CAS CS 591 B: Communication Complexity

Nadya Voronova Fall 2019

Lecture Notes 22:

Property Testing and Communication Complexity

Talk based on the paper "Property testing lower bounds via communication complexity" by Eric Blais, Joshua Brody, and Kevin Matulef.

The field of property testing seeks to formalize the question: what can we determine about a large object, with limited access to the object itself? In general, the large object may be anything – instance a graph on n nodes, a function on n variables, a string of numbers, an image, etc. In a typical property testing setup, a tester who has unbounded computational power is given query access to the large object. The tester's goal is to accept the object if it has some property $\mathcal{P}$ and reject it if it is "far" from having property $\mathcal{P}$.

The notion of testing boolean functions in this framework goes back to the seminal work of Rubinfeld & Sudan (1996) and has several connections to complexity theory (in particular PCPs and hardness of approximation), as well as computational learning theory (Ron 2008). Over the last two decades, researchers have exerted a considerable amount of effort to determine the query complexity of testing properties of a function f, such as

- whether f is a linear function (Blum et al. 1993)
- whether f is isomorphic to a given function (Alon & Blais 2010; Blais & O'Donnell 2010; Chakraborty et al. 2011b),
- whether f is a k-junta (Blais 2008, 2009; Fischer et al. 2004),
- whether f is a monotone function (Fischer et al. 2002; Goldreich et al. 2000),
- whether f is a dictator (Parnas et al. 2002), etc.

Starting with the ground-breaking work of Goldreich et al. (1998), there has also been much effort directed at determining the query complexity for testing properties of graphs and, more generally, of combinatorial objects. (See, e.g., Goldreich 2010b; Ron 2009.)

Over the course of this effort, a variety of techniques have been developed for designing property testing algorithms, thus proving testing upper bounds. However, as is often the case in theoretical computer science, lower bounds are harder to come by. Although several lower bounds for specific problems are known, few general techniques are known beyond the use of Yao's minimax lemma and the communication complexity technique that we're going to discuss.

Property testing and communication complexity have striking similarities. Both involve parties with unbounded computational power (in one case, the tester, and in the other case, the communicating players), both involve algorithms which are restricted by the parties' limited access to their input and both are interested in the amount of the "information" parties exchange.

We will show that there is indeed a strong connection between testing and communication complexity. More specifically, we will show how to reduce certain communication problems to property testing problems. This reduction method represents an approach to proving testing lower bounds.

1 Property testing model

For functions $f, g : \{0,1\}^n \to R$ where $R \subseteq \text{CC}$ let distance between those functions $\text{dist}(f,g)$ be the fraction of inputs $x \in \{0,1\}^n$ for which $f(x) \neq g(x)$.

Definition 1. A *property* $\mathcal{P}$ of the functions $\{0,1\}^n \to R$ is a subset of those functions. The *distance* between function $f : \{0,1\}^n \to R$ and a property $\mathcal{P}$ is $\min_{g \in P} \text{dist}(f,g)$. When $\text{dist}(f,\mathcal{P}) \geq \epsilon$, we say that f is ϵ-*far* from the property, otherwise it's ϵ-*close*.

Definition 2. An (ϵ, q)-*tester* for a property $\mathcal{P}$ of a function $\{0,1\}^n \to R$ is a randomized algorithm that has query access to the input $f : \{0,1\}^n \to R$, makes at most q queries and

- Accepts f with probability at least $\frac{2}{3}$ if $f \in \mathcal{P}$
- Rejects f with probability at least $\frac{2}{3}$ if f is ϵ-far from $\mathcal{P}$

Definition 3. A tester is called *nonadaptive* if it can select its q queries before observing the value of f on any of those queries. Otherwise, it's called *adaptive*.

Definition 4. A tester that always accepts any function from $\mathcal{P}$ has *1-sided error*. Otherwise, it has *2-sided error*.

We're going to denote the query complexity of the property $\mathcal{P}$ as $Q_\epsilon(\mathcal{P})$ where $Q_\epsilon(\mathcal{P})$ is the minimum value of q for which $\mathcal{P}$ has (ϵ, q)-tester. Similarly, $Q^1_\epsilon(\mathcal{P})$ and $Q^{na}_\epsilon(\mathcal{P})$ denote the minimum number of queries required to ϵ-test the property $\mathcal{P}$ with one-sided error and nonadaptive testers, respectively.

Thought this talk, we will assume that ϵ is a small fixed constant.

2 Main reduction lemma

We define a class of property testing communication games and show how communication lower bounds for these games yield query complexity lower bounds for property testers. These communication games are based on what we call combining operators.

