Beyond the Worst-Case Analysis of Algorithms

There are no silver bullets in algorithm design, and no single algorithmic idea is powerful and flexible enough to solve every computational problem. Nor are there silver bullets in algorithm analysis, as the most enlightening method for analyzing an algorithm often depends on the problem and the application. However, typical algorithms courses rely almost entirely on a single analysis framework, that of worst-case analysis, wherein an algorithm is assessed by its worst performance on any input of a given size.

The purpose of this book is to popularize several alternatives to worst-case analysis and their most notable algorithmic applications, from clustering to linear programming to neural network training. Forty leading researchers have contributed introductions to different facets of this field, emphasizing the most important models and results, many of which can be taught in lectures to beginning graduate students in theoretical computer science and machine learning.

Tim Roughgarden is a professor of computer science at Columbia University. For his research, he has been awarded the ACM Grace Murray Hopper Award, the Presidential Early Career Award for Scientists and Engineers (PECASE), the Kalai Prize in Computer Science and Game Theory, the Social Choice and Welfare Prize, the Mathematical Programming Society’s Tucker Prize, and the EATCS-SIGACT Gödel Prize. He was an invited speaker at the 2006 International Congress of Mathematicians, the Shapley Lecturer at the 2008 World Congress of the Game Theory Society, and a Guggenheim Fellow in 2017. His other books include Twenty Lectures on Algorithmic Game Theory (2016) and the Algorithms Illuminated book series (2017–2020).
Beyond the Worst-Case Analysis of Algorithms

Edited by

Tim Roughgarden

Columbia University, New York
# Contents

| Preface | page xiii |
| List of Contributors | xv |

## 1 Introduction

*Tim Roughgarden*

1.1 The Worst-Case Analysis of Algorithms | 1
1.2 Famous Failures and the Need for Alternatives | 3
1.3 Example: Parameterized Bounds in Online Paging | 8
1.4 Overview of the Book | 12
1.5 Notes | 20

## PART ONE REFINEMENTS OF WORST-CASE ANALYSIS

### 2 Parameterized Algorithms

*Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi*

2.1 Introduction | 27
2.2 Randomization | 31
2.3 Structural Parameterizations | 34
2.4 Kernelization | 35
2.5 Hardness and Optimality | 39
2.6 Outlook: New Paradigms and Application Domains | 42
2.7 The Big Picture | 46
2.8 Notes | 47

### 3 From Adaptive Analysis to Instance Optimality

*Jérémy Barbay*

3.1 Case Study 1: Maxima Sets | 52
3.2 Case Study 2: Instance-Optimal Aggregation Algorithms | 60
3.3 Survey of Additional Results and Techniques | 64
3.4 Discussion | 65
3.5 Selected Open Problems | 66
3.6 Key Takeaways | 67
3.7 Notes | 68
CONTENTS

4 Resource Augmentation 72
Tim Roughgarden
4.1 Online Paging Revisited 72
4.2 Discussion 75
4.3 Selfish Routing 77
4.4 Speed Scaling in Scheduling 81
4.5 Loosely Competitive Algorithms 86
4.6 Notes 89

PART TWO  DETERMINISTIC MODELS OF DATA

5 Perturbation Resilience 95
Konstantin Makarychev and Yury Makarychev
5.1 Introduction 95
5.2 Combinatorial Optimization Problems 98
5.3 Designing Certified Algorithms 101
5.4 Examples of Certified Algorithms 106
5.5 Perturbation-Resilient Clustering Problems 108
5.6 Algorithm for 2-Perturbation-Resilient Instances 111
5.7 $(3 + \varepsilon)$-Certified Local Search Algorithm for $k$-Medians 113
5.8 Notes 115

6 Approximation Stability and Proxy Objectives 120
Avrim Blum
6.1 Introduction and Motivation 120
6.2 Definitions and Discussion 121
6.3 The $k$-Median Problem 125
6.4 $k$-Means, Min-Sum, and Other Clustering Objectives 132
6.5 Clustering Applications 133
6.6 Nash Equilibria 134
6.7 The Big Picture 135
6.8 Open Questions 136
6.9 Relaxations 137
6.10 Notes 137

7 Sparse Recovery 140
Eric Price
7.1 Sparse Recovery 140
7.2 A Simple Insertion-Only Streaming Algorithm 142
7.3 Handling Deletions: Linear Sketching Algorithms 143
7.4 Uniform Algorithms 148
7.5 Lower Bound 154
7.6 Different Measurement Models 155
7.7 Matrix Recovery 158
7.8 Notes 160
## CONTENTS

