Lecture 8

Lecturer: Sofya Raskhodnikova *Scribe(s): Meiram Murzabulatov*

1 Two-Distribution Version of Yao's Principle

In the previous lecture, we saw that known lower bounds in communication complexity can be used to prove lower bounds for property testing algorithms. Today we will see another method of proving lower bounds for randomized algorithms. This method is summarized in Theorem 1 and it is based on Yao's Principle [Yao77]. We will apply this method to show a lower bound for any nonadaptive algorithm that tests a *nontrivial* graph property (see Definition 2) in the bounded-degree model.

1.1 Proving Lower Bounds for a Randomized Algorithm

Let $a_1...a_q(x)$ denote the answers to the queries of an algorithm for a given property on input x. For any distribution $\mathcal{D}$, define $\mathcal{D}$-view to be the distribution of $a_1...a_q(x)$ when x is selected according to $\mathcal{D}$. Statistical difference between two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ over Ω is

$$SD(\mathcal{D}_1, \mathcal{D}_2) = \max_{S \subseteq \Omega} \left(\left| \Pr_{x \leftarrow \mathcal{D}_1}[x \in S] - \Pr_{x \leftarrow \mathcal{D}_2}[x \in S] \right| \right).$$

Theorem 1 (Yao's Principle)**.** *To prove a lower bound q on the query complexity of a randomized algorithm for a given property, it is enough to give two distributions on inputs:*

- *$\mathcal{P}$ on positive instances, and*
- *$\mathcal{N}$ on negative instances*

such that it is hard for any q-query deterministic algorithm to distinguish $\mathcal{P}$ from $\mathcal{N}$, i.e.,

$$SD(\mathcal{P}\text{-view}, \mathcal{N}\text{-view}) < 1/3.$$

Proof. (Directly taken from [RS06].) Let $\mathcal{A}$ be any (adaptive) deterministic q-query tester. Given distributions $\mathcal{P}, \mathcal{N}$ with $SD(\mathcal{P}\text{-view}, \mathcal{N}\text{-view}) < \frac{1}{3}$ for all such testers $\mathcal{A}$, we define a distribution $\mathcal{D}$, as required in the mainstream version of Yao's Principle. Namely, to get a sample from distribution $\mathcal{D}$, with probability $1/2$ we draw a sample from $\mathcal{P}$ and with probability $1/2$ we draw a sample from $\mathcal{N}$.

Let S be the set of strings $a_1...a_q$ on which $\mathcal{A}$ accepts.

$$\begin{aligned} \left| \Pr_{x \leftarrow \mathcal{P}}[\mathcal{A}(x) = 1] - \Pr_{x \leftarrow \mathcal{N}}[\mathcal{A}(x) = 1] \right| &= \left| \Pr_{a \leftarrow \mathcal{P}\text{-view}}[a \in S] - \Pr_{a \leftarrow \mathcal{N}\text{-view}}[a \in S] \right| \\ &\leq SD(\mathcal{P}\text{-view}, \mathcal{N}\text{-view}) < \frac{1}{3}. \end{aligned}$$

Now we calculate the probability that algorithm $\mathcal{A}$ is correct on inputs distributed according to $\mathcal{D}$:

$$\begin{aligned} \Pr_{x \leftarrow \mathcal{D}}[\mathcal{A}(x) \text{ is correct}] &= \frac{1}{2} \Pr_{x \leftarrow \mathcal{P}}[\mathcal{A}(x) \text{ accepts}] + \frac{1}{2} \Pr_{x \leftarrow \mathcal{N}}[\mathcal{A}(x) \text{ rejects}] \\ &= \frac{1}{2} + \frac{1}{2} \left(\Pr_{x \leftarrow \mathcal{P}}[\mathcal{A}(x) \text{ accepts}] - \Pr_{x \leftarrow \mathcal{N}}[\mathcal{A}(x) \text{ accepts}] \right) \\ &< \frac{1}{2} + \frac{1}{2} = \frac{2}{3}. \end{aligned}$$

That is, every deterministic q-query algorithm is correct on distribution $\mathcal{D}$ with probability less than $\frac{2}{3}$. In other words, there is no deterministic q-query algorithm that works well on distribution $\mathcal{D}$, so we can conclude, by the original formulation of Yao's Principle that there is no q-query probabilistic algorithm for the problem. □

1.2 Limitation of Nonadaptive Algorithms in the Bounded Degree Model

We use Theorem 1 to find a lower bound on the query complexity of every nonadaptive algorithm for a *nontrivial* graph property in the bounded-degree model[1].

