



Data Structures and Algorithms in Java

Third Edition





Data Structures and Algorithms in Java

Third Edition

Michael T. Goodrich

Department of Computer Science
University of California, Irvine

Roberto Tamassia

Department of Computer Science
Brown University



John Wiley & Sons, Inc.

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

ACQUISITIONS EDITOR	Paul Crockett
ASSOCIATE MARKETING DIRECTOR	Ilse Wolfe
EDITORIAL ASSISTANT	Simon Durkin
SENIOR PRODUCTION EDITOR	Ken Santor
COVER DESIGNER	David Levy

About the cover: Anatjari Tjampitjinpa, "Snake Dreaming," 1988. Reproduced courtesy of Patrick Corbally Stourton of Corbally Stourton Contemporary Art Ltd.

This book was set in L^AT_EX by the authors and printed and bound by Von Hoffmann, Inc. The cover was printed by Von Hoffmann, Inc.

This book is printed on acid free paper. ∞

Trademark Acknowledgments: *Java* is a trademark of Sun Microsystems, Inc. *UNIX*[®] is a registered trademark in the United States and other countries, licensed through X/Open Company, Ltd. *Powerpoint*[®] is a trademark of Microsoft Corporation. All other product names mentioned herein are the trademarks of their respective owners.

Copyright © 2004, John Wiley & Sons, Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc. 222 Rosewood Drive, Danvers, MA 01923, (978)750-8400, fax (978)750-4470.

Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201)748-6011, fax (201)748-6008, E-Mail: PERMREQ@WILEY.COM.

To order books or for customer service please call 1-800-CALL WILEY (225-5945).

ISBN 0-471-46983-1

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1



To Karen, Paul, Anna, and Jack
– *Michael T. Goodrich*

To Isabel
– *Roberto Tamassia*





Preface to the Third Edition

This third edition is designed to provide an introduction to data structures and algorithms, including their design, analysis, and implementation. With respect to computer science and computer engineering curricula, we have written this book to be primarily for Freshman-Sophomore level data structures courses. In terms of curricula based on the *IEEE/ACM 2001 Computing Curriculum*, this book is appropriate for use in the courses CS102 (I/O/B versions), CS103 (I/O/B versions), CS111 (A version), and CS112 (A/I/O/F/H versions). In terms of curricula based on the *1978 ACM Computer Science Curriculum*, this book is intended for the CS2 course. We discuss its use for such courses in more detail later in this preface.

The major changes, with respect to the second edition, are the following:

- Expanded coverage of recursion
- Streamlined mathematics to the seven most used functions
- Inclusion of HTML tag matching and favorites list applications
- Simplification of the Java interfaces associated with ADTs
- Better integration with the Java Collections Framework
- Simplified binary tree and priority queue ADTs
- Addition of the map ADT
- Addition of splay trees
- Addition of the union/find partition data structure
- Simplified graph ADT
- Expanded and revised exercises.

This book is related to the following books:

- M.T. Goodrich, R. Tamassia, and D.M. Mount, *Data Structures and Algorithms in C++*, John Wiley & Sons, Inc., 2004. This book has a similar overall structure to the present book, but uses C++ as the underlying language (with some modest, but necessary pedagogical differences required by this approach). Thus, it could make for a handy companion book in a curriculum that allows for either a Java or C++ track in the introductory courses.
- M.T. Goodrich and R. Tamassia, *Algorithm Design: Foundations, Analysis, and Internet Examples*, John Wiley & Sons, Inc., 2002. This is a textbook for a more advanced algorithms and data structures course, such as CS7 in the ACM 1978 curriculum or CS210 (T/W/C/S versions) in the IEEE/ACM 2001 curriculum.

Use as a Textbook

The design and analysis of efficient data structures has long been recognized as a vital subject in computing, for the study of data structures is part of the core of every collegiate computer science and computer engineering major program we are familiar with. Typically, the introductory courses are presented as a two- or three-course sequence. Elementary data structures are often briefly introduced in the first programming or introduction to computer science course and this is followed by a more in-depth introduction to data structures in the following course(s). Furthermore, this course sequence is typically followed at a later point in the curriculum by a more in-depth study of data structures and algorithms. We feel that the central role of data structure design and analysis in the curriculum is fully justified, given the importance of efficient data structures in most software systems, including the Web, operating systems, databases, compilers, and scientific simulation systems.

