

# Combinatorial Pattern Matching

27th Annual Symposium on Combinatorial Pattern Matching,  
June 27-29, 2016, Tel Aviv, Israel

Edited by

Roberto Grossi

Moshe Lewenstein



*Editors*

Roberto Grossi	Moshe Lewenstein
Dipartimento di Informatica	Department of Computer Science
Università di Pisa	Bar Ilan University
<a href="mailto:grossi@di.unipi.it">grossi@di.unipi.it</a>	<a href="mailto:moshe@cs.biu.ac.il">moshe@cs.biu.ac.il</a>

*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 978-3-95977-012-5**

*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/978-3-95977-012-5>.

*Publication date*

June 2016

*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.2016.0

**ISBN 978-3-95977-012-5**

**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*

- Susanne Albers (TU München)
- Chris Hankin (Imperial College London)
- Deepak Kapur (University of New Mexico)
- Michael Mitzenmacher (Harvard University)
- Madhavan Mukund (Chennai Mathematical Institute)
- Catuscia Palamidessi (INRIA)
- Wolfgang Thomas (*Chair*, RWTH Aachen)
- Pascal Weil (CNRS and University Bordeaux)
- Reinhard Wilhelm (Saarland University)

**ISSN 1868-8969**

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



## ■ Contents

Preface	
<i>Roberto Grossi and Moshe Lewenstein</i> .....	0:vii
Deterministic Sub-Linear Space LCE Data Structures With Efficient Construction	
<i>Yuka Tanimura, Tomohiro I, Hideo Bannai, Shunsuke Inenaga, Simon J. Puglisi, and Masayuki Takeda</i> .....	1:1–1:10
Space-Efficient Dictionaries for Parameterized and Order-Preserving Pattern Matching	
<i>Arnab Ganguly, Wing-Kai Hon, Kunihiko Sadakane, Rahul Shah, Sharma V. Thankachan, and Yilin Yang</i> .....	2:1–2:12
Encoding Two-Dimensional Range Top-k Queries	
<i>Seungbum Jo, Rahul Lingala, and Srinivasa Rao Satti</i> .....	3:1–3:11
Efficient Index for Weighted Sequences	
<i>Carl Barton, Tomasz Kociumaka, Solon P. Pissis, and Jakub Radoszewski</i> .....	4:1–4:13
Faster Longest Common Extension Queries in Strings over General Alphabets	
<i>Pawel Gawrychowski, Tomasz Kociumaka, Wojciech Rytter, and Tomasz Waleń</i> ..	5:1–5:13
Succinct Online Dictionary Matching with Improved Worst-Case Guarantees	
<i>Tsvi Kopelowitz, Ely Porat, and Yaron Rozen</i> .....	6:1–6:13
Graph Motif Problems Parameterized by Dual	
<i>Guillaume Fertin and Christian Komusiewicz</i> .....	7:1–7:12
Truly Subquadratic-Time Extension Queries and Periodicity Detection in Strings with Uncertainties	
<i>Costas S. Iliopoulos and Jakub Radoszewski</i> .....	8:1–8:12
Estimating Statistics on Words Using Ambiguous Descriptions	
<i>Cyril Nicaud</i> .....	9:1–9:12
Reconstruction of Trees from Jumbled and Weighted Subtrees	
<i>Dénes Bartha, Péter Burcsi, and Zsuzsanna Lipták</i> .....	10:1–10:13
A $7/2$ -Approximation Algorithm for the Maximum Duo-Preservation String Mapping Problem	
<i>Nicolas Boria, Gianpiero Cabodi, Paolo Camurati, Marco Palena, Paolo Pasini, and Stefano Quer</i> .....	11:1–11:8
Fast Compatibility Testing for Rooted Phylogenetic Trees	
<i>Yun Deng and David Fernández-Baca</i> .....	12:1–12:12
Hardness of RNA Folding Problem With Four Symbols	
<i>Yi-Jun Chang</i> .....	13:1–13:12