Definition 2 (Nontrivial Graph Property)**.** *We call a graph property* nontrivial *if it does not depend only on the degree distribution of the nodes: namely, for all sufficiently large n there is some degree sequence $d_1, \ldots, d_n \in \{0, 1, \ldots, d\}$ such that there is at least one graph, G_1, with node degrees $d_1, \ldots, d_n$ with the property and at least one, G_2, that is ϵ-far.*

We prove the following theorem.

Theorem 3. *Every nonadaptive tester for a* nontrivial *property in the* adjacency lists model *requires $\Omega(\sqrt{n}/d)$ queries.*

Proof. We give the tester more power: every time it queries a neighbor i of vertex v, it will get the entire adjacency list of vertex v, i.e., it will find out a "star" portion of the graph with v in the center and its neighbors connected to it. The main idea is that with $q = o(\sqrt{n}/d)$ queries even the enhanced (nonadaptive) tester will see only a disjoint collection of stars with high probability. Therefore, the only information the tester will be able to collect is the degrees of q vertices.

Let G_1 and G_2 be as stated in Definition 2. Let $\mathcal{P}$ be a random isomorphic copy of G_1 and $\mathcal{N}$ be a random isomorphic copy of G_2. Recall the definition of $\mathcal{P}$-view and $\mathcal{N}$-view from the previous section. For two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$, let $\mathcal{D}_1 \approx_\delta \mathcal{D}_2$ denote that the statistical difference between the two distributions is at most δ. By Theorem 1, it is enough to show that $\mathcal{P}$-view $\approx_{\frac{1}{4}} \mathcal{N}$-view.

Let BAD denote the event that two stars centered at query points intersect, namely, that for some pair of queries v and u there is some vertex s such that both (u, s) and (v, s) are edges in the input graph. Let I be a random variable that denotes the number of such intersecting pairs. If v_i is the node that got mapped to node i under random isomorphism of graph G then the set containing v_i, the neighbors of v_i and the neighbors of neighbors of v_i has at most $d^2 + 1$ nodes. Under random isomorphism, the probability that one of these nodes is mapped to node j is at most $\frac{d^2+1}{n}$. Therefore, for both distributions $\mathcal{P}$-view and $\mathcal{N}$-view,

$$E[I] \leq \binom{q}{2} \frac{d^2+1}{n} \leq \frac{1}{17}, \text{ for sufficiently large } n.$$

Consequently,

$$\Pr[BAD] = \Pr[I > 0] = \sum_{i=1}^{\infty} \Pr[I = i] \leq \sum_{i=1}^{\infty} i \Pr[I = i] = E[I] \leq \frac{1}{17}.$$

We will show that conditioned on BAD not occurring, (1) the tester learns only the degrees of the queried nodes, and (2) the distributions on the degree list seen by the tester are similar under $\mathcal{P}$ and $\mathcal{N}$. For a graph G, let $d_1, ..., d_q(G)$ be the degrees of the vertices of G queried by the tester. Let $\mathcal{P}$-degs be the distribution $d_1, ..., d_q(G)$ when G is selected according to $\mathcal{P}$. Similarly, define $\mathcal{N}$-degs. Observe that $\mathcal{P}$-degs $= \mathcal{N}$-degs because of our condition on the degrees of G_1 and G_2.

To show that knowing adjacency lists does not give any advantage over knowing only the degree list when BAD does not occur, define a randomized algorithm $\mathcal{A}$ that converts a degree list to a possible set of answers. On input $d_1, ..., d_q$, $\mathcal{A}$ picks $d = \sum_{i=1}^{q} d_i$ random numbers from $\{q + 1, ..., n\}$ *without replacement* and outputs those numbers in order as elements of the adjacency lists of the nodes $1, \ldots, q$. Note that $\mathcal{A}$ always produces non-intersecting adjacency lists (i.e., $\mathcal{A}$ simulates a world where BAD never happens).

Claim 4. Conditioned on BAD not occurring*, the output of $\mathcal{A}$ is distributed according to*

[1]The rest of the text is almost directly taken from [RS06].

- $\mathcal{P}$-view *when its input is distributed according to* $\mathcal{P}$-degs*;*
- $\mathcal{N}$-view *when its input is distributed according to* $\mathcal{N}$-degs*.*

That is, in symbols, $\mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}}) = \mathcal{P}\text{-view}|_{\overline{BAD}}$ *and* $\mathcal{A}(\mathcal{N}\text{-degs}|_{\overline{BAD}}) = \mathcal{N}\text{-view}|_{\overline{BAD}}$.

Proof. We prove the claim only for distribution $\mathcal{P}$. The same proof works for $\mathcal{N}$. First, observe that the distribution on lists of degrees in $\mathcal{P}\text{-view}|_{\overline{BAD}}$ and in $\mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}})$ is the same: in both cases it is $\mathcal{P}\text{-degs}|_{\overline{BAD}}$, by definition. Thus, it is sufficient to prove that for each possible degree list $d_1, ..., d_q$, the distribution on neighbor lists $a_1, ..., a_q$ is the same in both distributions.

