

Table of Contents

Invited Papers

Patterns on the Web	1
<i>Krishna Bharat</i>	

Current Challenges in Bioinformatics	16
<i>João Meidanis</i>	

Web Algorithms

What's Changed? Measuring Document Change in Web Crawling for Search Engines	28
<i>Halil Ali and Hugh E. Williams</i>	

Link Information as a Similarity Measure in Web Classification	43
<i>Marco Cristo, Pavel Calado, Edleno Silva de Moura, Nivio Ziviani, and Berthier Ribeiro-Neto</i>	

A Three Level Search Engine Index Based in Query Log Distribution	56
<i>Ricardo Baeza-Yates and Felipe Saint-Jean</i>	

Bit-Parallel Algorithms

Row-wise Tiling for the Myers' Bit-Parallel Approximate String Matching Algorithm	66
<i>Kimmo Fredriksson</i>	

Alternative Algorithms for Bit-Parallel String Matching	80
<i>Hannu Peltola and Jorma Tarhio</i>	

Bit-Parallel Approximate String Matching Algorithms with Transposition	95
<i>Heikki Hyvrö</i>	

Compression

Processing of Huffman Compressed Texts with a Super-Alphabet	108
<i>Kimmo Fredriksson and Jorma Tarhio</i>	

(S,C)-Dense Coding: An Optimized Compression Code for Natural Language Text Databases	122
<i>Nieves R. Brisaboa, Antonio Fariña, Gonzalo Navarro, and María F. Esteller</i>	

Linear-Time Off-Line Text Compression by Longest-First Substitution . . .	137
<i>Shunsuke Inenaga, Takashi Funamoto, Masayuki Takeda, and Ayumi Shinohara</i>	

SCM: Structural Contexts Model for Improving Compression in Semistructured Text Databases	153
<i>Joaquín Adiego, Gonzalo Navarro, and Pablo de la Fuente</i>	

Categorization and Ranking

Ranking Structured Documents Using Utility Theory in the Bayesian Network Retrieval Model.	168
<i>Fabio Crestani, Luis M. de Campos, Juan M. Fernández-Luna, and Juan F. Huete</i>	

An Empirical Comparison of Text Categorization Methods	183
<i>Ana Cardoso-Cachopo and Arlindo L. Oliveira</i>	

Improving Text Retrieval in Medical Collections Through Automatic Categorization	197
<i>Rodrigo F. Vale, Berthier Ribeiro-Neto, Luciano R.S. de Lima, Alberto H.F. Laender, and Hermes R.F. Junior</i>	

Music Retrieval

A Bit-Parallel Suffix Automaton Approach for (δ, γ) -Matching in Music Retrieval.	211
<i>Maxime Crochemore, Costas S. Iliopoulos, Gonzalo Navarro, and Yoan J. Pinzon</i>	

Flexible and Efficient Bit-Parallel Techniques for Transposition Invariant Approximate Matching in Music Retrieval	224
<i>Kjell Lemström and Gonzalo Navarro</i>	

Multilingual Information Retrieval

FindStem: Analysis and Evaluation of a Turkish Stemming Algorithm . . .	238
<i>Hayri Sever and Yüctan Bitirim</i>	

Non-adjacent Digrams Improve Matching of Cross-Lingual Spelling Variants	252
<i>Heikki Keskitalo, Ari Pirkola, Kari Visala, Erkka Leppänen, and Kalervo Järvelin</i>	

The Implementation and Evaluation of a Lexicon-Based Stemmer	266
<i>Gilberto Silva and Claudia Oliveira</i>	

French Noun Phrase Indexing and Mining for an Information Retrieval System	277
<i>Hatem Haddad</i>	

Subsequences and Distributed Algorithms

New Refinement Techniques for Longest Common Subsequence Algorithms	287
<i>Lasse Bergroth, Harri Hakonen, and Juri Väisänen</i>	
The Size of Subsequence Automaton	304
<i>Zdeněk Troníček and Ayumi Shinohara</i>	
Distributed Query Processing Using Suffix Arrays	311
<i>Mauricio Marín and Gonzalo Navarro</i>	

Algorithms on Strings and Trees

BFT: Bit Filtration Technique for Approximate String Join in Biological Databases	326
<i>S. Alireza Aghili, Divyakant Agrawal, and Amr El Abbadi</i>	
A Practical Index for Genome Searching	341
<i>Heikki Hyvrö and Gonzalo Navarro</i>	
Using WordNet for Word Sense Disambiguation to Support Concept Map Construction	350
<i>Alberto J. Cañas, Alejandro Valerio, Juan Lalande-Pulido, Marco Carvalho, and Marco Arguedas</i>	
Memory-Adaptative Dynamic Spatial Approximation Trees	360
<i>Diego Arroyuelo, Francisca Muñoz, Gonzalo Navarro, and Nora Reyes</i>	
Large Edit Distance with Multiple Block Operations	369
<i>Dana Shapira and James A. Storer</i>	

Author Index	379
---------------------------	------------

String Processing and Information Retrieval

10th International Symposium, SPIRE 2003, Manaus,
Brazil, October 8-10, 2003, Proceedings

Nascimento, M.A.; Moura, E.S. de; Oliveira, A.L. (Eds.)
2003, CCCXCVI, 384 p., Softcover

ISBN: 978-3-540-20177-9