Definition 5. A *combining operator* is an operator ψ that takes as input two functions $f, g : \{0,1\}^n \to Z$ and returns a function $h : \{0,1\}^n \to R$. A combining operator ψ is *simple* if for all f, g, and for all x, the query $h(x)$ can be computed given only x and the queries $f(x)$ and $g(x)$.

For example, when the base functions are boolean, the combining operator defined by $\psi(f,g) = f \oplus g$ is clearly simple – each $h(x) = f(x) \oplus g(x)$ can trivially be computed from $f(x)$ and $g(x)$.

On the other hand, the combining operator ψ that returns the function defined by $h(x) = \bigoplus_{y \in B_x}[f(y) \cdot g(y)]$, where B_x is a Hamming ball centered at x of radius 1, is not simple, since computing $h(x)$ requires knowledge of $f(y)$ and $g(y)$ for several y.

Given a combining operator ψ and a property $\mathcal{P}$, we define $C^{\mathcal{P}}_\psi$ to be the following property testing communication game. Alice receives a function f. Bob receives a function g. They need to compute

$$C^{\mathcal{P}}_\psi = \begin{cases} 1 \text{ if } \psi(f,g) \in \mathcal{P} \\ 0 \text{ if } \psi(f,g) \text{ is } \epsilon\text{-far from } \mathcal{P} \end{cases}$$

We prove all of our testing lower bounds by reducing from an associated communication game $C_\psi^{\mathcal{P}}$. This reduction is simple – Alice and Bob solve $C_\psi^{\mathcal{P}}$ by emulating a testing algorithm for $\mathcal{P}$ on $h = \psi(f, g)$. Note that neither Alice nor Bob have enough information to evaluate a query $h(x)$, because h depends on both f and g. Instead, they must communicate to jointly compute $h(x)$.

Lemma 6 (Main reduction lemma)**.** *Let Z to be a finite set. For any simple combining operator ψ on functions $f, g : \{0,1\}^n \to Z$ and any property over $\{0,1\}^n \to R$, we have*

1. $\mathbf{BPP}(C_\psi^{\mathcal{P}}) \leq 2Q(\mathcal{P}) \cdot \lceil \log |Z| \rceil$
2. $\mathbf{coRP}(C_\psi^{\mathcal{P}}) \leq 2Q^1(\mathcal{P}) \cdot \lceil \log |Z| \rceil$
3. $\mathbf{BPP}^{\rightarrow}(C_\psi^{\mathcal{P}}) \leq Q^{na}(\mathcal{P}) \cdot \lceil \log |Z| \rceil$
4. $\mathbf{coRP}^{\rightarrow}(C_\psi^{\mathcal{P}}) \leq Q^{na,1}(\mathcal{P}) \cdot \lceil \log |Z| \rceil$

where $\mathbf{coRP}(t)$ denotes protocols that compute $t(x, y)$ with certainty if $t(x, y) = 1$ and with probability at least $\frac{2}{3}$ when $t(x, y) = 0$, and $\mathbf{BPP}^{\rightarrow}, \mathbf{coRP}^{\rightarrow}$ denotes one-way communication protocols. All of the protocols have error $\leq \frac{1}{3}$.

Proof. We begin by proving (3). Let A be an optimal q-query nonadaptive tester for $\mathcal{P}$. Let's create a one-way protocol for $C_\psi^{\mathcal{P}}$ using A. Firstly, Alice and Bob generate queries $x_1, \ldots x_q$ using public randomness. Then, Alice computes $f(x_1) \ldots f(x_q)$ and sends them to Bob using $q \cdot \lceil \log |Z| \rceil$ bits. For each i Bob computes $g(x_i)$ and combines it with $f(x_i)$ to compute $h(x_i)$. Finally, Bob emulates A using $h(x_1) \ldots h(x_q)$ and outputs 1 iff A accepts h.

If A has two-sided error then by the correctness of A, protocol computes $\mathcal{C}_\psi^{\mathcal{P}}$ with probability at least $\frac{2}{3}$. Hence, $\mathbf{BPP}^{\rightarrow}(C_\psi^{\mathcal{P}}) \leq q \cdot \lceil \log |Z| \rceil = Q^{na}(\mathcal{P}) \cdot \lceil \log |Z| \rceil$, which proves proposition (3).