With the emergence of the object-oriented paradigm as the framework of choice for building robust and reusable software, we have tried to take a consistent object-oriented viewpoint throughout this text. One of the main ideas of the object-oriented approach is that data should be presented as being encapsulated with the methods that access and modify them. That is, rather than simply viewing data as a collection of bytes and addresses, we think of data as instances of an *abstract data type (ADT)* that includes a repertory of methods for performing operations on the data. Likewise, object-oriented solutions are often organized utilizing common *design patterns*, which facilitate software reuse and robustness. Thus, we present each data structure using ADTs and their respective implementations and we introduce important design patterns as means to organize those implementations into classes, methods, and objects.

For each ADT presented in this book, we provide an associated Java interface. Also, concrete data structures realizing the ADTs are provided as Java classes implementing the above interfaces. We also give Java implementations of fundamental algorithms (such as sorting and graph traversals) and of sample applications of data structures (such as HTML tag matching and a photo album). Due to space limitations, we sometimes show only code fragments in the book and make the full source code available on the companion Web site <http://java.datastructures.net>.

The Java code in this the book implementing fundamental data structures is organized in a single Java package, `net.datastructures`. This package forms a coherent library of data structures and algorithms in Java specifically designed for educational purposes in a way that is complementary with the Java Collections Framework.

Web Added-Value Education

This book is accompanied by an extensive Web site:

<http://java.datastructures.net>

Students are encouraged to use this site along with the book, to help with exercises and increase understanding of the subject. Instructors are likewise welcome to use the site to help plan, organize, and present their course materials.

For the Student

For all readers, and specifically for students, we include:

- All the Java source code presented in this book
- The student version of the `net.datastructures` package
- Slide handouts (four-per-page) in PDF format
- A database of hints to *all* exercises, indexed by problem number
- Java animations and interactive applets for data structures and algorithms
- Hyperlinks to other data structures and algorithms resources.

We feel that the Java animations and interactive applets should be of particular interest, since they allow readers to interactively “play” with different data structures, which leads to better understanding of the different ADTs. In addition, the hints should be of considerable use to anyone needing a little help getting started on certain exercises.

For the Instructor

For instructors using this book, we include the following additional teaching aids:

- Solutions to over two hundred of the book’s exercises
- A keyword-searchable database of additional exercises
- The complete `net.datastructures` package
- Additional Java source code
- Slides in Powerpoint and PDF (one-per-page) format
- Self-contained special-topic supplements, including discussions on convex hulls, range trees, and orthogonal segment intersection.

The slides are fully editable, so as to allow an instructor using this book full freedom in customizing his or her presentations.

For the Instructor

This book contains many Java-code and pseudo-code fragments, and over six hundred exercises, which are divided into roughly 40% reinforcement exercises, 40% creativity exercises, and 20% programming projects.

Relation to the IEEE/ACM 2001 Computing Curriculum

This book can be used for the CS2 course in the 1978 ACM Computer Science Curriculum or courses CS102 (I/O/B versions), CS103 (I/O/B versions), CS111 (A version), and/or CS112 (A/I/O/F/H versions) in the IEEE/ACM 2001 Computing Curriculum, with instructional units as outlined in Table 0.1.

Instructional Unit	Relevant Material
PL1. Overview of Programming Languages	Chapters 1 & 2
PL2. Virtual Machines	Sections 4.2.3, 4.3.4, & 12.4.4
PL3. Introduction to Language Translation	Section 1.9
PL4. Declarations and Types	Sections 1.1, 2.4, & 2.4.4
PL5. Abstraction Mechanisms	Sections 2.4, 4.2, 4.3, 4.5, 5.1.1, 5.2, 5.3, 5.5, 6.1, 6.3.1, 7.1, 8.1, 8.3, 10.6, & 12.1
PL6. Object-Oriented Programming	Chapters 1 & 2 and Sections 5.2.2, 5.5, 6.3.5, 7.1.2, & 12.3.1
PF1. Fundamental Programming Constructs	Chapters 1 & 2
PF2. Algorithms and Problem-Solving	Sections 1.9, 3.2, & 3.4
PF3. Fundamental Data Structures	Sections 1.5, 4.2–4.4, 4.5, , 5.1–5.3, 6.1, 6.3, 6.4, 7.1, 7.3, 8.1–8.4, 9.1–9.6, 12.1, & 12.2
PF4. Recursion	Sections 2.5.1 & 4.1
SE1. Software Design	Chapter 2 and Sections 5.2.2, 5.5, 6.3.5, 7.1.2, & 12.3.1
SE2. Using APIs	Sections 2.4, 2.4.4, 4.2, 4.3, 4.5, 5.1.1, 5.2, 5.3, 5.5, 6.1, 6.3.1, 7.1, 8.1, 8.3, 10.6, & 12.1
AL1. Basic Algorithmic Analysis	Chapter 3
AL2. Algorithmic Strategies	Sections 10.1.1, 10.7.1, 11.2.1, 11.4.2, & 11.5.2
AL3. Fundamental Computing Algorithms	Sections 7.1.4, 7.2.3, 7.3.5, 8.2, & 8.3.3, and Chapters 10, 11, & 12
DS1. Functions, Relations, and Sets	Sections 3.3, 7.1, & 10.6
DS3. Proof Techniques	Sections 3.5, 5.1.3, 6.3.3, 7.3, 9.2, 9.3, 9.4, 9.5, 10.2.1, 10.3, 10.6.2, 12.1, 12.3.1, 12.4, & 12.5
DS4. Basics of Counting	Sections 2.2.3 & 10.1.5
DS5. Graphs and Trees	Chapters 6, 7, 9, & 12
DS6. Discrete Probability	Appendix A and Sections 8.2.2, 8.4.2, 10.2.1, & 10.7

