The Design of Approximation Algorithms

Discrete optimization problems are everywhere, from traditional operations research planning problems, such as scheduling, facility location, and network design; to computer science problems in databases; to advertising issues in viral marketing. Yet most such problems are NP-hard. Thus unless P = NP, there are no efficient algorithms to find optimal solutions to such problems. This book shows how to design approximation algorithms: efficient algorithms that find provably near-optimal solutions.

The book is organized around central algorithmic techniques for designing approximation algorithms, including greedy and local search algorithms, dynamic programming, linear and semidefinite programming, and randomization. Each chapter in the first part of the book is devoted to a single algorithmic technique, which is then applied to several different problems. The second part revisits the techniques but offers more sophisticated treatments of them. The book also covers methods for proving that optimization problems are hard to approximate.

Designed as a textbook for graduate-level algorithms courses, the book will also serve as a reference for researchers interested in the heuristic solution of discrete optimization problems.

David P. Williamson is a Professor at Cornell University with a joint appointment in the School of Operations Research and Information Engineering and in the Department of Information Science. Prior to joining Cornell, he was a Research Staff Member at the IBM T.J. Watson Research Center and a Senior Manager at the IBM Almaden Research Center. He has won several awards for his work on approximation algorithms, including the 2000 Fulkerson Prize, sponsored by the American Mathematical Society and the Mathematical Programming Society. He has served on several editorial boards, including *ACM Transactions on Algorithms, Mathematics of Operations Research*, the SIAM *Journal on Computing*, and the SIAM *Journal on Discrete Mathematics*.

David B. Shmoys is a Professor at Cornell University with appointments in both the School of Operations Research and Information Engineering and the Department of Computer Science. He is currently Associate Director of the Institute for Computational Sustainability at Cornell. He is a Fellow of the ACM, was an NSF Presidential Young Investigator, and has served on numerous editorial boards, including *Mathematics of Operations Research* (for which he is currently an associate editor), *Operations Research*, *ORSA Journal on Computing, Mathematical Programming*, and both the SIAM *Journal on Computing* and *Journal on Discrete Mathematics*; he also served as editor-in-chief for the latter.

The Design of Approximation Algorithms

David P. Williamson

Cornell University

David B. Shmoys

Cornell University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521195270

First published 2011

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Williamson, David P.
The design of approximation algorithms / David P. Williamson, David B. Shmoys.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-521-19527-0 (hardback)
1. Approximation theory. 2. Mathematical optimization.
I. Shmoys, David Bernard. II. Title.
QA221.W55 2011
518′.5–dc22 2011001945

ISBN 978-0-521-19527-0 Hardback

Cambridge University Press
978-0-521-19527-0 - The Design of Approximation Algorithms
David P. Williamson and David B. Shmoys
Frontmatter
More information

Contents

Preface

This book is designed to be a textbook for graduate-level courses in approximation algorithms. After some experience teaching minicourses in the area in the mid-1990s, we sat down and wrote out an outline of the book. Then one of us (DPW), who was at the time an IBM Research Staff Member, taught several iterations of the course following the outline we had devised, in Columbia University's Department of Industrial Engineering and Operations Research in Spring 1998, in Cornell University's School of Operations Research and Industrial Engineering in Fall 1998, and at the Massachusetts Institute of Technology's Laboratory for Computer Science in Spring 2000. The lecture notes from these courses were made available, and we got enough positive feedback on them from students and from professors teaching such courses elsewhere that we felt we were on the right track. Since then, there have been many exciting developments in the area, and we have added many of them to the book; we taught additional iterations of the course at Cornell in Fall 2006 and Fall 2009 in order to field test some of the writing of the newer results.

The courses were developed for students who have already had a class, undergraduate or graduate, in algorithms, and who were comfortable with the idea of mathematical proofs about the correctness of algorithms. The book assumes this level of preparation. The book also assumes some basic knowledge of probability theory (for instance, how to compute the expected value of a discrete random variable). Finally, we assume that the reader knows something about NP-completeness, at least enough to know that there might be good reason for wanting fast, approximate solutions to NP-hard discrete optimization problems. At one or two points in the book, we do an NP-completeness reduction to show that it can be hard to find approximate solutions to such problems; we include a short appendix on the problem class NP and the notion of NP-completeness for those unfamiliar with the concepts. However, the reader unfamiliar with such reductions can also safely skip over such proofs.

In addition to serving as a graduate textbook, this book is a way for students to get the background to read current research in the area of approximation algorithms. In particular, we wanted a book that we could hand our own Ph.D. students just starting in the field and say, "Here, read this."

We further hope that the book will serve as a reference to the area of approximation algorithms for researchers who are generally interested in the heuristic solution of discrete optimization problems; such problems appear in areas as diverse as traditional operations research planning problems (such as facility location and network design) to computer science problems in database and programming language design to advertising issues in viral marketing. We hope that the book helps researchers understand the techniques available in the area of approximation algorithms for approaching such problems.

