Contents

Preface

Computational complexity theory has been a central area of theoretical computer science since its early development in the mid-1960s. Its subsequent rapid development in the next three decades has not only established itself as a rich, exciting theory but also shown strong influence on many other related areas in computer science, mathematics, and operations research. We may roughly divide its development into three periods. The works from the mid-1960s to the early 1970s paved a solid foundation for the theory of feasible computation. The Turing machine-based complexity theory and the axiomatic complexity theory established computational complexity as an independent mathematics discipline. The identification of polynomial-time computability with feasible computability and the discovery of the *NP*-complete problems consolidated the *P* versus *NP* question as the central issue of the complexity theory.

From the early 1970s to the mid-1980s, research in computational complexity expanded at an exponential rate both in depth and in breadth. Various new computational models were proposed as alternatives to the traditional deterministic models. Finer complexity hierarchies and complexity classes were identified from these new models and more accurate classifications have been obtained for the complexity of practical algorithmic problems. Parallel computational models, such as alternating Turing machines and parallel random access machines, together with the *NC* hierarchy, provide a tool for the classification of the complexity of feasible problems. Probabilistic Turing machines are a model for the complexity theory of distribution-independent randomized algorithms. Interactive proof systems, an extension of probabilistic Turing machines, and communication complexity study the complexity aspect of distributed or interactive computing. The study of one-way functions led to a breakthrough in cryptography. A theory of average-case completeness, based on the notion of distribution-faithful reductions, aims at establishing the foundation for the distribution-dependent average-case complexity. Boolean and threshold circuits are models for nonuniform complexity in which algebraic and topological methods have found interesting applications. Oracle Turing machines are the model for the theory of relativization in which combinatorial techniques meet recursion-theoretic techniques. Program-size complexity (or, Kolmogorov complexity) formalizes the notion of descriptive complexity and has strong connections with computational complexity. Although the central questions remain open, these developments demonstrate that computational complexity is a rich discipline with both a deep mathematical theory and a diverse area of applications.

Beginning in the mid-1980s, we have seen a number of deep, surprising results using diverse, sophisticated proof techniques. In addition, seemingly independent subareas have found interesting connections. The exponential lower bounds for monotone circuits and constant-depth circuits have been found using probabilistic and algebraic methods. The connection between constant-depth circuits and relativization led to the relativized separation of the polynomial-time hierarchy. The technique of nondeterministic iterative counting has been used to collapse the nondeterministic space hierarchy. The study of probabilistic reductions gave us the surprising result about the power of the counting class $\#P$ versus the polynomial-time hierarchy. Arithmetization of Boolean computation in interactive proof systems collapses the class of polynomial space to the class of sets with interactive proof systems. Further development of this research, together with techniques of coding theory, have led to strong negative results in combinatorial approximation.

As outlined above, complexity theory has grown fast both in breadth and in depth. With so many new computational models, new proof techniques and applications in different areas, it is simply not possible to cover all important topics of the theory in a single book. The goal of this book is therefore not to provide a comprehensive treatment of complexity theory. Instead, we only select some of the fundamental areas which we believe represent the most important recent advances in complexity theory, in particular, on the P versus NP problem, and present the complete treatment of these subjects. The presentation follows the approach of traditional mathematics textbooks. With a small number of exceptions, all theorems are presented with rigorous mathematical proofs.

We divide the subjects of this book into three parts, each representing a different approach to computational complexity. In Part I, we develop the theory of uniform computational complexity, which is based on the worst-case complexity measures on traditional models of Turing machines, both deterministic ones and nondeterministic ones. The central issue here is the P versus NP question, and we apply the notions of reducibility and completeness to develop a complexity hierarchy. We first develop in Chapter 1 the notion of time and space complexity and complexity classes. Two basic proof techniques, simulation and diagonalization, including Immerman and Szelepcsényi's iterative counting technique, are presented. The knowledge of recursion theory is useful here but is not required. Chapter 2 presents the notion of *NP*-completeness, including Cook's Theorem and a few well-known *NP*-complete problems. The relations between decision problems versus search problems are carefully discussed through the notion of polynomial-time Turing reducibility. Chapter 3 extends the theory of *NP*-completeness to the polynomial-time hierarchy and polynomial space. In addition to complete problems for these complexity classes, we also present the natural characterizations of these complexity classes by alternating Turing machines and alternating quantifiers. In Chapter 4, the structure of the class *NP* is analyzed in several different views. We present both the abstract proof that there exist problems in *NP* that are neither *NP*-complete nor in *P*, assuming *NP* does not collapse to *P*, and some natural problems as the candidates of such problems, as well as their applications in public-key cryptography. The controversial theory of relativization and their interpretations are also introduced and discussed in this chapter.

