, Ahmad Retha, Giovanna Rosone, Fatima Vayani, and Luca Versari</i>	9:1–9:14
A Family of Approximation Algorithms for the Maximum Duo-Preservation String Mapping Problem	
<i>Bartłomiej Dudek, Paweł Gawrychowski, and Piotr Ostropolski-Nalewaja</i>	10:1–10:14
Revisiting the Parameterized Complexity of Maximum-Duo Preservation String Mapping	
<i>Christian Komusiewicz, Mateus de Oliveira Oliveira, and Meirav Zehavi</i>	11:1–11:17
Clique-Based Lower Bounds for Parsing Tree-Adjoining Grammars	
<i>Karl Bringmann and Philip Wellnitz</i>	12:1–12:14
Communication and Streaming Complexity of Approximate Pattern Matching	
<i>Tatiana Starikovskaya</i>	13:1–13:11

28th Annual Symposium on Combinatorial Pattern Matching (CPM 2017).

Editors: Juha Kärkkäinen, Jakub Radoszewski, and Wojciech Rytter

Leibniz International Proceedings in Informatics



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

The Longest Filled Common Subsequence Problem <i>Mauro Castelli, Riccardo Dondi, Giancarlo Mauri, and Italo Zoppis</i>	14:1–14:13
Lempel-Ziv Compression in a Sliding Window <i>Philip Bille, Patrick Hagge Cording, Johannes Fischer, and Inge Li Gørtz</i>	15:1–15:11
Time-Space Trade-Offs for Lempel-Ziv Compressed Indexing <i>Philip Bille, Mikko Berggren Ettienne, Inge Li Gørtz, and Hjalte Wedel Vildhøj</i> ..	16:1–16:17
From LZ77 to the Run-Length Encoded Burrows-Wheeler Transform, and Back <i>Alberto Policriti and Nicola Prezza</i>	17:1–17:10
Longest Common Extensions with Recompression <i>Tomohiro I</i>	18:1–18:15
Fast and Simple Jumbled Indexing for Binary Run-Length Encoded Strings <i>Luís Cunha, Simone Dantas, Travis Gagie, Roland Wittler, Luis Kowada, and Jens Stoye</i>	19:1–19:9
Faster STR-IC-LCS Computation via RLE <i>Keita Kuboi, Yuta Fujishige, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda</i>	20:1–20:12
Gapped Pattern Statistics <i>Philippe Duchon, Cyril Nicaud, and Carine Pivoteau</i>	21:1–21:12
Computing All Distinct Squares in Linear Time for Integer Alphabets <i>Hideo Bannai, Shunsuke Inenaga, and Dominik Köppl</i>	22:1–22:18
Palindromic Length in Linear Time <i>Kirill Borozdin, Dmitry Kosolobov, Mikhail Rubinchik, and Arseny M. Shur</i>	23:1–23:12
Tight Bounds on the Maximum Number of Shortest Unique Substrings <i>Takuya Mieno, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda</i>	24:1–24:11
Can We Recover the Cover? <i>Amihood Amir, Avivit Levy, Moshe Lewenstein, Ronit Lubin, and Benny Porat</i> ...	25:1–25:15
Approximate Cover of Strings <i>Amihood Amir, Avivit Levy, Ronit Lubin, and Ely Porat</i>	26:1–26:14
Beyond Adjacency Maximization: Scaffold Filling for New String Distances <i>Laurent Bulteau, Guillaume Fertin, and Christian Komusiewicz</i>	27:1–27:17
On the Weighted Quartet Consensus Problem <i>Manuel Lafond and Celine Scornavacca</i>	28:1–28:18
Optimal Omnitig Listing for Safe and Complete Contig Assembly <i>Massimo Cairo, Paul Medvedev, Nidia Obscura Acosta, Romeo Rizzi, and Alexandru I. Tomescu</i>	29:1–29:12
Dynamic Elias-Fano Representation <i>Giulio Ermanno Pibiri and Rossano Venturini</i>	30:1–30:14
Synergistic Solutions on MultiSets <i>Jérémy Barbay, Carlos Ochoa, and Srinivasa Rao Satti</i>	31:1–31:14

■ Preface

The Annual Symposium on Combinatorial Pattern Matching is an international forum for research in combinatorial pattern matching and related applications. It addresses issues of searching and matching strings and more complicated patterns such as trees, regular expressions, graphs, point sets, and arrays. The goal is to derive combinatorial properties of such structures and to exploit these properties in order to achieve more efficient algorithms for the corresponding computational problems. The meeting deals with problems in bioinformatics and computational biology, coding and data compression, combinatorics on words, data mining, information retrieval, natural language processing, pattern discovery, string algorithms, string processing in databases, symbolic computing, and text searching and indexing.

