LECTURE 16: COMMUNICATION COMPLEXITY II

L16:I

Recall

- Alice gets $x \in \mathcal{X}$, Bob gets $y \in \mathcal{Y}$, want to compute $f(x,y)$
- Only care about communication
- Protocol P = binary tree
  - internal node: who speaks? sends which bit?
  - leaf: labelled by function value
- Communication complexity (deterministic) $D(f)$ of f is min cost of any protocol P computing f
  Cost of P max # bits communicated for any input (x,y)
- Combinatorial rectangle $R = A \times B$ $A \subseteq \mathcal{X}$ $B \subseteq \mathcal{Y}$
- R monochromatic if $\exists z \; \forall (x,y) \in R \; f(x,y) = z$

$$\boxed{\text{Any protocol } P \text{ for } f: \mathcal{X} \times \mathcal{Y} \to \mathcal{Z} \text{ induces a partition of } \mathcal{X} \times \mathcal{Y} \text{ into monochromatic rectangles.}}$$

- Lower bounds on # monochromatic rectangles needed, say at least t $\Longrightarrow$ $D(f) \geq \log t$
- Most common set-up $\mathcal{X} = \mathcal{Y} = \{0,1\}^n$; $\mathcal{Z} = \{0,1\}$.
  Trivial upper bound $D(f) \leq n + 1$
  Trivial lower bound $D(f) \geq 1$
- Identify $x \in \{0,1\}^n \iff x \subseteq [n]$
  $x_i = 1 \iff i \in x$
- $EQ(x,y) = [x = y]$ $\qquad D(EQ) = n + 1$
- $DISJ(x,y) = [x \cap y = \emptyset]$ $\qquad D(DISJ) = n + 1$

So far, Alice and Bob computationally all powerful, but deterministic. What if they are allowed to use randomness ("toss coins")?

Communication transcript on (x, y) becomes random variable

Output of protocol also random variable

Several different conceivable flavours

(1) ERROR

(a) "Las Vegas protocols" always correct

(b) "Monte Carlo protocols" correct with high probability

(c) One-sided error Always correct when $f(x,y) = 0$, say, and whp when $f(x) = 1$

(2) RANDOMNESS

(a) private coin Alice has her private random string r_A
Bob has his —"— r_B

(b) public coin Alice & Bob see the same common random string r

(3) COST on input (x,y) for fixed protocol P

(a) worst case over all choices of random strings

(b) average case expected cost (<u>over all random strings</u> for fixed input).

We will focus on

1 (b) Monte Carlo with two-sided error

2 (b) public coin

[difference from private coin not too big, though]

3 (a) worst-case cost

[doesn't really matter, actually]

DEF 1

$R^{pub}_{\varepsilon}(f)$: Worst-case cost over all inputs for best two-sided error public-coin randomized communication protocol P for f such that

$$Pr[P(x,y) \neq f(x,y)] \leq \varepsilon$$

with probability taken over random coin tosses only

$R^{priv}_{\varepsilon}(f)$: Same thing, but for private coins

If we just write $R_{\varepsilon}(f)$, mean public-coin

Error parameter $\varepsilon < 1/2$ doesn't matter too much as long as it is bounded away from 1/2.

So... How hard is $EQ(x,y)$ in this setting?
How hard is $DISJ(x,y)$ in this setting?

Ex 2 Public-coin protocol for EQ

Joint n-bit uniformly random string $r = r_1 r_2 \cdots r_n$

Alice computes inner product

$$\langle x, r \rangle = \sum_{i=1}^{n} x_i \cdot r_i \pmod 2$$

and sends to Bob — 1 bit

Bob computes

$$\langle y, r \rangle = \sum_{i=1}^{n} y_i \cdot r_i \pmod 2$$

and checks if $\langle x, r \rangle = \langle y, r \rangle$

If yes, Bob outputs 1 = equal, else 0 = not equal

— 1 bit

Total of 2 bits of communication

Analysis

If $x = y$, clearly Bob always outputs 1

(so protocol has one-sided error)

Suppose $x \neq y$. Fix arbitrary position j such that $x_j \neq y_j$

Consider values $a = \sum_{i \neq j} x_i r_i$ and $b = \sum_{i \neq j} y_i r_i$

Two cases

(a) $a = b$: If so, with prob $= 1/2$ $r_j = 1$ and $\langle x, r \rangle \neq \langle y, r \rangle$

(b) $a \neq b$: If so, with prob $= 1/2$ $r_j = 0$ and $\langle x, r \rangle = a \neq b = \langle y, r \rangle$

So in both cases error probability exactly $= 1/2$.