Consider any two non-intersecting sequences of adjacency lists $a_1, ..., a_q$ and $a'_1, ..., a'_q$ which correspond to the same degree list $d_1, \ldots, d_q$. Since $\mathcal{A}(d_1, ..., d_q)$ selects a non-intersecting sequence uniformly at random (from the set of non-intersecting sequences with degrees $d_1, \ldots, d_q$), it outputs both sequences with the same probability. We will show that they also arise with same probability under $\mathcal{P}$-view. Note that there exists some permutation π of G_1, such that the nodes in $a'_1, ...a'_q$ are the images of the nodes in $a_1, ..., a_q$ under π (since, in both cases, no node appears twice). We can now set up a 1-to-1 correspondence between permutations that give rise to $a_1, ..., a_q$ and permutations that give rise to $a'_1, ..., a'_q$: for any permutation σ of $\{1, ..., n\}$ such that $\sigma(G_1)$ has adjacency lists $a_1, ..., a_q$, the permutation $\pi \circ \sigma$ produces $a'_1, ..., a'_q$; similarly, if σ led to $a'_1, ..., a'_q$, then $\pi^{-1} \circ \sigma$ would lead to $a_1, ..., a_q$. This correspondence is 1-to-1 since we can never have $\pi \circ \sigma_1 = \pi \circ \sigma_2$ unless $\sigma_1 = \sigma_2$.

Because of this correspondence, the two sequences of adjacency lists arise with same probability under $\mathcal{P}$-view, and so $\mathcal{P}\text{-view}|_{\overline{BAD}} = \mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}})$. (Note that the equality would not hold without conditioning on $\overline{BAD}$). □

Conditioning on $\overline{BAD}$ does not significantly change our distributions, as formalized in claim 5.

Claim 5. *Let* E *be an event that happens with probability at least* $1 - \delta$ *under the distribution* $\mathcal{D}$ *and let* $\mathcal{B}$ *denote distribution* $\mathcal{D}|_E$*. Then* $\mathcal{B} \approx_{\delta'} \mathcal{D}$ *where* $\delta' = \frac{1}{1-\delta} - 1$.

Proof. It is enough to show that $\Pr_{\mathcal{B}}[S] \leq \Pr_{\mathcal{D}}[S] + \delta'$ for every event S.

$$\Pr_{\mathcal{B}}[S] = \Pr_{\mathcal{D}}[S|E] = \frac{\Pr_{\mathcal{D}}[S \wedge E]}{\Pr_{\mathcal{D}}[E]} \leq \frac{\Pr_{\mathcal{D}}[S]}{\Pr_{\mathcal{D}}[E]} \leq \frac{\Pr_{\mathcal{D}}[S]}{1-\delta} = \Pr_{\mathcal{D}}[S](1+\delta') \leq \Pr_{\mathcal{D}}[S] + \delta'.$$ □

In particular, if $\delta = \frac{1}{17}$ then $\delta' = \frac{1}{16}$. We will apply the claim four times with these parameters to prove that $\mathcal{P}\text{-view} \approx_{\frac{1}{4}} \mathcal{N}\text{-view}$. First, $\mathcal{P}\text{-view} \approx_{\frac{1}{16}} \mathcal{P}\text{-view}|_{\overline{BAD}}$ and $\mathcal{P}\text{-degs} \approx_{\frac{1}{16}} \mathcal{P}\text{-degs}|_{\overline{BAD}}$. The second statement implies that $\mathcal{A}(\mathcal{P}\text{-degs}) \approx_{\frac{1}{16}} \mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}})$. Putting the two statements together and applying Claim 4 gives $\mathcal{P}\text{-view} \approx_{\frac{1}{8}} \mathcal{A}(\mathcal{P}\text{-degs})$. Similarly, $\mathcal{N}\text{-view} \approx_{\frac{1}{8}} \mathcal{A}(\mathcal{N}\text{-degs})$. It remains to use that $\mathcal{P}\text{-degs} = \mathcal{N}\text{-degs}$ and, consequently, $\mathcal{A}(\mathcal{P}\text{-degs}) = A(\mathcal{N}\text{-degs})$, yielding $\mathcal{P}\text{-view} \approx_{\frac{1}{4}} \mathcal{N}\text{-view}$, as required. □

References

[RS06] Sofya Raskhodnikova and Adam Smith. A note on adaptivity in testing properties of bounded-degree graphs. *ECCC, TR06-089*, 2006.

[Yao77] Andrew C. Yao. Probabilistic computation, towards a unified measure of complexity. In *Proceedings of the Eighteenth Annual Symposium on Foundations of Computer Science*, pages 222–227, 1977.