Table 0.1: Material for Units in the IEEE/ACM 2001 Computing Curriculum.

Teaching Options

This book is also structured to allow instructors a great deal of freedom in organizing and presenting the material. In Table 0.2, we illustrate some possible uses of this book for an intermediate-paced data structures course.

Chapter	Possible Options
1. Java Programming	skip if students know it already
2. Object-Oriented Design	omit the adapter pattern
3. Analysis Tools	omit justification methods
4. Stacks, Queues & Recursion	omit Java-thread queue application
5. Vectors, Lists, and Sequences	omit list iterators
6. Trees	omit template method pattern
7. Priority Queues	omit bottom-up heap construction
8. Maps and Dictionaries	omit extensions of dictionaries
9. Search Trees	omit (2,4) and red-black trees or splay trees
10. Sorting, Sets, and Selection	omit sorting lower bound
11. Text Processing	omit tries, compression, and/or LCS
12. Graphs	omit directed graphs or MSTs

Table 0.2: Options for a Freshman-Sophomore data structures course.

Prerequisites

We have written this book assuming that the reader comes to it with certain knowledge. That is, we assume that the reader is at least vaguely familiar with a high-level programming language, such as C, C++, or Java, and that he or she understands the main constructs from such a high-level language, including:

- Variables and expressions
- Methods (also known as functions or procedures)
- Decision structures (such as if-statements and switch-statements)
- Iteration structures (for-loops and while-loops).

For readers who are familiar with these concepts, but not with how they are expressed in Java, we provide a primer on the Java language in Chapter 1. Still, this book is primarily a data structures book, not a Java book; hence, it does not provide a comprehensive treatment of Java. Nevertheless, we do not assume that the reader is necessarily familiar with object-oriented design or with linked structures, such as linked lists, for these topics are covered in the core chapters of this book.

In terms of mathematical background, we assume the reader is somewhat familiar with topics from high-school mathematics. Even so, in Chapter 3, we discuss the seven most-important functions for algorithm analysis. In fact, sections that use something other than one of these seven functions are considered optional, and are indicated with a star (★). We give a summary of other useful mathematical facts, including elementary probability, in Appendix A.

About the Authors

Professors Goodrich and Tamassia are well-recognized researchers in algorithms and data structures, having published many papers in this field, with applications to Internet computing, information visualization, computer security, and geometric computing. They have served as principal investigators in several joint projects sponsored by the National Science Foundation, the Army Research Office, and the Defense Advanced Research Projects Agency. They are also active in educational technology research, with special emphasis on algorithm visualization systems.

Michael Goodrich received his Ph.D. in Computer Science from Purdue University in 1987. He is currently a professor in the Department of Computer Science at University of California, Irvine. Previously, he was a professor at Johns Hopkins University. He is an editor for the *International Journal of Computational Geometry & Applications* and *Journal of Graph Algorithms and Applications*.

Roberto Tamassia received his Ph.D. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 1988. He is currently a professor in the Department of Computer Science at Brown University. He is editor-in-chief for the *Journal of Graph Algorithms and Applications* and an editor for *Computational Geometry: Theory and Applications*. He previously served on the editorial board of *IEEE Transactions on Computers*.