Repeat twice $\Rightarrow$ error probability $= 1/4 < 1/2$

So $R^{pub}_{1/4}(EQ) = 4$.

What about set disjointness DISJ ?
Much trickier question — in fact, so tricky that we won't quite be able to deal with it today. Will cover some (vaguely) related results instead.

Public-coin protocol P can sample random string r at each node in protocol tree to make random choices.
But fixing random string r, this yields a completely deterministic protocol P_r

Hence, we can view a public-coin randomized protocol P as a distribution $\{P_r\}_{r \sim \Pi}$ over deterministic protocols, where r is a random string distributed according to Π.
In fact, it's useful to define them this way

DEF 3 A RANDOMIZED PUBLIC-COIN PROTOCOL P is a probability distribution over deterministic protocols $\{P_r\}$. The success probability on input (x,y) is the probability of choosing a (deterministic) protocol P_r according to this distribution such that $P_r(x,y) = f(x,y)$. Then $R_\varepsilon^{pub}(f) = R_\varepsilon(f)$ is the min cost of any public-coin protocol that computes f with success probability $\geq 1 - \varepsilon$

DISTRIBUTIONAL COMPLEXITY

Technique to prove lower bounds for randomized protocols with 2-sided error.

Different setting

- Protocols are deterministic
- Inputs are random according to some probability distribution μ

DEF 4 Let μ be a probability distribution on $\mathcal{X}\times\mathcal{Y}$. The (μ,ε) - DISTRIBUTIONAL COMMUNICATION COMPLEXITY of f, denoted $D^{\mu}_{\varepsilon}(f)$ is the cost of an optimal deterministic protocol P that gives the correct answer for f for at least a fraction $(1-\varepsilon)$ of all inputs in $\mathcal{X}\times\mathcal{Y}$ as weighted by μ, i.e.,

$$\Pr_{(x,y)\sim\mu}\left[P(x,y)\neq f(x,y)\right] \leq \varepsilon$$

EX 5 Let $\mathcal{U}$ denote uniform distribution over $\mathcal{X}\times\mathcal{Y}$.

Let $GT(x,y) = [x>y] = \begin{cases} 1 & \text{if } x>y \\ 0 & \text{otherwise}\end{cases}$

Then $D^{\mathcal{U}}_{1/4}(GT) \leq 2$

numbers in binary

Alice has $x=(x_{n-1},x_{n-2},\ldots,x_1,x_0)$ Bob has $y=(y_{n-1},\ldots,y_0)$

Alice sends Bob most significant bit x_{n-1} of x

Bob compares with y_{n-1}

If $x_{n-1}\neq y_{n-1}$, Bob knows answer. Happens with prob $1/2$.

If $x_{n-1}=y_{n-1}$, Bob answers 0, say. Correct with prob $1/2$ and this scenario happens with prob $1/2$.

Total success probability $= 1/2 + 1/2\cdot 1/2 = 3/4$ $\boxed{\text{Lw}}$

So what is the connection?

Turns out that distributional communication complexity lower bounds <u>completely determine</u> randomized public-coin communication complexity.

THM 6 [Yao '83]

$$R_\varepsilon^{pub}(f) = \max_\mu D_\varepsilon^\mu(f)$$

<u>Proof</u> <u>($\geq$ direction)</u>

Let $\mathcal{P}$ protocol achieving cost $R_\varepsilon^{pub}(f)$. $\mathcal{P}$ is correct for <u>every input</u> with probability $\geq 1-\varepsilon$

Thus

$$\Pr_{r \in \Pi, (x,y)\sim\mu}\left[\mathcal{P}_r(x,y) = f(x,y)\right] \geq 1-\varepsilon$$

Switch order of summation

$$\sum_{r^*} \Pr[r = r^*] \Pr_{(x,y)\sim\mu}\left[\mathcal{P}_r(x,y) = f(x,y)\right] \geq 1-\varepsilon$$

But if this sum is $\geq 1-\varepsilon$, then in particular there must exist some fixed r^* s.t.

$$\Pr_{(x,y)\sim\mu}\left[\mathcal{P}_{r^*}(x,y) = f(x,y)\right] \geq 1-\varepsilon$$

This deterministic protocol $\mathcal{P}_{r^*}$ shows that

$$D_\varepsilon^\mu(f) \leq R_\varepsilon^{pub}(f).$$

<u>($\leq$ direction)</u>

Requires some basic game theory: zero-sum games and von Neumann's Minimax theorem. Not very hard, but beyond the scope of this course. (And we won't need it today)

DISCREPANCY

Recall that for deterministic communication complexity, we showed that if for a function f the matrix M_f [with $M(x,y) = f(x,y)$] has no large monochromatic rectangles, then # rectangles in partition of $X \times Y$ large $\Rightarrow$ $D(f)$ large

For distributional comm cplx, we say that a rectangle $S \times T$, $S \subseteq X$, $T \subseteq Y$, is "unbalanced" if it contains mostly 1s or mostly 0s w.r.t. μ.