### PART THREE  SEMIRANDOM MODELS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Distributional Analysis</td>
<td>167</td>
</tr>
<tr>
<td>9</td>
<td>Introduction to Semirandom Models</td>
<td>189</td>
</tr>
<tr>
<td>10</td>
<td>Semirandom Stochastic Block Models</td>
<td>212</td>
</tr>
<tr>
<td>11</td>
<td>Random-Order Models</td>
<td>234</td>
</tr>
<tr>
<td>12</td>
<td>Self-Improving Algorithms</td>
<td>259</td>
</tr>
</tbody>
</table>

8 Distributional Analysis  
Tim Roughgarden  
8.1 Introduction  
8.2 Average-Case Justifications of Classical Algorithms  
8.3 Good-on-Average Algorithms for Euclidean Problems  
8.4 Random Graphs and Planted Models  
8.5 Robust Distributional Analysis  
8.6 Notes

9 Introduction to Semirandom Models  
Uriel Feige  
9.1 Introduction  
9.2 Why Study Semirandom Models?  
9.3 Some Representative Work  
9.4 Open Problems

10 Semirandom Stochastic Block Models  
Ankur Moitra  
10.1 Introduction  
10.2 Recovery via Semidefinite Programming  
10.3 Robustness Against a Monotone Adversary  
10.4 Information Theoretic Limits of Exact Recovery  
10.5 Partial Recovery and Belief Propagation  
10.6 Random versus Semirandom Separations  
10.7 Above Average-Case Analysis  
10.8 Semirandom Mixture Models

11 Random-Order Models  
Anupam Gupta and Sahil Singla  
11.1 Motivation: Picking a Large Element  
11.2 The Secretary Problem  
11.3 Multiple-Secretary and Other Maximization Problems  
11.4 Minimization Problems  
11.5 Related Models and Extensions  
11.6 Notes

12 Self-Improving Algorithms  
C. Seshadhri  
12.1 Introduction  
12.2 Information Theory Basics  
12.3 The Self-Improving Sorter  
12.4 Self-Improving Algorithms for 2D Maxima  
12.5 More Self-Improving Algorithms  
12.6 Critique of the Self-Improving Model
CONTENTS

PART FOUR SMOOTHED ANALYSIS

13 Smoothed Analysis of Local Search 285
   Bodo Manthey
   13.1 Introduction 285
   13.2 Smoothed Analysis of the Running Time 286
   13.3 Smoothed Analysis of the Approximation Ratio 301
   13.4 Discussion and Open Problems 304
   13.5 Notes 305

14 Smoothed Analysis of the Simplex Method 309
   Daniel Dadush and Sophie Huiberts
   14.1 Introduction 309
   14.2 The Shadow Vertex Simplex Method 310
   14.3 The Successive Shortest Path Algorithm 315
   14.4 LPs with Gaussian Constraints 319
   14.5 Discussion 329
   14.6 Notes 330

15 Smoothed Analysis of Pareto Curves in Multiobjective Optimization 334
   Heiko Röglin
   15.1 Algorithms for Computing Pareto Curves 334
   15.2 Number of Pareto-optimal Solutions 342
   15.3 Smoothed Complexity of Binary Optimization Problems 352
   15.4 Conclusions 354
   15.5 Notes 355

PART FIVE APPLICATIONS IN MACHINE LEARNING AND STATISTICS

16 Noise in Classification 361
   Maria-Florina Balcan and Nika Haghtalab
   16.1 Introduction 361
   16.2 Model 362
   16.3 The Best Case and the Worst Case 363
   16.4 Benefits of Assumptions on the Marginal Distribution 365
   16.5 Benefits of Assumptions on the Noise 374
   16.6 Final Remarks and Current Research Directions 378

17 Robust High-Dimensional Statistics 382
   Ilias Diakonikolas and Daniel M. Kane
   17.1 Introduction 382
   17.2 Robust Mean Estimation 384
   17.3 Beyond Robust Mean Estimation 396
   17.4 Notes 399
## CONTENTS

18 Nearest Neighbor Classification and Search  
*Sanjoy Dasgupta and Samory Kpotufe*

18.1 Introduction  
18.2 The Algorithmic Problem of Nearest Neighbor Search  
18.3 Statistical Complexity of $k$-Nearest Neighbor Classification  
18.4 Notes

19 Efficient Tensor Decompositions  
*Aravindan Vijayaraghavan*

19.1 Introduction to Tensors  
19.2 Applications to Learning Latent Variable Models  
19.3 Efficient Algorithms in the Full-Rank Setting  
19.4 Smoothed Analysis and the Overcomplete Setting  
19.5 Other Algorithms for Tensor Decompositions  
19.6 Discussion and Open Questions