If A has one-sided error then whenever $h \in \mathcal{P}$ the protocol correctly outputs 1, and when h is ϵ-far from $\mathcal{P}$ the protocol correctly outputs 0 with probability at least $\frac{2}{3}$. Therefore, $\mathbf{coRP}^{\rightarrow}(C_\psi^{\mathcal{P}}) \leq q \cdot \lceil \log |Z| \rceil = Q^{na,1}(\mathcal{P}) \cdot \lceil \log |Z| \rceil$, as we wanted in (4).

Now, suppose A is an optimal q-query adaptive tester for $\mathcal{P}$. Again, Alice and Bob will use public randomness to generate queries, but this time they will do it one at a time since A is adaptive. Each time x_i is generated, Alice and Bob exchange $f(x_i)$ and $g(x_i)$, which is enough information for Alice and Bob to individually compute $h(x_i)$ because ψ is simple, which gives them enough information to generate the next query with the appropriate distribution (which depends on previous queries). When $h(x_1) \ldots h(x_q)$ have all been computed, Alice and Bob output 1 iff A accepts h. This protocol costs $2q \cdot \lceil \log |Z| \rceil$ bits of communication and thus $\mathbf{coRP}(C_\psi^{\mathcal{P}}) \leq 2q \cdot \lceil \log |Z| \rceil = 2Q^1(\mathcal{P}) \cdot \lceil \log |Z| \rceil$, which proves proposition (1).

Similarly, if A is an optimal adaptive tester with one-sided error then $\mathbf{coRP}(C_\psi^{\mathcal{P}}) \leq 2Q^1(\mathcal{P}) \cdot \lceil \log |Z| \rceil$, which proves proposition (2).

□

3 Lower bounds

Now we can prove some lower bounds for a number of properties that we define below.

Definition 7. The function $f : \{0,1\}^n \to \{0,1\}$ is *linear* if there exists $S \subseteq [n]$ such that for every $x \in \{0,1\}^n$, $f(x) = \bigoplus_{i \in S} x_i$. When $|S| = k$, we say that f is a *k-linear* function.

Definition 8. The function $f : \{0,1\}^n \to \{0,1\}$ is *k-junta* if there is a set $J \subseteq [n]$, $|J| \leq k$ such that for every $x, y \in \{0,1\}^n$ such that $x_j = y_j$ for every $j \in J$ we have $f(x) = f(y)$.

Definition 9. For convenience when discussing Fourier degree we will represent boolean functions using range $\{-1,1\}$ instead of $\{0,1\}$. It is well known that every boolean function $f : \{0,1\}^n \to \{-1,1\}$ has a unique representation of form $f(x) = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S(x)$, where $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$ and $\hat{f}(S) \in \mathbb{R}$. The terms $\hat{f}(S)$ are the Fourier coefficients of f, and the *Fourier degree* of f is the maximum value of $k \geq 0$ such that $\hat{f}(S) \neq 0$ for some S of size k.

We are interested in functions that have Fourier degree at most k.

Definition 10. Every boolean function $f : \{0,1\}^n \to \{0,1\}$ also has a unique representation as a polynomial over GF(2), where $\oplus$ is used as an addition operation, and logical AND is used as a multiplication operation. We say that f is a *k-sparse* polynomial if its representation over GF(2) has at most k terms.

One useful fact is true for all of these properties.

Proposition 11. *Fix $n > 2$ and $1 \leq k \leq n-2$. If $f : \{0,1\}^n \to \{0,1\}$ is $(k+2)$-linear, then f is*

1. *$\frac{1}{2}$-far from k-linear functions,*
2. *$\frac{1}{2}$-far from k-juntas,*
3. *$\frac{1}{2}$-far from functions of Fourier degree at most k,*
4. *$\frac{1}{20}$-far from k-sparse polynomial.*

To prove lower bounds on the communication games associated with these properties, we will reduce to a balanced version of disjointness called k-balanced-DISJ. In this version, Alice receives a set $A \subseteq [n]$ of size $|A| = \lfloor \frac{k}{2} \rfloor + 1$, Bob receives a set $B \subseteq [n]$ of size $|B| = \lceil \frac{k}{2} \rceil + 1$, and there is a promise that $|A \cap B| \leq 1$.

Lemma 12. *For all $0 \leq k \leq n-2$*

$$\mathbf{BPP}(k\text{-balanced-DISJ}_n) = \Omega(\min\{k, n-k\})$$

One can find proofs of proposition 11 and lemma 12 in the original paper.

Now we can prove lower bounds on query the complexity of k-linearity and other properties mentioned above. To do this we're doing to show that if query complexity of these properties are small then we can solve k-balanced-DISJ easier than lemma 12 states because we will provide the protocol for k-balanced-DISJ using an optimal tester for one of properties above.