To show large distributional comm cplx, sufficient to prove that unbalanced rectangles must be small. (If a protocol is mostly correct w.r.t. μ, then intuitively most rectangles should be clearly unbalanced)

Let's formalize this!

<u>Nice notation</u>: For probability distr μ over D and $D' \subseteq D$, let $\mu(D') = \sum_{\mathbf{z} \in D'} \mu(\mathbf{z})$

$\mu(D')$ = probability to see $\mathbf{z} \in D'$ when sampling randomly according to μ.

DEF 7 $f : X \times Y \to \{0,1\}$ function
R rectangle
μ probability distribution on $X \times Y$

The DISCREPANCY of f on R ~~w.r.t.~~ according to μ is

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

The DISCREPANCY OF f according to μ is

$$\mathrm{Disc}_\mu(f) = \max_R \left(\mathrm{Disc}_\mu(R,f) \right)$$

where max taken over all rectangles $R \subseteq X \times Y$.

LEMMA 8

For every function $f : X \times Y \to \{0,1\}$, every probability distribution μ on $X \times Y$, and every $\varepsilon \in (0, 1/2)$ it holds that

$$D^{\mu}_{\frac{1}{2}-\varepsilon}(f) \geq \log_2\left(\frac{2\varepsilon}{\mathrm{Disc}_\mu(f)} \right)$$

So if we can prove strong upper bounds on discrepancy, then we get strong lower bounds on randomized public-coin communication cplx.

Proof Let $\mathcal{P}$ be a protocol of cost c that computes f correctly with prob $\geq \frac{1}{2} + \varepsilon$ (w.r.t. μ). Then we can upper-bound the extra advantage ε over "random guessing" as follows.

$$2\varepsilon = \left(\frac{1}{2} + \varepsilon\right) - \left(\frac{1}{2} - \varepsilon\right)$$

$$\leq \Pr_{\mu}\left[P(x,y) = f(x,y)\right] - \Pr_{\mu}\left[P(x,y) \neq f(x,y)\right]$$

$$= \sum_{\substack{\text{leaf } \ell \text{ of} \\ \text{protocol } P}} \Big(\Pr_{\mu}\left[(x,y) \in R_\ell \text{ and } P(x,y) = f(x,y)\right] - \Pr_{\mu}\left[(x,y) \in R_\ell \text{ and } P(x,y) \neq f(x,y)\right]\Big) \quad (\dagger)$$

For each leaf ℓ the deterministic protocol P gives a specific output 0 or 1. Hence, we can upper bound

$$(\dagger) \leq \sum_{\ell} \Big| \Pr_{\mu}\left[(x,y) \in R_\ell \text{ and } f(x,y) = 1\right] - \Pr_{\mu}\left[(x,y) \in R_\ell \text{ and } f(x,y) = 0\right] \Big|$$

$$= \sum_{\ell} \left| \mu\left(f^{-1}(1) \cap R_\ell\right) - \mu\left(f^{-1}(0) \cap R_\ell\right) \right| \quad (\ddagger)$$

By definition, each term in $(\ddagger)$ is bounded from above by $\mathrm{Disc}_{\mu}(f)$.

At most 2^c leaves = terms. So:

$$(\ddagger) \leq 2^c \cdot \mathrm{Disc}_{\mu}(f)$$

That is,
$$2^c \geq \frac{2\varepsilon}{\mathrm{Disc}_{\mu}(f)}$$

or taking logarithms

$$c \geq \log_2\left(\frac{2\varepsilon}{\mathrm{Disc}_{\mu}(f)}\right) \quad \boxed{\square}$$

<u>Inner product</u>

$$IP(x,y) = \sum_{i=1}^{n} x_i y_i \pmod 2$$

<u>THM 9</u> [Chor & Goldreich '85 (?)]

$$R_{\varepsilon}^{pub}(IP) = \Omega(n)$$

Follows immediately if we can prove strong upper bound on discrepancy. Will use <u>uniform distribution</u> $\mathcal{U}$ over $X \times Y = \{0,1\}^n \times \{0,1\}^n$

so we set $\mu = \mathcal{U}$

<u>LEMMA 10</u>

$$Disc_{\mathcal{U}}(IP) = 2^{-\Omega(n)}$$

We will conclude our discussion of communication complexity by proving this lemma.