20 Topic Models and Nonnegative Matrix Factorization  
*Rong Ge and Ankur Moitra*

20.1 Introduction  
20.2 Nonnegative Matrix Factorization  
20.3 Topic Models  
20.4 Epilogue: Word Embeddings and Beyond

21 Why Do Local Methods Solve Nonconvex Problems?  
*Tengyu Ma*

21.1 Introduction  
21.2 Analysis Technique: Characterization of the Landscape  
21.3 Generalized Linear Models  
21.4 Matrix Factorization Problems  
21.5 Landscape of Tensor Decomposition  
21.6 Survey and Outlook: Optimization of Neural Networks  
21.7 Notes

22 Generalization in Overparameterized Models  
*Moritz Hardt*

22.1 Background and Motivation  
22.2 Tools to Reason About Generalization  
22.3 Overparameterization: Empirical Phenomena  
22.4 Generalization Bounds for Overparameterized Models  
22.5 Empirical Checks and Holdout Estimates  
22.6 Looking Ahead  
22.7 Notes
# CONTENTS

## 23 Instance Optimal Distribution Testing and Learning

*Gregory Valiant and Paul Valiant*

- 23.1 Testing and Learning Discrete Distributions 506
- 23.2 Instance Optimal Distribution Learning 507
- 23.3 Identity Testing 516
- 23.4 Digression: An Automatic Inequality Prover 519
- 23.5 Beyond Worst-Case Analysis for Other Testing Problems 522
- 23.6 Notes 523

## PART SIX FURTHER APPLICATIONS

## 24 Beyond Competitive Analysis

*Anna R. Karlin and Elias Koutsoupias*

- 24.1 Introduction 529
- 24.2 The Access Graph Model 530
- 24.3 The Diffuse Adversary Model 534
- 24.4 Stochastic Models 537
- 24.5 Direct Comparison of Online Algorithms 540
- 24.6 Where Do We Go from Here? 541
- 24.7 Notes 542

## 25 On the Unreasonable Effectiveness of SAT Solvers

*Vijay Ganesh and Moshe Y. Vardi*

- 25.1 Introduction: The Boolean SAT Problem and Solvers 547
- 25.2 Conflict-Driven Clause Learning SAT Solvers 550
- 25.3 Proof Complexity of SAT Solvers 554
- 25.4 Proof Search, Automatizability, and CDCL SAT Solvers 557
- 25.5 Parameteric Understanding of Boolean Formulas 558
- 25.6 Proof Complexity, Machine Learning, and Solver Design 562
- 25.7 Conclusions and Future Directions 563

## 26 When Simple Hash Functions Suffice

*Kai-Min Chung, Michael Mitzenmacher, and Salil Vadhan*

- 26.1 Introduction 567
- 26.2 Preliminaries 571
- 26.3 Hashing Block Sources 575
- 26.4 Application: Chained Hashing 576
- 26.5 Optimizing Block Source Extraction 577
- 26.6 Application: Linear Probing 578
- 26.7 Other Applications 580
- 26.8 Notes 581

## 27 Prior-Independent Auctions

*Inbal Talgam-Cohen*

- 27.1 Introduction 586
- 27.2 A Crash Course in Revenue-Maximizing Auctions 587
CONTENTS

27.3 Defining Prior-Independence 591
27.4 Sample-Based Approach: Single Item 593
27.5 Competition-Based Approach: Multiple Items 598
27.6 Summary 602
27.7 Notes 603

28 Distribution-Free Models of Social Networks 606
Tim Roughgarden and C. Seshadhri
28.1 Introduction 606
28.2 Cliques of $c$-Closed Graphs 607
28.3 The Structure of Triangle-Dense Graphs 612
28.4 Power-Law Bounded Networks 615
28.5 The BCT Model 619
28.6 Discussion 621
28.7 Notes 623

29 Data-Driven Algorithm Design 626
Maria-Florina Balcan
29.1 Motivation and Context 626
29.2 Data-Driven Algorithm Design via Statistical Learning 628
29.3 Data-Driven Algorithm Design via Online Learning 639
29.4 Summary and Discussion 644

30 Algorithms with Predictions 646
Michael Mitzenmacher and Sergei Vassilvitskii
30.1 Introduction 646
30.2 Counting Sketches 649
30.3 Learned Bloom Filters 650
30.4 Caching with Predictions 652
30.5 Scheduling with Predictions 655
30.6 Notes 660

Index 663
Preface

There are no silver bullets in algorithm design – no one algorithmic idea is powerful and flexible enough to solve every computational problem of interest. The emphasis of an undergraduate algorithms course is accordingly on the next-best thing: a small number of general algorithm design paradigms (such as dynamic programming, divide-and-conquer, and greedy algorithms), each applicable to a range of problems that span multiple application domains.

