

Lecture Notes in Computer Science

Edited by G. Goos, J. Hartmanis and J. van Leeuwen

1448

Martin Farach-Colton (Ed.)

Combinatorial Pattern Matching

9th Annual Symposium, CPM 98
Piscataway, New Jersey, USA, July 20-22, 1998
Proceedings



Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editor

Martin Farach-Colton
Bell Labs and
Rutgers University, Department of Computer Science
Piscataway, NJ 08855, USA
E-mail: farach@research.bell-labs.com

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Combinatorial pattern matching : 9th annual symposium ;
proceedings / CPM 98, Piscataway, New Jersey, USA, July 20 - 22,
1998. Martin Farach-Colton (ed.). - Berlin ; Heidelberg ; New York ;
Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ;
Singapore ; Tokyo : Springer, 1998
(Lecture notes in computer science ; Vol. 1448)
ISBN 3-540-64739-2

CR Subject Classification (1991): F.2.2, I.5.4, I.5.0, I.7.3, H.3.3, E.4, G.2.1

ISSN 0302-9743

ISBN 3-540-64739-2 Springer-Verlag 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-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1998
Printed in Germany

Typesetting: Camera-ready by author
SPIN 10638148 06/3142 - 5 4 3 2 1 0 Printed on acid-free paper

Program Committee

Martin Farach-Colton (Committee Chair)

Ricardo Baeza-Yates

Brenda Baker

Gary Benson

Richard Cole

Dannie Durand

Christos Faloutsos

Paolo Ferragina

Paola Merlo

S. Muthukrishnan

Mike Paterson

Prabhakar Raghavan

Jeanette Schmidt

Ron Shamir

Table of Contents

Session I

A Fast Bit-Vector Algorithm for Approximate String Matching
Based on Dynamic Programming 1
G. Myers

A Bit-Parallel Approach to Suffix Automata: Fast Extended String
Matching 14
G. Navarro and M. Raffinot

A Dictionary Matching Algorithm Fast on the Average for Terms of
Varying Lengths 34
M. Ziv-Ukelson and A. Kershenbaum

A Very Fast String Matching Algorithm for Small Alphabets and
Long Patterns 55
C. Charras, T. Lecroq and J. D. Pehoushek

Session II

Approximate Word Sequence Matching over Sparse Suffix Trees 65
K. M. Risvik

Efficient Parallel Algorithm for the Editing Distance Between
Ordered Trees 80
K. Zhang

Reporting Exact and Approximate Regular Expression Matches 91
E. G. Myers, P. Oliva and K. Guimarães

Session III

An Approximate Oracle for Distance in Metric Spaces 104
Y. Yang, K. Zhang, X. Wang, J. T. L. Wang and D. Shasha

A Rotation Invariant Filter for Two-Dimensional String Matching 118
K. Fredriksson and E. Ukkonen

Constructing Suffix Arrays for Multi-Dimensional Matrices 126
D. K. Kim, Y. A. Kim and K. Park

Session IV

Simple and Flexible Detection of Contiguous Repeats Using a
Suffix Tree 140
J. Stoye and D. Gusfield

Comparison and Coding DNA 153
C. N. S. Pedersen, R. Lyngs and J. Hein

Fixed Topology Alignment with Recombination <i>B. Ma, L. Wang and M. Li</i>	174
Aligning Alignments <i>J. D. Kececioglu and W. Zhang</i>	189
Session V	
Efficient Special Cases of Pattern Matching with Swaps <i>A. Amir, G. M. Landau, M. Lewenstein and N. Lewenstein</i>	209
Aligning DNA Sequences to Minimize the Changes in Protein <i>Y. Hua, T. Jiang and B. Wu</i>	221
Genome Halving <i>N. El-Mabrouk, J. H. Nadeau and D. Sankoff</i>	235
Author Index	251