We have taken several particular perspectives in writing the book. The first is that we wanted to organize the material around certain principles of designing approximation algorithms, around algorithmic ideas that have been used in different ways and applied to different optimization problems. The title *The Design of Approximation Algorithms* was carefully chosen. The book is structured around these design techniques. The introduction applies several of them to a single problem, the set cover problem. The book then splits into two parts. In the first part, each chapter is devoted to a single algorithmic idea (e.g., "greedy and local search algorithms," "rounding data and dynamic programming"), and the idea is then applied to several different problems. The second part revisits all of the same algorithmic ideas, but gives more sophisticated treatments of them; the results covered here are usually much more recent. The layout allows us to look at several central optimization problems repeatedly throughout the book, returning to them as a new algorithmic idea leads to a better result than the previous one. In particular, we revisit such problems as the uncapacitated facility location problem, the prize-collecting Steiner tree problem, the bin-packing problem, and the maximum cut problem several times throughout the course of the book.

The second perspective is that we treat linear and integer programming as a central aspect in the design of approximation algorithms. This perspective is from our background in the operations research and mathematical programming communities. It is a little unusual in the computer science community, and students coming from a computer science background may not be familiar with the basic terminology of linear programming. We introduce the terms we need in the first chapter, and we include a brief introduction to the area in an appendix.

The third perspective we took in writing the book is that we have limited ourselves to results that are simple enough for classroom presentation while remaining central to the topic at hand. Most of the results in the book are ones that we have taught ourselves in class at one point or another. We bent this rule somewhat in order to cover the recent, exciting work by Arora, Rao, and Vazirani [22] applying semidefinite programming to the uniform sparsest cut problem. The proof of this result is the most lengthy and complicated of the book.

We are grateful to a number of people who have given us feedback about the book at various stages in its writing. We are particularly grateful to James Davis, Lisa Fleischer, Isaac Fung, Rajiv Gandhi, Igor Gorodezky, Nick Harvey, Anna Karlin, Vijay Kothari, Katherine Lai, Gwen Spencer, and Anke van Zuylen for very detailed comments on a number of sections of the book. Additionally, the following people spotted typos, gave us feedback, helped us understand particular papers, and made useful suggestions: Bruno Abrahao, Hyung-Chan An, Matthew Andrews, Eliot Anshelevich, Sanjeev Arora, Ashwinkumar B.V., Moses Charikar, Chandra Chekuri, Joseph Cheriyan, Chao

Ding, Dmitriy Drusvyatskiy, Michel Goemans, Sudipto Guha, Anupam Gupta, Sanjeev Khanna, Lap Chi Lau, Renato Paes Leme, Jan Karel Lenstra, Jiawei Qian, Roman Rischke, Gennady Samorodnitsky, Daniel Schmand, Yogeshwer Sharma, Viktor Simjanoski, Mohit Singh, Éva Tardos, Mike Todd, Di Wang, and Ann Williamson. We also thank a number of anonymous reviewers who made useful comments. Eliot Anshelevich, Joseph Cheriyan, Lisa Fleischer, Michel Goemans, Nicole Immorlica, and Anna Karlin used various drafts of the book in their courses on approximation algorithms and gave us useful feedback about the experience of using the book. We received quite a number of useful comments from the students in Anna's class: Benjamin Birnbaum, Punyashloka Biswal, Elisa Celis, Jessica Chang, Mathias Hallman, Alyssa Joy Harding, Trinh Huynh, Alex Jaffe, Karthik Mohan, Katherine Moore, Cam Thach Nguyen, Richard Pang, Adrian Sampson, William Austin Webb, and Kevin Zatloukal. Frans Schalekamp generated the image on the cover; it is an illustration of the tree metric algorithm of Fakcharoenphol, Rao, and Talwar [106] discussed in Section 8.5. Our editor at Cambridge, Lauren Cowles, impressed us with her patience in waiting for this book to be completed and gave us a good deal of useful advice.

We would like to thank the institutions that supported us during the writing of this book, including our home institution, Cornell University, and the IBM T.J. Watson and Almaden Research Centers (DPW), as well as TU Berlin (DPW) and the Sloan School of Management at MIT and the Microsoft New England Research Center (DBS), where we were on sabbatical leave when the final editing of the book occurred. We are grateful to the National Science Foundation for supporting our research in approximation algorithms.

We are also grateful to our wives and children – to Ann, Abigail, Daniel, and Ruth, and to Éva, Rebecca, and Amy – for their patience and support during the writing of this volume.

Additional materials related to the book (such as contact information and errata) can be found at the website www.designofapproxalgs.com.

Finally, we hope the book conveys some of our enthusiasm and enjoyment of the area of approximation algorithms. We hope that you, dear reader, will enjoy it too.

David P. Williamson
David B. Shmoys
January 2011