This volume contains the papers presented at the 28th Annual Symposium on Combinatorial Pattern Matching (CPM 2017) held on July 4-6, 2017 in Warsaw, Poland.

The conference programme included 28 contributed papers and three invited talks by Artur Jeż (University of Wrocław, Poland), Giovanni Manzini (University of Eastern Piedmont and IIT-CNR, Italy), and Marcin Mucha (University of Warsaw, Poland). Contributions of the invited lectures are also included in this volume.

The contributed papers were selected out of 49 submissions, corresponding to an acceptance ratio of about 57%. Each submission received at least three reviews. We thank the members of the Programme Committee and all the additional external reviewers that are listed below for their hard and invaluable work that resulted in an excellent scientific programme.

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, and Tel Aviv. From the 3rd to the 26th meeting, all proceedings were published in the LNCS (Lecture Notes in Computer Science) series. The 27th meeting in 2016 was the first to have its proceedings appear in the LIPIcs (Leibniz International Proceedings in Informatics) series, as volume 54.

The whole submission and review process was carried out with the help of the EasyChair conference system. We thank the CPM Steering Committee for supporting Warsaw as the site for CPM 2017 and for their advice and help in different issues. We thank Tomasz Kociumaka and Tomasz Waleń from the University of Warsaw for their extensive involvement in the organisation of the conference and Hanna Bargieł and Monika Goszczycka from Global Congress, Poland for the local arrangements. We would like to thank the Warsaw Center of Mathematics and Computer Science for providing generous financial support to the conference.

Juha Kärkkäinen
Jakub Radoszewski
Wojciech Rytter



■ Programme Committee

Juha Kärkkäinen (Co-Chair)	University of Helsinki, Finland
Jakub Radoszewski (Co-Chair)	King's College London, UK and University of Warsaw, Poland
Wojciech Rytter (Co-Chair)	University of Warsaw, Poland
Hideo Bannai	Kyushu University, Japan
Philip Bille	Technical University of Denmark, Denmark
Maxime Crochemore	King's College London, UK and Université Paris-Est, France
Gabriele Fici	University of Palermo, Italy
Johannes Fischer	TU Dortmund, Germany
Jan Holub	Czech Technical University in Prague, Czech Republic
Stepan Holub	Charles University in Prague, Czech Republic
Moshe Lewenstein	Bar Ilan University, Israel
Gonzalo Navarro	University of Chile, Chile
Kunsoo Park	Seoul National University, South Korea
Marcin Piątkowski	Nicolaus Copernicus University in Toruń, Poland
Nadia Pisanti	University of Pisa, Italy and Erable Team INRIA, France
Simon Puglisi	University of Helsinki, Finland
Eric Rivals	CNRS and Université de Montpellier, France
Cenk Sahinalp	Indiana University, Bloomington, USA
Rahul Shah	Louisiana State University, USA
Ayumi Shinohara	Tohoku University, Japan
Arseny Shur	Ural Federal University, Russia
Tatiana Starikovskaya	Paris Diderot University, France
Gabriel Valiente	Technical University of Catalonia, Spain



■ External Reviewers

Badkobeh, Golnaz
Belazzougui, Djamal
Brubach, Brian
Canovas, Rodrigo
Cardona, Gabriel
Chakraborty, Diptarka
Chateau, Annie
Christiansen, Anders Roy
Cording, Patrick Hagge
Didier, Gilles
Faro, Simone
Ferrada, Hector
Freydenberger, Dominik
Gagie, Travis
Ganguly, Arnab
Georgiadis, Loukas
Hon, Wing-Kai
I, Tomohiro
Inenaga, Shunsuke
Jansson, Jesper
Kaniecki, Mariusz
Kempa, Dominik
Kockan, Can
Konow, Roberto
Kosolobov, Dmitry
Kucherov, Gregory
Lecroq, Thierry
Lozano, Antoni
Malikic, Salem
Manea, Florin
Mantaci, Sabrina

