18.405J/6.841J: Advanced Complexity Theory Spring 2016

Lecture 12: Randomized Communication

Mohammad Bavarian *Scribe: Andrew He*
Scribe Date: Spring 2016

1 Introduction

Recall from last time our definitions of randomized communication complexity: we define $\mathsf{R}_\varepsilon(f)$, $\mathsf{R}^{pub}_\varepsilon(f)$ to be the complexity of randomized protocols which correct correctly compute f with probability at least $1-\varepsilon$, with either private (Alice and Bob each have independent sources of randomness) or public randomness (Alice and Bob have access to a shared random string), respectively. Randomness is free, and does not contribute to the cost of the protocol.

2 Public versus private randomness in communication

Recall that private randomness can easily be simulated by private randomness, by splitting the publicly random bits between Alice and Bob. Moreover, we saw that $\mathsf{R}(\mathsf{EQ}_n) = O(\log n)$, but $\mathsf{R}^{pub}(\mathsf{EQ}_n) = O(1)$, so public randomness can be stronger than private randomness. In this lecture, we show that private randomness is nearly as strong as public randomness.

Theorem 1 (Newmann)**.** *For all $f : X \times Y \to \{0,1\}$,*

$$\mathsf{R}_{\varepsilon+\delta}(f) \leq \mathsf{R}^{pub}_\varepsilon(f) + O(\log n + \log 1/\delta) \ .$$

In other words, we can make randomness private at the small, additive logarithmic cost and a small error (which can be boosted to only ε for a small constant factor).

Proof. Consider a public-randomness protocol $\mathcal{P}$ such that $f(x,y) = \mathcal{P}(x,y,r)$ with high probability, and let π denote the set of all strings of randomness r. We would like to construct a private-randomness protocol to solve f.

If we knew that $|\pi| \leq \mathsf{poly}(n)\mathsf{poly}(1/\delta) \iff |r| = O(\log n + \log 1/\delta)$, then we are done by trivial simulation, as Alice can generate private random bits and send them to Bob.

This motivates a reduction approach: in general, we will show that there exists an "equivalent" (up to an error of δ) protocol $\mathcal{P}'$ which has smaller randomness complexity $|\pi'| \leq \mathsf{poly}(n)\mathsf{poly}(1/\delta)$.

Let $r_1, r_2, \ldots, r_t$ be t elements of π selected independently at random (with replacement). For each such choice, $\mathcal{P}$ induces an alternative protocol $\mathcal{P}' = \mathcal{P}_{r_1,\ldots,r_t}$: $\mathcal{P}'$ runs $\mathcal{P}$, except it samples from $\{r_1, \ldots, r_t\}$ instead of from π. Note that $|\pi'| = t$. We now show that, with high probability over our choice of the r_i, $\mathcal{P}'$ makes not much error.

Let $Z(x,y,r) = 1$ if $\mathcal{P}$ makes error on input (x,y) and random string r. By definition, for all (x,y), $\mathbb{E}_{r\sim\pi}[Z(x,y,r)] \leq \varepsilon$. Thus, by Chernoff, for all (x,y),

$$\mathop{\mathbb{P}}_{r_i}\left[\frac{1}{t}\sum_{i=1}^{t} Z(x,y,r_i) > \varepsilon + \delta\right] \leq 2e^{-2\delta^2 t}$$

Thus, if $t \geq \frac{n+1}{\delta^2} \ln 2$, by a union bound,

$$\mathop{\mathbb{P}}_{r_i}\left[\exists(x,y):\ \frac{1}{t}\sum_{i=1}^{t} Z(x,y,r_i) > \varepsilon + \delta\right] \leq 2^{2n}\left(2e^{-2\delta^2 t}\right) < 1$$

so there must exist some choice of $r_1, \ldots r_t$ such that, for all (x,y), $\mathcal{P}'$ fails with probability at most $\varepsilon + \delta$.

Now, π' has only $t = O(n/\delta^2) = \mathsf{poly}(n)\mathsf{poly}(1/\delta)$ different random strings, so we can sample from it using $O(\log n + \log 1/\delta)$ bits of randomness. Thus, we can simulate $\mathcal{P}'$ with private randomness using only $O(\log n + \log 1/\delta)$ extra bits of communication. □

3 Lower bounds on randomized communication complexity

We now try to prove some lower bounds on actual problems. In particular, we will show the inner product problem is not in $\mathsf{BPP}^{\mathsf{cc}}$.

We first define a different sense of randomized communication complexity.

Definition 2 (Distributional communication complexity)**.** *Let μ be a probability distribution on $X \times Y$. The (μ, ε)-distributional communication complexity of f, denote $\mathsf{D}^{\mu}_{\varepsilon}(f)$ is the cost of the best deterministic protocol with error $\leq \varepsilon$ with inputs weighted according to distribution μ.*

Claim 3. $\mathsf{R}^{pub}_{\varepsilon}(f) \geq \max_{\mu} \mathsf{D}^{\mu}_{\varepsilon}(f)$.