In addition to their research accomplishments, the authors also have extensive experience in the classroom. For example, Dr. Goodrich has taught data structures and algorithms courses, including Data Structures as a freshman-sophomore level course and Introduction to Algorithms as an upper level course. He has earned several teaching awards in this capacity. His teaching style is to involve the students in lively interactive classroom sessions that bring out the intuition and insights behind data structuring and algorithmic techniques. Dr. Tamassia has taught Data Structures and Algorithms as an introductory freshman-level course since 1988. One thing that has set his teaching style apart is his effective use of interactive hypermedia presentations integrated with the Web.

The instructional Web sites, datastructures.net and algorithmdesign.net, supported by Drs. Goodrich and Tamassia are used as reference material by students, teachers, and professionals worldwide.

Acknowledgments

There are a number of individuals who have made contributions to this book.

We are grateful to all our research collaborators and teaching assistants, who provided feedback on early drafts of chapters and have helped us in developing exercises, programming assignments, and algorithm animation systems. In particular, we would like to thank Jeff Achter, James Baker, Ryan Baker, Benjamin Boer, Mike Boilen, Devin Borland, Lubomir Bourdev, Stina Bridgeman, Bryan Cantrill, Yi-Jen Chiang, Robert Cohen, David Emory, Jody Fanto, Ashim Garg, Natasha Gelfand, Mark Handy, Michael Horn, Benoît Hudson, Jovanna Ignatowicz, Seth Padowitz, James Piechota, Dan Polivy, Susannah Raub, Andy Schwerin, Michael Shapiro, Michael Shin, Christian Straub, Galina Shubina, Nikos Triandopoulos, and Luca Vismara. Lubomir Bourdev, Mike Demmer, Mark Handy, Michael Horn, and Scott Speigler developed a basic Java tutorial, which ultimately led to Chapter 1, Java Programming.

Special thanks go to Eric Zamore, who contributed to the development of the Java code examples in this book and to the design, implementation and testing of the `net.datastructures` library of data structures and algorithms in Java.

Many students and instructors have used the two previous editions of this book and their experiences and responses have helped shape this third edition.

There have been a number of friends and colleagues whose comments have led to improvements in the text. We are particularly thankful to Karen Goodrich, Art Moorshead, David Mount, Scott Smith and Ioannis Tollis for their insightful comments. In addition, contributions by David Mount to Section 2.5.1 and to several figures are gratefully acknowledged.

We are also truly indebted to the outside reviewers and readers for their copious comments, emails, and constructive criticism, which were extremely useful in writing the third edition. We specifically thank the following reviewers for their comments and suggestions: Divy Agarwal, University of California, Santa Barbara; Terry Andres, University of Manitoba; Bobby Blumofe, University of Texas, Austin; Michael Clancy, University of California, Berkeley; Larry Davis, University of Maryland; Scott Drysdale, Dartmouth College; Arup Guha, University of Central Florida; Chris Ingram, University of Waterloo; Stan Kwasny, Washington University; Calvin Lin, University of Texas at Austin; John Mark Mercer, McGill University; Laurent Michel, University of Connecticut; Leonard Myers, California Polytechnic State University, San Luis Obispo; David Naumann, Stevens Institute of Technology; Robert Pastel, Michigan Technological University; Bina Ramamurthy, SUNY Buffalo; Ken Slonneger, University of Iowa; C.V. Ravishankar, University of Michigan; Val Tannen, University of Pennsylvania; Paul Van Aragon, Messiah College; and Christopher Wilson, University of Oregon.

We are grateful to our editor, Paul Crockett, for his enthusiastic support of this project. The team at Wiley has been great. Many thanks go to Lilian Brady, Ken Santor, Simon Durkin, Madelyn Lesure, Dawn Stanley, and Jeri Warner.

The computing systems and excellent technical support staff in the departments of computer science at Brown University and University of California, Irvine gave us reliable working environments. This manuscript was prepared primarily with the L^AT_EX typesetting package for the text and Adobe FrameMaker[®] and Microsoft Visio[®] for the figures.

Finally, we would like to warmly thank Isabel Cruz, Karen Goodrich, Giuseppe Di Battista, Franco Preparata, Ioannis Tollis, and our parents for providing advice, encouragement, and support at various stages of the preparation of this book. We also thank them for reminding us that there are things in life beyond writing books.

Michael T. Goodrich
Roberto Tamassia