Marcus, Shoshana
Messeguer, Xavier
Mikulski, Łukasz
Nekrich, Yakov
Patil, Manish
Pereira, Alberto Ordóñez
Pibiri, Giulio Ermanno
Ponty, Yann
Prezza, Nicola
Raffinot, Mathieu
Rosone, Giovanna
Russo, Luis M. S.
Serna, Maria
Smyczyński, Sebastian
Sugimoto, Shiho
Tabei, Yasuo
Tarhio, Jorma
Tomescu, Alexandru I.
Vojtěchovský, Petr
Weller, Mathias
Zhang, Qin
Zhu, Binhai
Zhu, Kaiyuan

28th Annual Symposium on Combinatorial Pattern Matching (CPM 2017).

Editors: Juha Kärkkäinen, Jakub Radoszewski, and Wojciech Rytter



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

■ List of Authors

- Amir, Amihood (25, 26)
Bannai, Hideo (20, 22, 24)
Barbay, Jérémy (31)
Belazzougui, Djamal (7)
Bille, Philip (6, 15, 16)
Borozdin, Kirill (23)
Bringmann, Karl (12)
Bulteau, Laurent (27)
Cairo, Massimo (29)
Castelli, Mauro (14)
Cording, Patrick Hagge (15)
Cunha, Luís (19)
Cunial, Fabio (7)
Dantas, Simone (19)
Diptarama (8)
Dondi, Riccardo (14)
Duchon, Philippe (21)
Dudek, Bartłomiej (10)
Ettienne, Mikko Berggren (16)
Fertin, Guillaume (27)
Fischer, Johannes (15)
Fujishige, Yuta (20)
Gagie, Travis (5, 19)
Gawrychowski, Paweł (10)
Grossi, Roberto (9)
Gørtz, Inge Li (6, 15, 16)
He, Meng (5)
I, Tomohiro (18)
Iliopoulos, Costas (9)
Inenaga, Shunsuke (20, 22, 24)
Jež, Artur (2)
Katsura, Takashi (8)
Komusiewicz, Christian (11, 27)
Kosolobov, Dmitry (23)
Kowada, Luis (19)
Kuboi, Keita (20)
Köppl, Dominik (22)
Lafond, Manuel (28)
Levy, Avivit (25, 26)
Lewenstein, Moshe (25)
Liu, Chang (9)
Lubin, Ronit (25, 26)
Manzini, Giovanni (1)
Mauri, Giancarlo (14)
Medvedev, Paul (29)
Mieno, Takuya (24)
Mucha, Marcin (3)
Narisawa, Kazuyuki (8)
Navarro, Gonzalo (4, 5)
Nicaud, Cyril (21)
Obscura Acosta, Nidia (29)
Ochoa, Carlos (31)
Oliveira, Mateus de Oliveira (11)
Ostropolski-Nalewaja, Piotr (10)
Otomo, Yuhei (8)
Pibiri, Giulio Ermanno (30)
Pisanti, Nadia (9)
Pissis, Solon (9)
Pivoteau, Carine (21)
Policriti, Alberto (17)
Porat, Benny (25)
Porat, Ely (26)

Prezza, Nicola (17)
Retha, Ahmad (9)
Rizzi, Romeo (29)
Rosone, Giovanna (9)
Rubinchik, Mikhail (23)
Satti, Srinivasa Rao (31)
Scornavacca, Celine (28)
Shinohara, Ayumi (8)
Shur, Arseny (23)
Skjoldjensen, Frederik Rye (6)
Starikovskaya, Tatiana (13)
Stoye, Jens (19)
Takeda, Masayuki (20, 24)
Tomescu, Alexandru I. (29)
Vayani, Fatima (9)
Venturini, Rossano (30)
Versari, Luca (9)
Vildhøj, Hjalte Wedel (16)
Wellnitz, Philip (12)
Wittler, Roland (19)
Zehavi, Meirav (11)
Zoppis, Italo (14)