Lecture Notes in Computer Science 644
Edited by G. Goos and J. Hartmanis



A. Apostolico M. Crochemore Z. Galil
U. Manber (Eds.)

# Combinatorial Pattern Matching

Third Annual Symposium
Tucson, Arizona, USA, April 29-May 1, 1992
Proceedings

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Volume Editors

Alberto Apostolico
Dept. of Computer Sciences, Purdue University
1398 Comp. Sc. Bldg., West Lafayette, IN 47907-1398, USA

Maxime Crochemore
L. I. T. P., Université Paris VII
F-75251 Paris CX 05, France

Zvi Galil
Columbia University, New York, NY 10027, USA
and
Tel Aviv University
Ramat Aviv, Tel Aviv, Israel

Udi Manber
Computer Science Dept., University of Arizona
Gould-Simpson 721, Tucson, AZ 85721, USA

CR Subject Classification (1991): F.2.2, I.5.4, I.5.0, I.7.3, H.3.3, E.4

ISBN 3-540-56024-6 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-56024-6 Springer-Verlag New York Berlin Heidelberg



Typesetting: Camera ready by author/editor
45/3140-543210 - Printed on acid-free paper

# Foreword

The papers contained in this volume were presented at the third annual symposium on Combinatorial Pattern Matching, held April 29 to May 1, 1992 in Tucson, Arizona. They were selected from 39 abstracts submitted in response to the call for papers.

Combinatorial Pattern Matching addresses issues of searching and matching of strings and more complicated patterns such as trees, regular expressions, extended expressions, etc. The goal is to derive nontrivial combinatorial properties for such structures and then to exploit these properties in order to achieve superior performances for the corresponding computational problems. In recent years, a steady flow of high-quality scientific study of this subject has changed a sparse set of isolated results into a full-fledged area of algorithmics. Still, there is currently no central place for disseminating results in this area. We hope that CPM can grow to serve as the focus point.

This area is expected to grow even further due to the increasing demand for speed and efficiency that comes especially from molecular biology and the Genome project, but also from other diverse areas such as information retrieval (e.g., supporting complicated search queries), pattern recognition (e.g., using strings to represent polygons and string matching to identify them), compilers (e.g., using tree matching), data compression, and program analysis (e.g., program integration efforts). The stated objective of CPM gatherings is to bring together once a year the researchers active in the area for an informal and yet intensive exchange of information about current and future research in the area.

The first two meetings were held at the University of Paris in 1990 and at the University of London in 1991. These two meetings were informal and no proceedings were produced. We hope that these proceedings will contribute to the success and growth of this area.

The conference was supported in part by the National Science Foundation and the University of Arizona.

**Program Committee**

A. Apostolico
M. Crochemore
Z. Galil
G. Gonnet
D. Gusfield
D. Hirschberg
U. Manber, *chair*
E. W. Myers
F. Tompa
E. Ukkonen

**Standing back: (right to left):** Gary Benson, Ricardo Baeza-Yates, Andrew Hume, Jim Knight, Christian Burks, Hal Berghel, Andrzej Ehrenfeucht, David Roach, Udi Manber, Mike Waterman, Ramana Idury, Amihood Amir, Gene Lawler, Ethan Port, William Chang, Martin Vingron, Pavel Pevzner, Esko Ukkonen, Frank Olken, Xin Xu, Alberto Apostolico, Thierry Lecroq, and George Havas. **Standing First Row:** Boris Pittel, Mudita Jain, Dominique Revuz, Gene Myers, Alejandro Schaffer, Steve Seiden, Dinesh Mehta, Tariq Choudhary, Tak Cheung Ip, L.J. Cummings, Rob Irving, Sampath Kannan, John Kececioglu, Pekka Kilpelainen, Kaizhong Zhang, Sun Wu, Lucas Hui, Tandy Warnow, and Yuh Dauh Lyuu. **Sitting second row:** Robert Paige, Ladan Rostami, Jong Yong Kim, Martin Farach, Haim Wolfson, Gad Landau, Jeanette Schmid, Glen Herrmannsfeldt, David Sankoff, Ramesh Hariharan, Laura Toniolo, Cari Soderlund, Dan Gusfield, and Wojciech Szpankowski. **Sitting first row:** Maxime Crochemore, Deborah Joseph, Xiaoqui Huang, Mereille Regnier, Dan Hirschberg, Marcella McClure, Gary Lewandowski, Taha Vasi, Brenda Baker, Campbell Fraser, and Philippe Jacquet. **On the floor:** John Oommen, Guy Jacobson, and Kiem Phong Vo.

# Table of Contents