Efficient Non-Binary Gene Tree Resolution with Weighted Reconciliation Cost <i>Manuel Lafond, Emmanuel Noutahi, and Nadia El-Mabrouk</i> .....	14:1–14:12
Genomic Scaffold Filling Revisited <i>Haitao Jiang, Chenglin Fan, Boting Yang, Farong Zhong, Daming Zhu, and Binhai Zhu</i> .....	15:1–15:13
A Linear-Time Algorithm for the Copy Number Transformation Problem <i>Ron Shamir, Meirav Zehavi, and Ron Zeira</i> .....	16:1–16:13
On Almost Monge All Scores Matrices <i>Amir Carmel, Dekel Tsur, and Michal Ziv-Ukelson</i> .....	17:1–17:12
Tight Tradeoffs for Real-Time Approximation of Longest Palindromes in Streams <i>Paweł Gawrychowski, Oleg Merkurev, Arseny M. Shur, and Przemysław Uznański</i>	18:1–18:13
Finding Maximal 2-Dimensional Palindromes <i>Sara H. Gezhals and Dina Sokol</i> .....	19:1–19:12
Boxed Permutation Pattern Matching <i>Mika Amit, Philip Bille, Patrick Hagge Cording, Inge Li Gørtz, and Hjalte Wedel Vildhøj</i> .....	20:1–20:11
Longest Common Substring with Approximately $k$ Mismatches <i>Tatiana Starikovskaya</i> .....	21:1–21:11
Fully-online Construction of Suffix Trees for Multiple Texts <i>Takuya Takagi, Shunsuke Inenaga, and Hiroki Arimura</i> .....	22:1–22:13
Linear-time Suffix Sorting – A New Approach for Suffix Array Construction <i>Uwe Baier</i> .....	23:1–23:12
Color-Distance Oracles and Snippets <i>Tsvi Kopelowitz and Robert Krauthgamer</i> .....	24:1–24:10
The Nearest Colored Node in a Tree <i>Paweł Gawrychowski, Gad M. Landau, Shay Mozes, and Oren Weimann</i> .....	25:1–25:12
On the Benefit of Merging Suffix Array Intervals for Parallel Pattern Matching <i>Johannes Fischer, Dominik Köppl, and Florian Kurpicz</i> .....	26:1–26:11
Factorizing a String into Squares in Linear Time <i>Yoshiaki Matsuoka, Shunsuke Inenaga, Hideo Bannai, Masayuki Takeda, and Florin Manea</i> .....	27:1–27:12
Minimal Suffix and Rotation of a Substring in Optimal Time <i>Tomasz Kociumaka</i> .....	28:1–28:12
Optimal Prefix Free Codes with Partial Sorting <i>Jérémy Barbay</i> .....	29:1–29:13

## ■ Preface

The objective of the Annual Symposium on Combinatorial Pattern Matching is to provide 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 a superior performance for the corresponding computational problems. The meeting also 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.

This volume contains the papers presented at the 27th Annual Symposium on Combinatorial Pattern Matching (CPM 2016) held during June 27-29, 2016, in Tel Aviv, Israel

The conference program included 29 contributed papers and three invited talks by Gregory Kucherov, University Paris-Est Marne-la-Vallée, France, on “Stringology in action: bioinformatics examples”, Moni Naor, Weizmann Institute, Israel, on “How to share a secret, infinitely”, and Yoram Louzon, Bar-Ilan University, Israel, on “Node classification based on local and global sub-graph patterns”.

The contributed papers were selected out of 52 submissions, corresponding to an acceptance ratio of 55.8%. Each submission received at least three reviews. We thank the members of the Program Committee and all the additional external reviewers for their hard and invaluable work that resulted in an excellent scientific program.

The Annual Symposium on Combinatorial Pattern Matching started in 1990, and has since 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, and Ischia. From the 3rd to the 26th meeting, proceedings of all meetings have been published in the LNCS series, as volumes 644, 684, 807, 937, 1075, 1264, 1448, 1645, 1848, 2089, 2373, 2676, 3109, 3537, 4009, 4580, 5029, 5577, 6129, 6661, 7354, 7922, 8486, and 9133, respectively. From the current meeting, the proceedings will appear in the LIPIcs (Leibniz International Proceedings in Informatics) series.