Proof. Consider a protocol $\mathcal{P}$ for $\mathsf{R}^{pub}_{\varepsilon}(f)$. Now, by the guarantees of the randomized protocol,

$$\forall(x,y): \mathop{\mathbb{E}}_{r\sim D}[\mathcal{P}(x,y,r) \neq f(x,y)] \leq \varepsilon\ ,$$

so

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu, r\sim D}[\mathcal{P}(x,y,r) \neq f(x,y)] \leq \varepsilon\ .$$

Thus, there exists r such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}[\mathcal{P}(x,y,r) \neq f(x,y)] \leq \varepsilon$$

so we can "fix our randomness" to form a deterministic algorithm. □

Theorem 4. *In fact, this is an equality: $\mathsf{R}^{pub}_{\varepsilon}(f) = \max_{\mu} \mathsf{D}^{\mu}_{\varepsilon}(f)$.*

Proof idea. This is solved by the minimax theorem or LP duality. □

In analyzing deterministic communication complexity, we analyzed partitions into monochromatic rectangles. Now, we want to consider partitions into "almost" monochromatic rectangles. We now define a way to measure this.

Definition 5 (Discrepancy)**.** *Let $f : X \times Y \to \{0,1\}$. Then, the discrepancy of a rectangle R is*

$$\mathrm{Disc}_\mu(f,R) = \left|\mu(R \cap f^{-1}(0)) - \mu(R \cap f^{-1}(1))\right|$$

and the discrepancy of the function is

$$\mathrm{Disc}_\mu(f) = \max_{R=S\times T} \mathrm{Disc}_\mu(f,R) \ .$$

Note that higher discrepancy means that a rectangle is "bigger" and "more monochromatic", so higher discrepancy is better.

Proposition 6. $\mathsf{D}^\mu_\varepsilon(f) \geq \log \frac{1-2\varepsilon}{\mathrm{Disc}_\mu(f)}$

Proof. Consider $\mathcal{P}$ a protocol for $\mathsf{D}^\mu_\varepsilon(f)$. Then, we have

$$1 - 2\varepsilon \leq \mathop{\mathbb{P}}_{\mu}[f(x,y) = \mathcal{P}(x,y)] - \mathop{\mathbb{P}}_{\mu}[f(x,y) \neq \mathcal{P}(x,y)]$$

Let $\{R_1, R_2, \ldots R_m\}$ be the rectangles induces by $\mathcal{P}$. Then,

$$1 - 2\varepsilon \leq \sum_{l=1}^{m} \mathop{\mathbb{P}}_{\mu}[f(x,y) = \mathcal{P}(x,y) \wedge (x,y) \in R_l] - \mathop{\mathbb{P}}_{\mu}[f(x,y) \neq \mathcal{P}(x,y) \wedge (x,y) \in R_l]$$

so there exists some $R = R_l$ so that

$$\mathrm{Disc}_\mu(f) \geq \mathrm{Disc}_\mu(f,R) \geq \frac{1-2\varepsilon}{m} \implies \mathsf{D}^\mu_\varepsilon(f) \geq \log m \geq \log \frac{1-2\varepsilon}{\mathrm{Disc}_\mu(f)} \ .$$

□

Thus, we have a way to bound randomized communication complexity with discrepancy through distributional complexity. We are now ready to prove the main theorem.

Theorem 7. *Consider $\mathsf{IP}_n = \sum_{i=1}^n x_i y_i \bmod 2$. Then, $\mathsf{IP}_n \notin \mathsf{BPP}^{\mathsf{cc}}$. In particular, $\mathsf{R}(\mathsf{IP}_n) = \Omega(n)$.*

Proof. To apply our theorems now, we must pick a distribution μ to work on. In general, picking μ is the art of proving bounds on randomized communication complexity.

In our case, the distribution is easy: let μ be uniform over $\{0,1\}^n \times \{0,1\}^n$. We now show that $\mathrm{Disc}_\mu(\mathsf{IP}_n) \leq 2^{-n/2}$ which suffices in our bounds on R^{pub}.

Now, let $H(x,y) = (-1)^{\langle x,y\rangle} = (-1)^{\sum x_i y_i}$. Then,

$$\mathrm{Disc}_\mu(f) = \max_R \left| \sum_{(x,y)\in R} \mu(x,y) H(x,y) \right| .$$

Let $H_{sml} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Note that H, interpreted as a matrix, equals the tensor product $H_{sml}^{\otimes n}$. The eigenvalues of H_{sml} are $\pm\sqrt{2}$, so $\|H\| \leq 2^{n/2}$. Here, $\|H\|$ denotes the largest eigenvalue of H.

Thus,

$$\begin{aligned}
\mathrm{Disc}_\mu(f, R = S \times T) &= \frac{1}{2^{2n}} \sum_{(x,y)} H(x,y) 1_S(x) 1_T(y) \\
&= \frac{1}{2^{2n}} (1_S)^T H (1_T) \\
&\leq \frac{1}{2^{2n}} \|H\| \sqrt{|S|} \sqrt{|T|} \\
&\leq \frac{2^{n/2} \cdot 2^{n/2} \cdot 2^{n/2}}{2^{2n}} \\
&= 2^{-n/2}
\end{aligned}$$

where 1_S and 1_T are the indicator vectors of S and T. Thus, we have

$$\mathsf{R}(\mathsf{IP}_n) \geq \mathsf{R}^{pub}(\mathsf{IP}_n) = \mathsf{R}^{pub}_{\frac{1}{3}}(\mathsf{IP}_n) \geq \mathsf{D}^{\mu}_{\frac{1}{3}}(\mathsf{IP}_n) \geq \log \frac{1}{3 \cdot 2^{-n/2}} = \frac{n}{2} - O(1) \ .$$

□

Thus, we have shown that $\mathsf{IP}_n \notin \mathsf{BPP}^{\mathsf{cc}}$, which gives us a bound on randomized communication complexity.

18.405J / 6.841J Advanced Complexity Theory

Spring 2016