

Special Issue — COMPUTING PATTERNS IN STRINGS

Foreword

A *string* or *word* or *sequence* is an extremely simply mathematical object (data structure) that has only been deemed worthy of particular study — first by mathematicians, then later by computer scientists — since Axel Thue’s groundbreaking work nearly a century ago. Until about 35 years ago, the published work discussed mathematical properties of strings or of special classes of strings; however, the simultaneous advent of the computer and of major application areas (especially molecular biology and massive data storage/transmission) shifted the emphasis toward computer science and algorithms on strings. Today there are at least several hundred research papers published every year that deal with algorithms on strings and related methodology; an annual conference, Combinatorial Pattern Matching (CPM), has since 1987 published research contributions to this field of the highest quality. Since the mid-1990s there have been several books [1–8] published that seek to provide a unified view of algorithms on strings and/or their numerous applications. These application areas include natural and obvious ones such as computational biology, data compression and cryptography, but also many others: the identification of musical motifs, internet routing, pattern recognition, computer graphics, and so on. Meanwhile, during the last 10 years, in many countries around the world, undergraduate and graduate programmes have been established whose subject matter includes heavy emphasis on string algorithms — perhaps chiefly bio-informatics programmes, but also others.

Thus a special issue of *Fundamenta Informaticae* dealing with the computation of patterns in strings is both timely and relevant within the broader contexts of computer science and combinatorial algorithms. It is my hope that this issue will provide a snapshot of current research in string algorithms, perhaps motivate other researchers to take an interest in a dynamic and exciting research area. Certainly the papers collected here reflect the range of topics mentioned above: altogether six of them deal with subjects of particular interest in a biological context (four with some form of approximate pattern-matching, two with string alignment); three others treat aspects of data compression; and two final papers provide fresh insights into standard string-processing calculations (longest common prefix, suffix array).

Altogether 17 papers were submitted to the special issue, all of them of high quality; after an extended refereeing process, 11 were accepted for publication. It is a pleasure to thank the authors for their interest and for their original contributions, also to express my particular appreciation of the careful and diligent, but of necessity unattributable, work of a cadre of some 35 referees.

Bill Smyth

Algorithms Research Group
McMaster University

Department of Computing
Curtin University of Technology

- [1] Maxime Crochemore & Wojciech Rytter, *Text Algorithms*, Oxford University Press (1994) 412 pp.
- [2] Graham A. Stephen, *String Searching Algorithms*, World Scientific Publishing (1994) 243 pp.
- [3] Dan Gusfield, *Algorithms on Strings, Trees & Sequences*, Cambridge University Press (1997) 534 pp.
- [4] João Setubal & João Meidanis, *Introduction to Computational Molecular Biology*, PWS Publishing (1997) 296 pp.
- [5] Maxime Crochemore, Christophe Hancart & Thierry Lecroq, *Algorithmique du Texte*, Vuibert (2001) 347 pp.
- [6] Wojciech Szpankowski, *Average Case Analysis of Algorithms on Sequences*, Wiley-Interscience (2001) 551 pp.
- [7] Gonzalo Navarro & Mathieu Raffinot, *Flexible Pattern Matching in Strings*, Cambridge University Press (2002) 221 pp.
- [8] Bill Smyth, *Computing Patterns in Strings*, Pearson Addison-Wesley (2003) 423 pp.