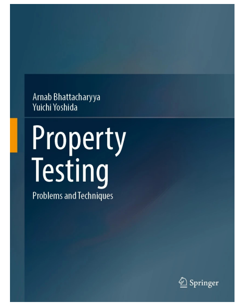
Review of[1]



Property Testing: Problems and Techniques

Arnab Bhattacharyya and Yuichi Yoshida
Published by Springer
Hardcover, 427 pages, $99.99

Review by
Sarvagya Upadhyay (supadhyay@fujitsu.com)
Fujitsu Research of America
350 Cobalt, Sunnyvale CA 94085, USA

1 Overview

Property testing concerns with testing whether an input object satisfies a certain property or not. Since it is a decision problem, the collection of potential input instances are partitioned into two sets: one set that satisfies the property, and the other set that is far from satisfying the property under a suitable distance measure. The set of instances which are not in these two cases is non-empty; however, they are promised not to be an input to the testing algorithm.

The primary goal of property testing is to design highly efficient testers for a given property. Such testers will hopefully read a part of the input and decide whether the property is satisfied or not. Since the algorithm does not read the whole input, it will err on certain cases. A tester reading the entire input will easily determine whether the property is satisfied or not. So the non-trivial aspect of property testing is to design either a sub-linear query algorithm for testing or demonstrate a non-trivial lower bound on the query complexity.

Since its inception and connections with several fields in theoretical computer science, the field has grown tremendously. This book is an attempt to survey the current state-of-the-art in this field. While there are topics that have not been covered in this book, I believe that the book has lot to offer for readers interested in property testing.

2 Summary of Contents

The book is divided into three parts. The first part contains two chapters that form the motivation and basic technical tools required to go through the book. The second part contains six chapters focusing on basic results in property testing. The final part contains five chapters that are geared towards advanced topics and results in property testing. A summary of each chapter in this book is given below. The few mathematical notations employed below are directly from the book.

Chapter 1: Introduction This chapter gives an introduction to property testing. It starts with a problem of monotonicity testing and discusses how we can test this property. While giving the testing algorithm, the authors highlight the naïve approaches and give counterexamples to why those approaches are not the best possible way to test the property. This is followed by a brief section motivating property testing. The rest of the chapter focuses on introducing property testing formally, the different aspects of property testing in view of its formal definition, variations of property testing, and formal proof of the monotonicity tester introduced in the first section. Finally, the chapter finishes with connection to different sub-fields of theoretical computer science and an organization of the book.

Chapter 2: Basic Techniques The second chapter is devoted to tools and techniques that are employed frequently in the book. It starts with illustrating analysis techniques by introducing few simple query problems. The second section of this chapter introduces gap-preserving local reductions between two properties, an important construct which is useful in proving lower bounds via reductions. The final section introduces Yao's minimax principle in the context of property testing. The authors show the power of minimax principle by demonstrating lower bounds for several property testing problems.

Chapter 3: Strings This chapter studies properties on strings. It covers four topics: palindromes, a slightly more involved variant of palindromes termed as two palindromes, the Dyck language, and monotonicity and permutation-freeness. Unsurprisingly, the tester for palindromes is very simple and the analysis not too hard. The authors' illustration of key points of the proof will be very illuminating to beginners in the field. The case of two palindromes is understandably highly non-trivial, and the authors provide a testing algorithm with almost optimal query complexity. They do this by showing an almost matching lower bound as well. The section on Dyck language is somewhat difficult to follow and requires a lot more attentive reading. The final section shows a non-adaptive tester for monotonicity of strings.

Chapter 4: Graphs in the Adjacency Matrix Model As the name of the chapter suggests, the focus is on representing graph by its adjacency matrix and analyze different properties of the underlying graph. The chapter starts with a very brief review on the basics of graph theory in the adjacency matrix model. There are several types of graph partitioning and cut problems whose property testing analogues have been studied in the literature. This forms one of the two main sections of the chapter. The other main section discusses the property of subgraph freeness. The two types of subgraphs considered are a square and a triangle. Apart from this, the chapter gives a very brief overview of additional topics involving digraphs and hypergraphs, graph isomorphism testers, and connection between testing algorithms and optimization problems in graph theory.

Chapter 5: Graphs in the Bounded-Degree Model This chapter focuses on graphs where each vertex has constant number of edges (hence, the name bounded-degree). In other words, the graph is sparse and the adjacency matrix representation is an overkill of resources. The first property testing problem considered in this model is the subgraph freeness. There is another section devoted to subgraph freeness when the subgraph is a cycle. In between these two sections, the chapter focuses on connectivity properties of graphs. All these properties can be tested by constant number of queries. The next two sections focus on testing graph colorability, and demonstrate that the number of queries require is at least $\Omega(\sqrt{n})$ (for 2-colorability) and $\Omega(n)$ (for 3-colorability), where n is the number of vertices in the graph. The final two sections devote to developing constant-time randomized approximation algorithms for certain types of graph optimization problems.

Chapter 6: Functions over Hypercubes This chapter focuses on functions over hypercubes (i.e., functions of the form $f : \{0, 1\}^n \to R \subseteq \mathbb{R}$). Of primary interest is the case when the range is Boolean. Studying such functional properties were the starting point of the field of property testing and have been widely used in inapproximability results, coding theory, and program checking. The chapter starts with testing monotonicity of Boolean functions and a brief review of Fourier analysis. Then the chapter focuses on linearity testing and testing juntas. The lower bounds are considered next by showing that testers for several function properties can be used to design two-party communication protocols, and then leverage known lower bounds in communication complexity to show query lower bounds on testers. The chapter's next section focuses on additional topics in function testing.

