

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo 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

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Paolo Ferragina Gad M. Landau (Eds.)

Combinatorial Pattern Matching

19th Annual Symposium, CPM 2008
Pisa, Italy, June 18-20, 2008
Proceedings

Volume Editors

Paolo Ferragina
University of Pisa, Dipartimento di Informatica
Largo B. Pontecorvo 3, 56127 Pisa, Italy
E-mail: ferragina@di.unipi.it

Gad M. Landau
University of Haifa, Department of Computer Science
Mount Carmel, Haifa 31905, Israel
E-mail: landau@cs.haifa.ac.il

Library of Congress Control Number: 2008928843

CR Subject Classification (1998): F.2.2, I.5.4, I.5.0, H.3.3, J.3, E.4, G.2.1, E.1

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

ISSN 0302-9743
ISBN-10 3-540-69066-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-69066-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2008
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12281893 06/3180 5 4 3 2 1 0

Preface

The papers contained in this volume were presented at the 19th Annual Symposium on Combinatorial Pattern Matching (CPM 2008) held at the University of Pisa, Italy, June 18–20, 2008.

All the papers presented at the conference are original research contributions on computational pattern matching and analysis. They were selected from 78 submissions. Each submission was reviewed by at least three reviewers. The committee decided to accept 25 papers. The programme also includes three invited talks by Daniel M. Gusfield from the University of California, Davis, USA, J. Ian Munro from the University of Waterloo, Canada, and Prabhakar Raghavan from Yahoo! Research, USA.

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, 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 and London (Canada). Starting from the third meeting, proceedings have been published in the LNCS series, volumes 644, 684, 807, 937, 1075, 1264, 1448, 1645, 1848, 2089, 2373, 2676, 3109, 3537, 4009, and 4580. Selected papers from the first meeting appeared in volume 92 of *Theoretical Computer Science*, from the 11th meeting in volume 2 of *Journal of Discrete Algorithms*, from the 12th meeting in volume 146 of *Discrete Applied Mathematics*, from the 14th meeting in volume 3 of *Journal of Discrete Algorithms*, from the 15th meeting in volume 368 of *Theoretical Computer Science* and from the 16th meeting in volume 5 of *Journal of Discrete Algorithms*.

The whole submission and review process was carried out with the help of the EasyChair system. The conference was sponsored by the University of Pisa and by Yahoo! Research. Special thanks are due to the members of the Program Committee who worked very hard to ensure the timely review of all the submitted manuscripts, and participated in a stimulating discussion that allowed all of us to choose the best papers to be presented at the conference.

Organization

Program Committee

Srinivasa Aluru	Iowa State University
Lars Arge	University of Aarhus
Ricardo Baeza-Yates	Yahoo! Research
Edgar Chavez	Universidad Michoacana
Ken Church	Microsoft Research
Raphael Clifford	University of Bristol
Paolo Ferragina (Co-chair)	University of Pisa
Roberto Grossi	University of Pisa
Tao Jiang	University of California, Riverside
Dong K. Kim	Hanyang University
Gregory Kucherov	CNRS, Lille
Gad M. Landau (Co-chair)	Haifa University and Polytechnic University
Stefano Lonardi	University of California, Riverside
Bin Ma	University of Western Ontario
Veli Makinen	University of Helsinki
Yossi Matias	Google, Tel Aviv
Alistair Moffat	University of Melbourne
Ian Munro	University of Waterloo
Gene Myers	Howard Hughes Medical Institute
Wojciech Plandowski	Uniwersytet Warszawski
Rajeev Raman	University of Leicester
Kunihiko Sadakane	Kyushu University
Peter Sanders	Karlsruhe University
Marinella Sciortino	University of Palermo
Dafna Sheinwald	IBM, Haifa
Stephane Vialette	University of Marne-la-Vallee
Frances Yao	City University of Hong Kong
Kaizhong Zhang	University of Western Ontario

Steering Committee

Alberto Apostolico	University of Padova and Georgia Tech
Maxime Crochemore	King's College
Zvi Galil	Tel Aviv University

Organizing Committee

Anna Bernasconi	Roberto Grossi
Paolo Ferragina	Danny Hermelin

Fabrizio Luccio
Igor Nitto
Linda Pagli
Nadia Pisanti

Giuseppe Prencipe
Romano Venturini
Oren Weimann

External Referees

Spyros Angelopoulos
Vo Anh
Frederique Bassino
Philip Bille
Marek Biskup
Guillaume Blin
Gerth Brodal
Marie-Pierre Bal
Arturo Carpi
Giuseppa Castiglione
Eric Chen
Shihyen Chen
Joselito Chua
Hagai Cohen
Shane Culpepper
Robert Dabrowski
Reza Dorri-Giv
Arash Farzan
Guillaume Fertin
Johannes Fischer
Gianni Franceschini
Kimmo Fredriksson
Travis Gagie
Mathieu Giraud
Massimiliano Goldwurm
Mark Greve
Franciszek Grzegorek
Sylvie Hamel
Elena Harris
Aram Harrow
Tzvika Hartman
Meng He
Danny Hermelin
Benjamin Jackson
Jesper Jansson
Anantharaman Kalyanaraman
Orgad Keller
Pang Ko

