Accepted Manuscript

London stringology days and london algorithmic workshop (LSD & LAW) 2015

Jakub Radoszewski, Tomasz Radzik

PII: S1570-8667(16)30010-7
DOI: http://dx.doi.org/10.1016/j.jda.2016.06.002
Reference: JDA 658

To appear in: Journal of Discrete Algorithms

Please cite this article in press as: J. Radoszewski, T. Radzik, London stringology days and london algorithmic workshop (LSD & LAW) 2015, J. Discret. Algorithms (2016), http://dx.doi.org/10.1016/j.jda.2016.06.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Editorial

London Stringology Days and London Algorithmic Workshop (LSD & LAW) 2015

The London Stringology Days and London Algorithmic Workshop (LSD & LAW) meeting gathers each year dozens of international researchers who work on theoretical or applied aspects of discrete algorithms, including string algorithms, automata theory, combinatorics on words, bioinformatics, graph and network algorithms, and computational complexity. It has been organized by the Algorithms and Bioinformatics Group at King's College London since 2000, initially under the name of London Stringology Days.

The 2015 edition of the LSD & LAW meeting featured 23 presentations and was attended by over 65 researchers from all six inhabited continents. This special issue contains a selection of six articles related to the research topics presented and discussed during this meeting.

The paper Approximate search of short patterns with high error rates using the 01*0 lossless seeds by C. Vroland, M. Salson, S. Bini, and H. Touzet deals with the problems of exact and approximate pattern matching using a well established technique of seeds. The authors introduce a new type of seeds that prove, both by a theoretical study of the filtering efficiency and a comparison with other methods employing seeds, to be especially well suited for short patterns and high error rates. An open source implementation of the algorithm prepared by the authors is also available.

In the article Lightweight LCP Construction for Very Large Collections of Strings the authors A. J. Cox, F. Garofalo, G. Rosone, and M. Sciortino present a lightweight algorithm for the simultaneous computation of the longest common prefix array and the Burrows-Wheeler transform of a very large collection of strings. An analysis of the algorithm in terms of time complexity, disk I/O operations, and memory and disk space is presented, as well as tests of the algorithm on real data.

Another included contribution, by W. H. A. Tustumi, S. Gog, G. P. Telles, and F. A. Louza, is entitled An improved algorithm for the all-pairs suffix-prefix problem. The main result is an optimized version of an asymptotically-optimal algorithm by Ohlebusch and Gog from 2010 for computing the longest overlaps between pairs of strings in a collection. In comparison with the earlier version it turns out to be 2.6 times faster and by 15% more space-efficient on average.

L. Bulteau, G. Fertin, and C. Komusiewicz’s paper (Prefix) Reversal Distance for (Signed) Strings with Few Blocks or Small Alphabets investigates in more detail the complexity of the NP-hard problem of computing the minimum number of substring reversals needed to transform a given string $S$ to another given string $T$. The paper shows that several restricted versions of this problem remain NP-hard and identifies an input parameter which makes these problems fixed-parameter tractable.

B. Cazaux and E. Rivals in their paper entitled A linear time algorithm for Shortest Cyclic Cover of Strings consider the problem of finding the Shortest Cyclic Cover of Strings (SCCS), which plays an important role in approximation algorithms for the shortest superstring problem. They analyze greedy solutions of the SCCS problem and use the developed insight to show the first linear-time SCCS algorithm.

The paper Complete Parsimony Haplotype Inference Problem and Algorithms Based on Integer Programming, Branch-and-Bound and Boolean Satisfiability by G. Jäger, S. Climer, and W. Zhang is a study of the optimization model for Haplotype Inference which minimizes the number of haplotypes explaining a given set of genotypes. The authors extend the methods previously proposed for finding one optimal solution for this model to the case when we want to compute all optimal solutions. They propose a number of interesting techniques for speeding up the computation and present a comprehensive experimental evaluation of their algorithms.
We would like to express our gratitude to all the anonymous reviewers for timely and thorough reviewing of the articles. Special thanks must also go to the authors for their patience and meticulous revisions during this process. Finally, we would also like to warmly thank the Editorial Board of the Journal of Discrete Algorithms for giving us the opportunity to present this special issue to a wider audience.

Guest Editors of the special issue
Jakub Radoszewski and Tomasz Radzik
Department of Informatics, King's College London, London, UK
E-mail addresses: jakub.radoszewski@kcl.ac.uk, tomasz.radzik@kcl.ac.uk