Theorem 13. *Fix $1 < k < n-1$. Then, at least $\Omega(\min\{k, n-k\})$ queries are required to test, if function $f : \{0,1\}^n \to \{0,1\}$*

1. *is k-linear,*
2. *is k-junta,*
3. *is a function of Fourier degree at most k,*

4. has a k-sparse representation in GF(2).

Proof. By Lemma 6

$$2Q(k\text{-linear}) \geq \mathbf{BPP}(C_{\oplus}^{k\text{-linear}})$$

where $C_{\oplus}^{k\text{-linear}}$ by definition is the communication game where the inputs are the functions $f, g : \{0,1\}^n \to \{0,1\}$ and the players must test whether the function $h = f \oplus g$ is k-linear.

By Lemma 12

$$\mathbf{BPP}(k\text{-balanced-DISJ}) = \Omega(\min\{k, n-k\})$$

To complete the proof of (1), we show that $\mathbf{BPP}(C_{\oplus}^{k\text{-linear}}) \geq \mathbf{BPP}(k\text{-balanced-DISJ})$ with a reduction from k-balanced-DISJ to $C_{\oplus}^{k\text{-linear}}$.

Let $A, B \subseteq [n]$ be the two sets of size $|A| = \lfloor \frac{k}{2} \rfloor + 1$ and $|B| = \lceil \frac{k}{2} \rceil + 1$ received by Alice and Bob as their inputs in k-balanced-DISJ.

Define $\text{Parity}_S = \sum_{i \in S} x_i$. Alice and Bob can construct the functions $\text{Parity}_A, \text{Parity}_B : \{0,1\}^n \to \{0,1\}$. When $|A \cap B| = 1$ the symmetric difference of two input sets has size $|A \Delta B| = |A| + |B| - 2|A \cap B| = k$ and $\text{Parity}_A \oplus \text{Parity}_B = \text{Parity}_{A \Delta B}$ is k-linear. When A and B are disjoint, the function $\text{Parity}_A \oplus \text{Parity}_B$ is $k+2$-linear function and by proposition 11, is $\frac{1}{2}$-far from k-linear functions. So Alice and Bob can solve their instance of k-balanced-DISJ using a communication protocol for $C_{\oplus}^{k\text{-linear}}$. This implies $\mathbf{BPP}(C_{\oplus}^{k\text{-linear}}) \geq \mathbf{BPP}(k\text{-balanced-DISJ})$, as we wanted.

The same reduction works for lower bounds for testing the other properties. Let $C_{\oplus}^{k\text{-junta}}$ denote the communication game where Alice and Bob receive boolean functions f, g and must decide if $h = f \oplus g$ is k-junta. Similarly, $C_{\oplus}^{\text{degree-}k}$ and $C_{\oplus}^{k\text{-sparse}}$ denote communication games where Alice and Bob need to decide if $h = f \oplus g$ has Fourier degree at most k or can be represented by a k-sparse GF(2) polynomial.

By proposition 11, $(k+2)$-linear function is $\frac{1}{2}$-far from k-junta, $\frac{1}{2}$-far from functions with Fourier degree at most k, and $\frac{1}{20}$-far from k-sparse polynomials. And k-linear function is k-junta, has Fourier degree at most k, and can be represented by a k-sparse GF(2) polynomial, so we can conclude that

$$\mathbf{BPP}(k\text{-balanced-DISJ}) \leq \min\{\mathbf{BPP}(C_{\oplus}^{k\text{-junta}}), \mathbf{BPP}(C_{\oplus}^{\text{degree-}k}), \mathbf{BPP}(C_{\oplus}^{k\text{-sparse}})\}$$

Thus,

$$\mathbf{BPP}(C_{\oplus}^{k\text{-linear}}) \geq \mathbf{BPP}(k\text{-balanced-DISJ}) = \Omega(\min\{k, n-k\}),$$

$$\mathbf{BPP}(C_{\oplus}^{k\text{-junta}}) \geq \mathbf{BPP}(k\text{-balanced-DISJ}) = \Omega(\min\{k, n-k\}),$$

$$\mathbf{BPP}(C_{\oplus}^{\text{degree-}k}) \geq \mathbf{BPP}(k\text{-balanced-DISJ}) = \Omega(\min\{k, n-k\}),$$

$$\mathbf{BPP}(C_{\oplus}^{k\text{-sparse}}) \geq \mathbf{BPP}(k\text{-balanced-DISJ}) = \Omega(\min\{k, n-k\}).$$

□