Editorial

The annual research meetings *London Stringology Days* and *London Algorithmic Workshop* have now been organized at King's College London for a number of years, usually together since it is always stimulating, and often unavoidable, to keep the stringology research within its wider algorithmic context. The idea of this special issue of *Journal of Discrete Algorithms* on *Stringology, Bioinformatics and Algorithms* was put forward at the 2010 edition of these meetings, which were held in February 2010. Consequently, the results in some of the papers in this issue can be traced back to the presentations and discussions during those meetings. We believe that the selected papers together illustrate well how stringology, which can be defined as research on combinatorial and computational properties of patterns in sequences of characters (called "strings" or "words"), is intrinsically based on general algorithmic techniques, in particular data structures, while forming itself an important part of the foundations of bioinformatics.

Two papers in this special issue are examples of the current research in the area of combinatorics on words: "The three squares lemma revisited" by Evguenia Kopylova and William Smyth, and "On the structure of compacted subword graphs of Thue–Morse words and their applications" by Jakub Radoszewski and Wojciech Rytter. While the direct aim of this fundamental part of stringology is to improve our insight into the structure of strings, the obtained combinatorial results often lead, sooner or later, to more efficient string algorithms.

The other stringology papers in this issue are focused more on computational aspects. Efficient algorithms for the classical string pattern matching problem are considered in two papers: "On the bit-parallel simulation of the nondeterministic Aho–Corasick and suffix automata for a set of patterns" by Domenico Cantone, Simone Faro and Emanuele Giaquinta, and "String matching with alphabet sampling" by Francisco Claude, Gonzalo Navarro, Hannu Peltola, Leena Salmela and Jorma Tarhio. In the first paper the efficiency of pattern matching computation is derived from fast operations on vectors of bits, which are provided by most processors. The second paper investigates the approach called "alphabet sampling", which projects strings on a suitably chosen subset of the alphabet to reduce their lengths and, in turn, to speed-up the running time of pattern matching computation.

String algorithms with applications to data compression, including compression of DNA sequences, are considered in "Efficient algorithms for three variants of the LPF table" by Maxime Crochemore, Costas Iliopoulos, Marcin Kubica, Wojciech Rytter and Tomasz Waleń, and "Searching for smallest grammars on large sequences and application to DNA" by Rafael Carrascosa, François Coste, Matthias Gallé and Gabriel Infante-Lopez. The first paper presents a theoretical analysis of the running time of the proposed algorithms, while the second one uses a DNA corpus to evaluate trade-offs between the running time and the compression ratio of the considered approach. The compression methods discussed in these papers show that in addition to the obvious benefit of saving the space, they can be also used for discovering structural properties of sequences.

The paper "Steiner tree reoptimization in graphs with sharpened triangle inequality", by Hans-Joachim Böckenhauer, Karin Freiermuth, Juraj Hromkovič, Tobias Mömke, Andreas Sprock and Björn Steffen, is an example of the research on the complexity of re-optimizing, that is, of updating a known solution when the input undergoes a small, local change. While this paper considers a classical graph-theoretical problem, the general idea of re-optimization is bound to take more prominent place in stringology, especially in its part which is directly motivated by bioinformatics. For some bioinformatics problems it may be natural to expect high similarity among different data inputs, for example, when the input represents a molecular sequence. In such cases re-optimization may be a promising approach. An example of re-optimization in stringology is the paper "On the number of elements to reorder when updating a suffix array" by Martine Léonard, Laurent Mouchard and Mikaël Salson, which analyzes the complexity of updating some data structures representing a large input string (a DNA sequence or a natural language text), when that string is modified.

We would like to thank the Editors-in-Chief of the Journal of Discrete Algorithm for inviting us to edit this special issue, and the authors for submitting their papers. All papers have undergone the rigorous refereeing process, and we are grateful to the referees for their professionalism and willingness to help.

Guest Editors

Tomasz Radzik *

King's College London, Department of Informatics, London, United Kingdom

E-mail address: tomasz.radzik@kcl.ac.uk

German Tischler *

Universität Würzburg, Lehrstuhl für Informatik II, Würzburg, Germany

E-mail address: tischler@informatik.uni-wuerzburg.de

---

* Corresponding authors.