Lecture Notes in Computer Science 7922

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
   Lancaster University, UK
Takao Kanade
   Carnegie Mellon University, Pittsburgh, PA, USA
Josef Kittler
   University of Surrey, Guildford, UK
Jon M. Kleinberg
   Cornell University, Ithaca, NY, USA
Alfred Kobsa
   University of California, Irvine, CA, USA
Friedemann Mattern
   ETH Zurich, Switzerland
John C. Mitchell
   Stanford University, CA, USA
Moni Naor
   Weizmann Institute of Science, Rehovot, Israel
Oscar Nierstrasz
   University of Bern, Switzerland
C. Pandu Rangan
   Indian Institute of Technology, Madras, India
Bernhard Steffen
   TU Dortmund University, Germany
Madhu Sudan
   Microsoft Research, Cambridge, MA, USA
Demetri Terzopoulos
   University of California, Los Angeles, CA, USA
Doug Tygar
   University of California, Berkeley, CA, USA
Gerhard Weikum
   Max Planck Institute for Informatics, Saarbruecken, Germany
Preface

This volume contains the papers presented at the 24th Annual Symposium on Combinatorial Pattern Matching (CPM 2013) held in Bad Herrenalb near Karlsruhe, Germany, during June 17–19, 2013. The hosting university was the Karlsruhe Institute of Technology.

There were 51 submissions from 22 countries. Each submission was reviewed by at least three Program Committee members. The committee decided to accept 21 papers, corresponding to an acceptance rate of 41%. We thank the members of the Program Committee and all additional external reviewers for their hard work that resulted in this excellent program. Their names are listed on the following pages. The whole submission and review process was carried out with the invaluable help of the EasyChair conference system.

The program also included two invited talks by Moshe Lewenstein from Bar Ilan University, Israel, and by Gene Myers from the MPI for Molecular Cell Biology and Genetics, Dresden, Germany. We thank the invited speakers for their contributions.

2013 marks the 40th anniversary of Peter Weiner’s foundational paper “Linear pattern matching algorithms” on suffix trees, in those days called “bi-trees” [14th Annual Symposium on Switching and Automata Theory (SWAT; nowadays FOCS), pp. 1–11, 1973]. CPM 2013 celebrated this event with a special session, organized by Martin Farach-Colton and S. Muthukrishnan, both from Rutgers University, USA. It included talks by Edward M. McCreight, Vaughan R. Pratt, Peter Weiner, and Jacob Ziv. This special session was accompanied by an invited contribution “Forty Years of Text Indexing” by Alberto Apostolico, Maxime Crochemore, Martin Farach-Colton, Zvi Galil, and S. Muthukrishnan.

The objective of the annual CPM meetings 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 non-trivial combinatorial properties of such structures and to exploit these properties in order to either achieve superior performance for the corresponding computational problems or pinpoint conditions under which searches cannot be performed efficiently. The meeting also deals with problems in computational biology, data compression and data mining, coding, information retrieval, natural language processing, and pattern recognition.

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, and Helsinki again. This year’s meeting was the first in Germany. Starting from
the third 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, and 7354.

We thank SAP (Walldorf, Germany) and the German Research Foundation (DFG) for their financial support.

