Algorithms and Data Structures

Helmut Knebl

Algorithms and Data Structures

Foundations and Probabilistic Methods for Design and Analysis

Springer

Helmut Knebl
Fakultät Informatik
Technische Hochschule Nürnberg
Nürnberg, Germany

Ursprünglich veröffentlicht in deutscher Sprache:
Algorithmen und Datenstrukturen von Prof. Dr. Helmut Knebl

ISBN 978-3-030-59757-3 ISBN 978-3-030-59758-0 (eBook)
https://doi.org/10.1007/978-3-030-59758-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Many practical problems can be solved by algorithms. For this reason, computer algorithms are diverse and ubiquitous today. The spectrum ranges from one of the oldest recorded algorithms, the algorithm of Euclid from the third century B.C., to algorithms for the investigation of large amounts of data, algorithms for communication and searching on the Internet, algorithms for imaging procedures and for diagnostics in medical technology, and algorithms for assistance systems in cars, engine control or the control of household appliances. Algorithms are the subject of intensive research and belong to the fundamental concepts of computer science. The design of efficient algorithms and their analysis with regard to resource requirements are fundamental for the development of computer programs. Therefore, the subject Algorithms and Data Structures is a central component of any computer science curriculum.

This book originates from lectures on algorithms and data structures for students of computer science, media and business informatics at the Technische Hochschule Nürnberg Georg Simon Ohm. The basic topics of the book are covered in the bachelor's courses. Advanced parts, such as randomized algorithms, are reserved for master's courses.

The algorithms of the first chapter, all of which are popular algorithms, are studied to introduce common design principles for the development of algorithms. The following Chapters 2 to 6 are organized by problem areas. We consider the problem of storing and retrieving elements of a set and problems that can be formulated with graphs. For the first problem, we use three techniques to efficiently implement these operations: Sorting with binary search, search trees and hashing. The first two methods require ordered sets, the last method requires that the elements of the set are uniquely identified by keys.

The sorting methods quicksort and heapsort, binary search and searching for the kth-smallest element are the subject of Chapter 2. Special attention is paid to the analysis of the running time of the algorithms. Throughout all chapters, the aim is to develop explicit formulas or precise estimates for the running time. Difference equations are used as a solution method. This allows exact and not only asymptotic statements to be made about the running times of algorithms. We use a standardized method for the running time calculation: First, establish a difference equation for the running time and then solve the equation with known methods.

Hash functions, in particular universal families of hash functions, methods for the treatment of collisions and a detailed analysis of hash procedures are the subject of Chapter 3.

Chapter 4 deals with binary search trees, AVL trees and randomized binary search trees. B-trees are used to store data on secondary storage. Code trees for the graphical representation of codes for data compression complete the chapter.

Graphs play a fundamental role in many areas of computer science. For many graph problems, solutions exist in the form of efficient algorithms. In Chapter 5 breadth-first search and depth-first search for graphs are studied and as an application topological sorting and the calculation of the strongly connected components. Fundamental optimization problems, such as the construction of minimum spanning trees and shortest paths as well as the flow problem in networks, are the contents of Chapter 6.

Probabilistic methods are fundamental for the construction of simple and efficient algorithms. In each chapter at least one problem is solved by using a randomized algorithm. In detail, it is about the verification of the identity of polynomials, the randomized version of quicksort and quickselect, universal families of hash functions and randomized binary search trees. Randomized algorithms for the computation of a minimal section in a graph and for the construction of a minimum spanning tree for a weighted graph are among the advanced topics.

The book focuses on algorithms. Data structures are discussed as far as they are needed for the implementation of the algorithms. The selection of the topics is mainly based on the aim to treat elementary algorithms which have a wide field of application. The aim is a detailed and in-depth study.

The text assumes experience in programming algorithms, especially with elementary data structures – such as chained lists, queues and stacks – in the scope of the contents of the programming lectures of the first year of computer science studies. Familiarity with mathematical methods covered in the first year is also desirable. For the convenience of the reader, the mathematical methods necessary for understanding, in particular elementary solution methods for difference equations and special probability distributions, are repeated in the first chapter and in the appendix.

The formulation of the algorithms by using pseudo-code focuses on the essentials and thus makes the idea of the algorithm clear. It is sufficiently precise to allow considerations on the correctness and calculations of the running time to be carried out. More than 100 figures illustrate the algorithms. Many examples help the reader to understand the individual steps of the algorithms. Numerous exercises complete each chapter and help the reader to practice and deepen the material. Answers to the exercises are provided on the webpage for this book: www.in.th-nuernberg.de/Knebl/Algorithms.

This book stems from lectures on algorithms and data structures that I taught at the Technische Hochschule Nürnberg Georg Simon Ohm for many years. During this time, the university changed its name twice and still re-

mained the same. During the preparation of the lecture I used the textbooks listed in Section 1.8.

I received a lot of support for the completion of the book. My colleagues Jens Albrecht, Christian Schiedermeier and especially Alexander Kröner have carefully looked through parts of it, which has led to the correction of mistakes and ambiguities. I owe Harald Stieber valuable suggestions and discussions, which have contributed to the improvement of the book. I would like to express my sincere thanks to all those who have supported me, including those not mentioned. I would especially like to thank my students, who have attended the lecture with dedication in the past years, worked diligently on exercises and helped to track down mistakes.

The book is essentially the translation of [Knebl19]

> Algorithmen und Datenstrukturen
> Grundlagen und probabilistische
> Methoden für den Entwurf und die Analyse

The content and structure of the text is the same as that of the German edition. During the translation, the book was also thoroughly reviewed and the presentation improved in many places. Inaccuracies of the German edition have been corrected.

I am grateful to Patricia Brockmann and Sebastian Knebl for their support in proofreading, and I thank Ronan Nugent and Sybille Thelen at Springer for their pleasant and valuable cooperation.

Nürnberg, September 2020 Helmut Knebl

Contents