Selected papers from the 1st meeting appeared in vol. 92 of Theoretical Computer Science, from the 11th meeting in vol. 2 of the Journal of Discrete Algorithms, from the 12th meeting in vol. 146 of Discrete Applied Mathematics, from the 14th meeting in vol. 3 of the Journal of Discrete Algorithms, from the 15th meeting in vol. 368 of Theoretical Computer Science, from the 16th meeting in vol. 5 of the Journal of Discrete Algorithms, from the 19th meeting in vol. 410 of Theoretical Computer Science, from the 20th meeting in vol. 9 of the Journal of Discrete Algorithms, from the 21st meeting in vol. 213 of Information and Computation, from the 22nd meeting in vol. 483 of Theoretical Computer Science, and from the 23rd meeting in vol. 25 of the Journal of Discrete Algorithms. Selected papers from this meeting will appear in a special issue of Algorithmica.

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 Tel Aviv as the site for CPM 2016, and for their advice and help in different issues. We thank Shay Golan, Avivit Levy, Noa Lewenstein and Ely Porat for the local arrangements. The conference was



sponsored by the I-CORE Program of the planning and budgeting committee and The Israel Science Foundation (grant number 4/11), Shenkar College of Engineering and Design, and Bar Ilan University. We thank them for their financial support.



## ■ List of Authors

Amit, Mika (20)  
Arimura, Hiroki (22)  
Baier, Uwe (23)  
Bannai, Hideo (1, 27)  
Barbay, Jérémy (29)  
Bartha, Dénes (10)  
Barton, Carl (4)  
Bille, Philip (20)  
Boria, Nicolas (11)  
Burcsi, Péter (10)  
Cabodi, Gianpiero (11)  
Camurati, Paolo (11)  
Carmel, Amir (17)  
Chang, Yi-Jun (13)  
Cording, Patrick Haggge (20)  
Deng, Yun (12)  
El-Mabrouk, Nadia (14)  
Fan, Chenglin (15)  
Fernández-Baca, David (12)  
Fertin, Guillaume (7)  
Fischer, Johannes (26)  
Ganguly, Arnab (2)  
Gawrychowski, Paweł (5, 18, 25)  
Gezhals, Sara H (19)  
Gørtz, Inge Li (20)  
Hon, Wing-Kai (2)  
I, Tomohiro (1)  
Iliopoulos, Costas (8)  
Inenaga, Shunsuke (1, 22, 27)  
Jiang, Haitao (15)  
Jo, Seungbum (3)  
Kociumaka, Tomasz (4, 5, 28)  
Komusiewicz, Christian (7)  
Kopelowitz, Tsvi (6, 24)  
Krauthgamer, Robert (24)  
Kurpicz, Florian (26)  
Köppl, Dominik (26)  
Lafond, Manuel (14)  
Landau, Gad M. (25)  
Lingala, Rahul (3)  
Lipták, Zsuzsanna (10)  
Manea, Florin (27)  
Matsuoka, Yoshiaki (27)  
Merkurev, Oleg (18)  
Mozes, Shay (25)  
Nicaud, Cyril (9)  
Noutahi, Emmanuel (14)  
Palena, Marco (11)  
Pasini, Paolo (11)  
Pissis, Solon (4)  
Porat, Ely (6)  
Puglisi, Simon (1)  
Quer, Stefano (11)  
Radoszewski, Jakub (4, 8)  
Rozen, Yaron (6)  
Rytter, Wojciech (5)  
Sadakane, Kunihiko (2)  
Satti, Srinivasa Rao (3)  
Shah, Rahul (2)  
Shamir, Ron (16)  
Shur, Arseny M. (18)  
Sokol, Dina (19)

27th Annual Symposium on Combinatorial Pattern Matching (CPM 2016).

Editors: Roberto Grossi and Moshe Lewenstein



Leibniz International Proceedings in Informatics

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

Starikovskaya, Tatiana (21)

Takagi, Takuya (22)

Takeda, Masayuki (1, 27)

Tanimura, Yuka (1)

Thankachan, Sharma V. (2)

Tsur, Dekel (17)

Uznański, Przemysław (18)

Vildhøj, Hjalte Wedel (20)

Waleń, Tomasz (5)

Weimann, Oren (25)

Yang, Boting (15)

Yang, Yilin (2)

Zehavi, Meirav (16)

Zeira, Ron (16)

Zhong, Farong (15)

Zhu, Binhai (15)

Zhu, Daming (15)

Ziv-Ukelson, Michal (17)