April 2013

Johannes Fischer
Peter Sanders
## Organization

### Program Committee

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolf Backofen</td>
<td>University of Freiburg, Germany</td>
</tr>
<tr>
<td>Philip Bille</td>
<td>Technical University of Denmark</td>
</tr>
<tr>
<td>Christina Boucher</td>
<td>Colorado State University, USA</td>
</tr>
<tr>
<td>Francisco Claude</td>
<td>University of Waterloo, Canada</td>
</tr>
<tr>
<td>Johannes Fischer</td>
<td>Karlsruhe Institute of Technology, Germany (Co-chair)</td>
</tr>
<tr>
<td>Travis Gagie</td>
<td>University of Helsinki, Finland</td>
</tr>
<tr>
<td>Pawel Gawrychowski</td>
<td>MPI Saarbrücken, Germany</td>
</tr>
<tr>
<td>Meng He</td>
<td>Dalhousie University, Canada</td>
</tr>
<tr>
<td>Jan Holub</td>
<td>Czech Technical University in Prague, Czech Republic</td>
</tr>
<tr>
<td>Jesper Jansson</td>
<td>Kyoto University, Japan</td>
</tr>
<tr>
<td>Juha Kärkkäinen</td>
<td>University of Helsinki, Finland</td>
</tr>
<tr>
<td>Tsvi Kopelowitz</td>
<td>Weizmann Institute of Science, Israel</td>
</tr>
<tr>
<td>Alejandro López-Ortiz</td>
<td>University of Waterloo, Canada</td>
</tr>
<tr>
<td>Gonzalo Navarro</td>
<td>University of Chile</td>
</tr>
<tr>
<td>Kunsoo Park</td>
<td>Seoul National University, South Korea</td>
</tr>
<tr>
<td>Mike Paterson</td>
<td>University of Warwick, UK</td>
</tr>
<tr>
<td>Rajeev Raman</td>
<td>University of Leicester, UK</td>
</tr>
<tr>
<td>Benjamin Sach</td>
<td>University of Warwick, UK</td>
</tr>
<tr>
<td>Kunihiko Sadakane</td>
<td>National Institute of Informatics, Japan</td>
</tr>
<tr>
<td>Peter Sanders</td>
<td>Karlsruhe Institute of Technology, Germany (Co-chair)</td>
</tr>
<tr>
<td>Jens Stoye</td>
<td>University of Bielefeld, Germany</td>
</tr>
<tr>
<td>Rossano Venturini</td>
<td>University of Pisa, Italy</td>
</tr>
<tr>
<td>Oren Weimann</td>
<td>University of Haifa, Israel</td>
</tr>
</tbody>
</table>

### Steering Committee

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alberto Apostolico</td>
<td>Georgia Institute of Technology, USA</td>
</tr>
<tr>
<td>Maxime Crochemore</td>
<td>Université Paris-Est, France, and King’s College London, UK</td>
</tr>
<tr>
<td>Zvi Galil</td>
<td>Georgia Institute of Technology, USA</td>
</tr>
</tbody>
</table>

### Organizing Committee

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timo Bingmann</td>
<td>Karlsruhe Institute of Technology, Germany</td>
</tr>
<tr>
<td>Johannes Fischer</td>
<td>Karlsruhe Institute of Technology, Germany</td>
</tr>
</tbody>
</table>
Peter Sanders  
Karlsruhe Institute of Technology, Germany

Nodari Sitchinava  
Karlsruhe Institute of Technology, Germany

Additional Reviewers

Amir, Amihood  
Manea, Florin

Amit, Mika  
Martínez-Prieto, Miguel A.

Arimura, Hiroki  
Mäkinen, Veli

Arroyuelo, Diego  
Nicholson, Patrick K.

Baier, Jan  
Osipov, Vitaly

Bannai, Hideo  
Pedersen, Christian Nørgaard Storm

Bingmann, Timo  
Philip, Geevarghese

Braga, Marilia  
Polach, Radomír

Canovas, Rodrigo  
Ponty, Yann

Chitsaz, Hamidreza  
Porat, Ely

Clifford, Raphael  
Pritchard, David

Cormode, Graham  
Puglisi, Simon

Crochemore, Maxime  
Sakamoto, Hiroshi

Cunial, Fabio  
Salinger, Alejandro

Cygan, Marek  
Salmela, Leena

Doerr, Daniel  
Savari, Serap

Dorrigiv, Reza  
Schmiedl, Christina

Fernau, Henning  
Seco, Diego

Goto, Keisuke  
Shibuya, Tetsuo

Guillemot, Sylvain  
Sirén, Jouni

Inenaga, Shunsuke  
Thankachan, Sharma

Jahn, Katharina  
Tiskin, Alexander

Jalsenius, Markus  
Transier, Frederik

Jeż, Artur  
Travnické, Jan

Jurdzinski, Tomasz  
Truszkowski, Jakub

Kempa, Dominik  
Vildhøj, Hjalte Wedel

Kennedy, Sean  
Walen, Tomasz

Krach, Peter  
Will, Sebastian

Kucherov, Gregory  
Wittler, Roland

Ladra, Susana  
Zhou, Gelin

Lokshtanov, Daniel  
Zhu, Binhai

Ziv-Ukelson, Michal
# Table of Contents

Forty Years of Text Indexing .............................................................. 1  
*Alberto Apostolico, Maxime Crochemore, Martin Farach-Colton, Zvi Galil, and S. Muthukrishnan*

LCP Magic ................................................................. 11  
*Moshe Lewenstein*

Discrete Methods for Image Analysis Applied to Molecular Biology .... 12  
*Gene Myers*

Locating All Maximal Approximate Runs in a String ..................... 13  
*Mika Amit, Maxime Crochemore, and Gad M. Landau*