Roman Kolpakov
Robert Krauthgamer
Inbok Lee
Avivit Levy
Weiming Li
Xiaowen Liu
Jingping Liu
Giovanni Manzini
Jerme Monnot
Ashley Montanaro
Hiroyoshi Morita
Gabriel Moruz
Joong Chae Na
Pat Nicholson
Francois Nicolas
Igor Nitto
Igor Nor
Heejin Park
Giulio Pavesi
Christian Storm Pedersen
Gemma Piella
Natasza Przulj
Simon Puglisi
Mathieu Raffinot
Srinivasa Rao
Antonio Restivo
Romeo Rizzi
Wojciech Rytter
Benjamin Sach
Cenk Sahinalp
Abhinav Sarje
Srinivasa Rao Satti
Roded Sharan
Johannes Singler
Ranjan Sinha
Jouni Siren
Vladimir Vacic
Rossano Venturini

Niko Valimaki
Tomasz Wale
Lusheng Wang
Bob Wang
William Webber
Zhan Wu

Xiao Yang
Qiaofeng Yang
Jie Zheng
Jaroslaw Zola
Roelof van Zwol
Yonghui Wu

Sponsoring Institutions

Yahoo! Research
University of Pisa

Table of Contents

Invited Talks

ReCombinatorics: Combinatorial Algorithms for Studying the History of Recombination in Populations	1
<i>Dan Gusfield</i>	
Lower Bounds for Succinct Data Structures	3
<i>J. Ian Munro</i>	
The Changing Face of Web Search	4
<i>Prabhakar Raghavan</i>	

Contributed Papers

Two-Dimensional Pattern Matching with Combined Scaling and Rotation	5
<i>Christian Hundt and Maciej Liśkiewicz</i>	
Searching for Gapped Palindromes	18
<i>Roman Kolpakov and Gregory Kucherov</i>	
Parameterized Algorithms and Hardness Results for Some Graph Motif Problems	31
<i>Nadja Betzler, Michael R. Fellows, Christian Komusiewicz, and Rolf Niedermeier</i>	
Finding Largest Well-Predicted Subset of Protein Structure Models	44
<i>Shuai Cheng Li, Dongbo Bu, Jinbo Xu, and Ming Li</i>	
HP Distance Via Double Cut and Join Distance	56
<i>Anne Bergeron, Julia Mixtacki, and Jens Stoye</i>	
Fixed Parameter Tractable Alignment of RNA Structures Including Arbitrary Pseudoknots	69
<i>Mathias Möhl, Sebastian Will, and Rolf Backofen</i>	
Faster Algorithm for the Set Variant of the String Barcoding Problem	82
<i>Leszek Gąsieniec, Cindy Y. Li, and Meng Zhang</i>	
Probabilistic Arithmetic Automata and Their Application to Pattern Matching Statistics	95
<i>Tobias Marschall and Sven Rahmann</i>	

Analysis of the Size of Antidictionary in DCA	107
<i>Julien Fayolle</i>	
Approximate String Matching with Address Bit Errors	118
<i>Amihod Amir, Yonatan Aumann, Oren Kapah, Avivit Levy, and Ely Porat</i>	
On-Line Approximate String Matching with Bounded Errors	130
<i>Marcos Kiwi, Gonzalo Navarro, and Claudio Telha</i>	
A Black Box for Online Approximate Pattern Matching	143
<i>Raphaël Clifford, Klim Efremenko, Benny Porat, and Ely Porat</i>	
An(other) Entropy-Bounded Compressed Suffix Tree	152
<i>Johannes Fischer, Veli Mäkinen, and Gonzalo Navarro</i>	
On Compact Representations of All-Pairs-Shortest-Path-Distance Matrices	166
<i>Igor Nitto and Rossano Venturini</i>	
Computing Inverse ST in Linear Complexity	178
<i>Ge Nong, Sen Zhang, and Wai Hong Chan</i>	
Dynamic Fully-Compressed Suffix Trees	191
<i>Luís M.S. Russo, Gonzalo Navarro, and Arlindo L. Oliveira</i>	
A Linear Delay Algorithm for Building Concept Lattices	204
<i>Martin Farach-Colton and Yang Huang</i>	
Matching Integer Intervals by Minimal Sets of Binary Words with <i>don't cares</i>	217
<i>Wojciech Fraczak, Wojciech Rytter, and Mohammadreza Yazdani</i>	
Fast Algorithms for Computing Tree LCS	230
<i>Shay Mozes, Dekel Tsur, Oren Weimann, and Michal Ziv-Ukelson</i>	
Why Greed Works for Shortest Common Superstring Problem	244
<i>Bin Ma</i>	
Constrained LCS: Hardness and Approximation	255
<i>Zvi Gotthilf, Danny Hermelin, and Moshe Lewenstein</i>	
Finding Additive Biclusters with Random Background	263
<i>Jing Xiao, Lusheng Wang, Xiaowen Liu, and Tao Jiang</i>	
An Improved Succinct Representation for Dynamic k -ary Trees	277
<i>Diego Arroyuelo</i>	
Towards a Solution to the “Runs” Conjecture	290
<i>Maxime Crochemore, Lucian Ilie, and Liviu Tinta</i>	

On the Longest Common Parameterized Subsequence..... 303
Orgad Keller, Tsvi Kopelowitz, and Moshe Lewenstein

Author Index 317