Let R' be any fixed rectangle

$$R' = \{ (x,y) \mid x \in \mathcal{X}', y \in Y' \}$$

for some $\mathcal{X}' \subseteq \mathcal{X}$, $Y' \subseteq Y$

Write
$$R'(x,y) = \begin{cases} 1 & \text{if } (x,y) \in R' \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{X}'(x) = \begin{cases} 1 & \text{if } x \in \mathcal{X}' \\ 0 & \text{otherwise} \end{cases}$$

$$Y'(y) = \begin{cases} 1 & \text{if } y \in Y' \\ 0 & \text{otherwise} \end{cases}$$

Switch notation to

$$IP : \{0,1\}^n \times \{0,1\}^n \to \{\pm 1\}$$

$$IP(x,y) = (-1)^{\sum x_i y_i}$$

Then

$$d := Dis_{\mathcal{U}}(R', IP) =$$

$$= \left| \Pr_{\mathcal{U}} [IP(x,y) = 1 \text{ and } R'(x,y) = 1] - \Pr [IP(x,y) = -1 \text{ and } R'(x,y) = 1] \right|$$

$$= \left| E_{\mathcal{U}} \left[IP(x,y) \cdot R'(x,y) \right] \right|$$

$$= \left| E_{\mathcal{U}} \left[(-1)^{x_1 y_1 + \dots + x_n y_n} X'(x) Y'(y) \right] \right|$$

$$\leq E_x \left[\underbrace{|X'(x)|}_{\leq 1} \cdot \left| E_y \left[(-1)^{x_1 y_1 + \dots + x_n y_n} \cdot Y'(y) \right] \right| \right]$$

$$\leq E_x [Z(x)] \qquad (*)$$

for $Z(x) = \left| E_y \left[(-1)^{\sum x_i y_i} \; Y'(y) \right] \right|$

We know that

$$E[Z^2] \geq (E[Z])^2$$

(E.g. observe that $0 \leq E\left[(Z - E[Z])^2 \right]$
$= E[Z^2] - (E[Z])^2$)

Combining this with $(*)$ we get

$$d^2 \le \left(\mathcal{E}_x [Z(x)] \right)^2$$

$$\le \mathcal{E}_x \left[Z^2(x) \right]$$

$$= \mathcal{E}_x \left[\mathcal{E}_u \left[(-1)^{\sum x_i u_i} Y'(u) \right] \cdot \mathcal{E}_v \left[(-1)^{\sum x_i v_i} Y'(v) \right] \right]$$

$$= \mathcal{E}_x \left[\mathcal{E}_{u,v} \left[(-1)^{\sum_i (u_i + v_i) x_i} \underbrace{Y'(u)\ Y'(v)}_{\le 1} \right] \right]$$

switch order of summation

$$\le \mathcal{E}_{u,v} \left[\mathcal{E}_x \left[(-1)^{(u_1+v_1)x_1 + \dots + (u_n+v_n)x_n} \right] \right] \quad (**)$$

This is starting to look vaguely familiar...

If $u \ne v$ then

$$\mathcal{E}_x \left[(-1)^{\sum_i (u_i + v_i) x_i} \right]$$

$$= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{\sum_i (u_i+v_i) x_i} = 0$$

This is just $\boxed{\text{problem 3 on pset 3}}$ $\boxed{\text{the hint in}}$ with slightly different notation.

If $u = v$ then

$$\mathcal{E}_x \left[(-1)^{\sum_i (u_i + v_i) x_i} \right] = \mathcal{E}_x \left[(-1)^{\sum_i 0 \cdot x_i} \right] = \underline{1}$$

But this happens only with probability 2^{-n}.

Hence, from (**) we get

$$d^2 = \left[Disc_U (R', IP) \right]^2 \le 2^{-n}$$

~~and~~ for any R' and hence

$$Disc_U (IP) \le 2^{-n/2}$$

which proves the lemma.

Unfortunately, discrepancy is **provably** too weak a tool to prove lower bounds for DISJ

Lower bound $\boxed{R_\varepsilon^{pub}(DISJ) = \Omega(n)}$

can be proven by other means, but this is beyond the scope of this course