On Minimal and Maximal Suffixes of a Substring ......................... 28  
*Maxim Babenko, Ignat Kolesnichenko, and Tatiana Starikovskaya*

Converting SLP to LZ78 in almost Linear Time ............................ 38  
*Hideo Bannai, Paweł Gawrychowski, Shunsuke Inenaga, and Masayuki Takeda*

A Bit-Parallel, General Integer-Scoring Sequence Alignment Algorithm .................................................. 50  
*Gary Benson, Yozen Hernandez, and Joshua Loving*

Compact q-Gram Profiling of Compressed Strings .......................... 62  
*Philip Bille, Patrick Hagge Cording, and Inge Li Gørtz*

A Constant-Space Comparison-Based Algorithm for Computing the Burrows–Wheeler Transform ............................................. 74  
*Maxime Crochemore, Roberto Grossi, Juha Kärkkäinen, and Gad M. Landau*

Pattern Matching with Variables: A Multivariate Complexity Analysis ................................................................. 83  
*Henning Fernau and Markus L. Schmid*

New Algorithms for Position Heaps .............................................. 95  
*Travis Gagie, Wing-Kai Hon, and Tsung-Han Ku*

Document Listing on Repetitive Collections .................................. 107  
*Travis Gagie, Kalle Karhu, Gonzalo Navarro, Simon J. Puglisi, and Jouni Sirén*
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximating Shortest Superstring Problem Using de Bruijn Graphs</td>
<td>120</td>
</tr>
<tr>
<td><em>Alexander Golovnev, Alexander S. Kulikov, and Ivan Mihajlin</em></td>
<td></td>
</tr>
<tr>
<td>Local Search for String Problems: Brute Force Is Essentially Optimal</td>
<td>130</td>
</tr>
<tr>
<td><em>Jiong Guo, Danny Hermelin, and Christian Komusiewicz</em></td>
<td></td>
</tr>
<tr>
<td>Space-Efficient Construction Algorithm for the Circular Suffix Tree</td>
<td>142</td>
</tr>
<tr>
<td><em>Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, and Sharma V. Thankachan</em></td>
<td></td>
</tr>
<tr>
<td>Efficient Lyndon Factorization of Grammar Compressed Text</td>
<td>153</td>
</tr>
<tr>
<td><em>Tomohiro I, Yuto Nakashima, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda</em></td>
<td></td>
</tr>
<tr>
<td>Approximation of Grammar-Based Compression via Recompression</td>
<td>165</td>
</tr>
<tr>
<td><em>Artur Jeż</em></td>
<td></td>
</tr>
<tr>
<td>Fast Algorithm for Partial Covers in Words</td>
<td>177</td>
</tr>
<tr>
<td><em>Tomasz Kociumaka, Solon P. Pissis, Jakub Radoszewski, Wojciech Rytter, and Tomasz Waleni</em></td>
<td></td>
</tr>
<tr>
<td>Linear Time Lempel-Ziv Factorization: Simple, Fast, Small</td>
<td>189</td>
</tr>
<tr>
<td><em>Juha Kärkkäinen, Dominik Kempa, and Simon J. Puglisi</em></td>
<td></td>
</tr>
<tr>
<td>External Memory Generalized Suffix and LCP Arrays Construction</td>
<td>201</td>
</tr>
<tr>
<td><em>Felipe A. Louza, Guilherme P. Telles, and Cristina Dutra De Aguiar Ciferri</em></td>
<td></td>
</tr>
<tr>
<td>Efficient All Path Score Computations on Grid Graphs</td>
<td>211</td>
</tr>
<tr>
<td><em>Ury Matarazzo, Dekel Tsur, and Michal Ziv-Ukelson</em></td>
<td></td>
</tr>
<tr>
<td>Time-Space Trade-Offs for the Longest Common Substring Problem</td>
<td>223</td>
</tr>
<tr>
<td><em>Tatiana Starikovskaya and Hjalte Wedel Vildhøj</em></td>
<td></td>
</tr>
<tr>
<td>A Succinct Grammar Compression</td>
<td>235</td>
</tr>
<tr>
<td><em>Yasuo Tabei, Yoshimasa Takabatake, and Hiroshi Sakamoto</em></td>
<td></td>
</tr>
<tr>
<td>Data Structure Lower Bounds on Random Access to Grammar-Compressed Strings</td>
<td>247</td>
</tr>
<tr>
<td><em>Elad Verbin and Wei Yu</em></td>
<td></td>
</tr>
<tr>
<td><strong>Author Index</strong></td>
<td>259</td>
</tr>
</tbody>
</table>