Nor are there silver bullets in algorithm analysis, as the most enlightening method for analyzing an algorithm often depends on the details of the problem and motivating application. However, the focus of a typical algorithms course rests almost entirely on a single analysis framework, that of worst-case analysis, wherein an algorithm is assessed by its worst performance on any input of a given size. The goal of this book is to redress the imbalance and popularize several alternatives to worst-case analysis, developed largely in the theoretical computer science literature over the past 20 years, and their most notable algorithmic applications. Forty leading researchers have contributed introductions to different facets of this field, and the introductory Chapter 1 includes a chapter-by-chapter summary of the book’s contents.

This book’s roots lie in a graduate course that I developed and taught several times at Stanford University. While the project has expanded in scope far beyond what can be taught in a one-term (or even one-year) course, subsets of the book can form the basis of a wide variety of graduate courses. Authors were requested to avoid comprehensive surveys and focus instead on a small number of key models and results that could be taught in lectures to second-year graduate students in theoretical computer science and theoretical machine learning. Most of the chapters conclude with open research directions as well as exercises suitable for classroom use. A free electronic copy of this book is available from the URL https://www.cambridge.org/9781108494311#resources (with the password ‘BWCA_CUP’).

Producing a collection of this size is impossible without the hard work of many people. First and foremost, I thank the authors for their dedication and timeliness in writing their own chapters and for providing feedback on preliminary drafts of other chapters. I thank Avrim Blum, Moses Charikar, Lauren Cowles, Anupam Gupta,

\footnote{Lecture notes and videos from this course, covering several of the topics in this book, are available from my home page (www.timroughgarden.org).}
PREFACE

Ankur Moitra, and Greg Valiant for their enthusiasm and excellent advice when this project was in its embryonic stages. I am also grateful to all the Stanford students who took my CS264 and CS369N courses, and especially to my teaching assistants Rishi Gupta, Joshua Wang, and Qiqi Yan. The cover art is by Max Greenleaf Miller. The editing of this book was supported in part by NSF award CCF-1813188 and ARO award W911NF1910294.
Contributors

Maria-Florina Balcan
Carnegie Mellon University

Jérémy Barbay
University of Chile

Avrim Blum
Toyota Technological Institute at Chicago

Kai-Min Chung
Institute of Information Science, Academia Sinica

Daniel Dadush
Centrum Wiskunde Informatica

Sanjoy Dasgupta
University of California at San Diego

Ilias Diakonikolas
University of Wisconsin-Madison

Uriel Feige
The Weizman Institute

Fedor Fomin
University of Bergen

Vijay Ganesh
University of Waterloo

Rong Ge
Duke University
LIST OF CONTRIBUTORS

Anupam Gupta
Carnegie Mellon University

Nika Haghtalab
Cornell University

Moritz Hardt
University of California at Berkeley

Sophie Huiberts
Centrum Wiskunde Informatica

Daniel Kane
University of California at San Diego

Anna R. Karlin
University of Washington at Seattle

Elias Koutsoupias
University of Oxford

Samory Kpotufe
Columbia University

Daniel Lokshtanov
University of California at Santa Barbara

Tengyu Ma
Stanford University

Konstantin Makarychev
Northwestern University

Yury Makarychev
Toyota Technological Institute at Chicago

Bodo Manthey
University of Twente

Michael Mitzenmacher
Harvard University

Ankur Moitra
Massachusetts Institute of Technology

xvi
LIST OF CONTRIBUTORS

Eric Price
The University of Texas at Austin

Heiko Röglin
University of Bonn

Tim Roughgarden
Columbia University

Saket Saurabh
Institute of Mathematical Sciences

C. Seshadhri
University of California at Santa Cruz

Sahil Singla
Princeton University

Inbal Talgam-Cohen
Technion–Israel Institute of Technology

Salil Vadhan
Harvard University

Gregory Valiant
Stanford University

Paul Valiant
Brown University

Moshe Vardi
Rice University

Sergei Vassilvitskii
Google, Inc.

Aravindan Vijayaraghavan
Northwestern University

Meirav Zehavi
Ben-Gurion University of the Negev