In Part II, we study the theory of nonuniform computational complexity, including the computational models of decision trees and Boolean circuits, and the notion of sparse sets. The nonuniform computational models grew out of our inability to solve the major open questions in the uniform complexity theory. It is hoped that the simpler structure of these nonuniform models will allow better lower bound results. Although the efforts so far are not strong enough to settle the major open questions in the area of uniform complexity theory, a number of nontrivial lower bound results have been obtained through new proof techniques. The emphasis of this part is thus not on the subjects themselves but on the proof techniques. In Chapter 5, we present both the algebraic and the topological techniques to prove the lower bounds for decision trees of Boolean functions, particularly Boolean functions representing monotone graph properties. In Chapter 6, we present two exponential lower bound results on circuits using the approximation circuit technique and the probabilistic method. The notion of sparse sets links the study of nonuniform complexity with uniform complexity theory. This interesting interconnection between uniform and nonuniform complexity theory, such as the question of *NC* versus *P*, is also studied in Chapter 6. Then, we present, in Chapter 7, the works on the Hartmanis-Berman conjecture about the polynomial-time isomorphism of *NP*-complete problems, which provide further insight into the structure of the complexity class *NP*.

Part III is about the theory of probabilistic complexity, which studies the complexity issues related to randomized computation. Randomization in algorithms started in the late 1970s, and has become increasingly popular. The computational model for randomized algorithms, the probabilistic Turing machine, and the corresponding probabilistic complexity classes are introduced in Chapter 8. The notion of probabilistic quantifiers is used to provide characterizations of these complexity classes, and their relations with deterministic and nondeterministic complexity classes are discussed. The counting problems and the complexity class $\#P$ may be viewed as an extension of probablistic computation, and they are the main subjects of Chapter 9. Valiant's proof of $\#P$-

completeness of the permanent problem, as well as Toda's theorem that all problems in the polynomial-time hierarchy are reducible to problems in $\#P$, are presented. The exponential lower bound of constant-depth circuits developed in Chapter 6 has an interesting application to the relativized separation of the complexity class $\#P$ from the polynomial-time hierarchy. This result is also presented in Chapter 9. Chapter 10 studies the notion of interactive proof systems, which is another extension of probabilistic computation. The collapse of the interactive proof systems hierarchy is presented, and the relations between the interactive proof systems and the Arthur-Merlin proof systems are discussed. Shamir's characterization of polynomial space by interactive proof systems is also presented as a prelude to the recent breakthrough on probabilistically checkable proofs. This celebrated result of Arora et al., that probabilistically checkable proofs with a constant number of queries characterize precisely the class *NP*, is presented in Chapter 11. We also present, in this chapter, the application of this result to various combinatorial approximation problems.

Although we have tried hard to include, within this framework, as many subjects in complexity theory as possible, many interesting topics inevitably have to be omitted. Two of the most important topics that we are not able to include here are program-size complexity (or, Kolmogorov complexity) and average-case completeness. Program-size complexity is a central theory which would provide a unified view of the other nonuniform models of Part II. However, this topic has grown into an independent discipline in recent years and has become too big to be included here. Interested readers are referred to the comprehensive book of Li and Vitányi [1997]. Average-case completeness provides a different view toward the notion of distribution-independent average-case complexity, and would complement the works studied in Part III about distribution-independent probabilistic complexity. This theory, however, seems to be still in the early development stage. Much research is needed before we can better understand its proof techniques and its relation to the worst-case complexity theory, and we reluctantly omit it here. We refer interested readers to Wang [1997] for a thorough review of this topic. Exercises at the end of each chapter often include additional topics that are worth studying but are omitted in the main text due to space limitations.

This book is grown out of authors' lecture notes developed in the past ten years at the University of Minnesota and the State University of New York at Stony Brook. We have taught from these notes in several different ways. For a one-semester graduate course in which the students have had limited exposure to theory, we typically cover most of Part I plus a couple of chapters from either Part II or Part III. For better prepared students, a course emphasizing the recent advances can be taught based mainly on either Part II or Part III. Seminars based on some advanced materials in Parts II and III, plus recent journal papers, have also been conducted for Ph.D. students.

We are grateful to all our colleagues and students who have made precious suggestions, corrections, and criticism on the earlier drafts of this book. We are also grateful to the following institutions for their financial support in preparing this book: the National Science Foundation of the United States, National Science Council of Taiwan, National Natural Science Foundation of China, National 973 Fundamental Research Program of China, City University of Hong Kong, and National Chiao-Tung University of Taiwan.

DING-ZHU DU
KER-I KO