Chapter 7: Massively Parameterized Model This chapter focuses on a model where a tester requires a large number of parameters (usually a function of the input size) to test a specified property. The first problem discussed in this framework is testing bipartiteness (or 2-colorability) and how it admits a constant-query algorithm. Then the chapter focuses on monotonicity testing of a labeling function $f : V \to 0, 1$ on a directed acyclic graph. The authors show a sub-linear query upper bound. The focus on the next section is on a type of property testing problem where no sub-linear query algorithm exists to test the property. The final section discusses constraint satisfaction problems (CSPs). The main focus is on sub-linear testing algorithm for 2-SAT and linear lower bound on 3-SAT .

Chapter 8: Vectors and Matrices over the Reals This chapter is somewhat different from the previous chapters in the sense that it deals with vectors and matrices over real numbers. The first topic discussed in this chapter is to test whether a given matrix is low-rank or not. The next section considers the problem of testing whether a given vector belongs to a low-dimensional subspace. The section gives both an upper bound and a lower bound on the query complexity of one-sided testers. The final section devotes to additional topics related to matrices and vectors.

Chapter 9: Graphs in the Adjacency Matrix Model: Characterizations via Regularity Lemma This chapter takes a general view on graph property testing in the adjacency matrix representation. It defines few notions of graph property and characterizes the query complexity based on the notion. For instance, the first theorem of this chapter states that every monotone property of a graph (if the property holds for the graph, then it holds for every subgraph too) is testable with constant queries. The next section defines the notion of hereditary property and shows that hereditary graph properties are also testable with constant number of queries. The next section introduces the notion of oblivious tester and provides a characterization of a property being constant-query testable by an oblivious tester. The final section provides another characterization of existence of constant-query oblivious tester for a graph property.

Chapter 10: Graphs in the Bounded-Degree Model: General Testability Results via Matroid Theory and Graph Minor Theory In this chapter, the emphasis is on graph properties in the bounded-degree models. The main goal of this chapter is to derive conditions under which a property is constant-query testable. The chapter studies two classes of properties: (i) properties that are closed under edge addition, and (ii) properties that are closed under edge-removal. In the former case, the authors introduce a notion of (k, ℓ)-fullness. This property encompasses several graph properties, and the authors show a constant-query property testing algorithm for (k, ℓ)-fullness. In the latter case, the authors consider minor-closed properties (properties closed under subgraph and edge-contraction) and again show a constant-query property testing algorithm. In both scenarios, the tester has two-sided errors.

Chapter 11: Affine-Invariant Properties of Functions As the title of the chapter suggests, the main focus of this chapter is on properties that are invariant under affine transformation on vector spaces over finite fields ($\mathbb{F}^n$). The chapter defines the notion of affine-invariant subspace-hereditary property which roughly says that the affine-invariant property restricted to any affine subspace remains affine-invariant. The rest of the chapter is devoted to showing that subspace-hereditary property are constant-query testable with one-sided error.

Chapter 12: Linear Properties of Functions This chapter focuses on linear properties $\mathcal{P}$ of functions of the form $f : D \to \mathbb{F}$. The property set $\mathcal{P}$ is linear if for any $f, g \in \mathcal{P}$ and any $\alpha, \beta \in \mathbb{F}$, $\alpha f + \beta g \in \mathcal{P}$. Such properties have a strong connection with locally testable codes (LTCs). LTCs are one of the most important topics in coding and complexity theory simply because they satisfy good error-correcting properties and have been central to the development of most constructions of probabilistically checkable proofs (PCPs). The first section gives an overview of coding theory followed by a section on a generic methodology of testability of linear codes. The next two sections are devoted to low-degree testing over finite fields (whether a polynomial is of low-degree or far from it). The next section deals with testing membership in a tensor product codes, and the following section focuses on testability of lifted codes. The emphasis of final section is on discussing constructions of locally testable codes with good error-correcting properties.

Chapter 13: Massively Parameterized Model: Classification of Boolean CSPs The focus of this chapter is on Boolean CSPs from the perspective of massively parameterized model. The focus is on characterization of Boolean CSPs with respect to the query complexity of their testers. The main tool used is universal algebra and gap-preserving local reductions between CSPs.

3 Evaluation and Opinion

I read the book as a non-expert in this field. My knowledge of the field is rudimentary at best. I didn't read this book in one go and have kept coming back to it after every few weeks. This has allowed me to read it at a leisurely pace and internalize the topics covered.

My main issue with the book is with its presentation style. At various places, it seemed that I am reading a research paper rather than a book. For example, section headings such as "Proof of Lemma ..." are not illuminating. It can be difficult to follow through for someone with a tangential interest in the topic. It would help if the authors gave an overview of dependency between the chapters and a rationale for the organization. Finally, the third chapter makes use of techniques from graph theory, which makes me wonder if it could have been discussed after discussing graph properties.

Property testing as a field has evolved over the years. The analytical techniques used to show the correctness of testers have become quite sophisticated and varied. The authors have done very well to give a brief overview of the techniques (wherever required). Personally, the section focusing on higher-order Fourier analysis and Gowers norm was great to read. I would like to mention that mathematical rigor and maturity are very important for reading this book. The definitions and proofs are presented in full rigor, and they can overwhelm a casual reader. There are instances in the book where a series of definitions are presented one by one; in these scenarios, it is important to understand each definition carefully before moving on. There are also topics in the book where as a reader I wanted to know more. Reading Chapter 8 (property testing on real vectors and matrices) left me desiring for more on property testing in linear algebra.

Overall, I will recommend this book as a reference to advanced graduate students. For beginners, I will recommend starting with surveys and other books in property testing before diving into this book. I personally learnt a lot from this book specially in Chapters 7 and 12, and I firmly believe that this book will be a great reference to a researcher who wants to know more on this topic or related topics.