Preface

This book is about algorithms on texts, also called algorithmic stringology. Text (word, string, sequence) is one of the main unstructured data types and the subject is of vital importance in computer science.

The subject is versatile because it is a basic requirement in many sciences, especially in computer science and engineering. The treatment of unstructured data is a very lively area and demands efficient methods owing both to their presence in highly repetitive instructions of operating systems and to the vast amount of data that needs to be analysed on digital networks and equipments. The latter is clear for information technology companies that manage massive data in their data centres but also holds for most scientific areas beyond Computer science.

The book presents a collection of the most interesting representative problems in stringology. They are introduced in a short and pleasant way and open doors to more advanced topics. They were extracted from hundreds of serious scientific publications, some of which are more than a hundred years old and some are very fresh and up to date. Most of the problems are related to applications while others are more abstract. The core part of most of them is an ingenious short algorithmic solution except for a few introductory combinatorial problems.

This is not just yet another monograph on the subject but a series of problems (puzzles and exercises). It is a complement to books dedicated to the subject in which topics are introduced in a more academic and comprehensive way. Nevertheless, most concepts in the field are included in the book, which fills a missing gap and is very expected and needed, especially for students and teachers, as the first problem-solving textbook of the domain.

https://doi.org/10.1017/9781108835831.001 Published online by Cambridge University Press

The book is organised into seven chapters:

'The Very Basics of Stringology' is a preliminary chapter introducing the terminology, basic concepts and tools for the next chapters and that reflects six main streams in the area.

'Combinatorial Puzzles' is about combinatorics on words, an important topic because many algorithms are based on combinatorial properties of their input.

'Pattern Matching' deals with the most classical subject, text searching and string matching.

'Efficient Data Structures' is about data structures for text indexing. They are used as fundamental tools in a large number of algorithms, such as special arrays and trees associated with texts.

'Regularities in Words' concerns regularities that occur in texts, in particular repetitions and symmetries, that have a strong influence on the efficiency of algorithms.

'Text Compression' is devoted to several methods of the practically important area of conservative text compression.

'Miscellaneous' contains various problems that do not fit in earlier chapters but certainly deserve presentation.

Problems listed in the book have been accumulated and developed over several years of teaching on string algorithms in our own different institutions in France, Poland, UK and USA. They have been taught mostly to master's students and are given with solutions as well as with references for further readings. The content also profits from the experience authors gained in writing previous textbooks.

Anyone teaching graduate courses on data structures and algorithms can select whatever they like from our book for their students. However, the overall book is not elementary and is intended as a reference for researchers, PhD and master's students, as well as for academics teaching courses on algorithms even if they are not directly related to text algorithms. It should be viewed as a companion to standard textbooks on the domain. The self-contained presentation of problems provides a rapid access to their understanding and to their solutions without requiring a deep background on the subject.

The book is useful for specialised courses on text algorithms, as well as for more general courses on algorithms and data structures. It introduces all required concepts and notions to solve problems but some prerequisites in bachelor- or sophomore-level academic courses on algorithms, data structures and discrete mathematics certainly help in grasping the material more easily.

https://doi.org/10.1017/9781108835831.001 Published online by Cambridge University Press