LECTURE 17

Communication complexity

Alice gets $x \in \mathcal{X}$, Bob gets $y \in \mathcal{Y}$
want to compute $f(x,y)$ for
some fixed $f: \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$
Today: $\mathcal{X} = \mathcal{Y} = \{0,1\}^n$ $\quad \mathcal{Z} = \{0,1\}$
Only care about communication between Alice & Bob
Computation is for free

Combinatorial rectangle $R = A \times B$; $A \subseteq \mathcal{X}$, $B \subseteq \mathcal{Y}$
R monochromatic if $\exists z \, \forall (x,y) \in R \;\; f(x,y) = z$

Any deterministic protocol P for f induces a partition of $\mathcal{X} \times \mathcal{Y}$ into monochromatic rectangles

Namely, every input (x,y) ends up in exactly one leaf of the protocol tree

<u>Randomized</u> communication complexity
Alice and Bob can flip coins
- Joint random string r
- Must have $\Pr[P(x,y) = f(x,y)] \geq 1 - \varepsilon$
  for <u>all inputs</u>; probability over random coins
- Cost of protocol: worst-case over inputs and random strings.

$R_\varepsilon^{pub}(f) = R_\varepsilon(f) =$ worst-case cost of best randomized protocol for computing f

How to prove lower bounds?
- No partition of inputs
- No rectangles in leaves
- No monochromaticity

Randomized protocol becomes deterministic once random string is fixed. Hence

A public-coin randomized protocol P is a probability distribution ~~over~~ $\{P_r\}_r$ over deterministic protocols P_r

$$\Pr[P \text{ computes } f(x,y) \text{ correctly}] =$$

$$\Pr\left[\begin{array}{l}\text{Choosing deterministic protocol } P_r \\ \text{such that } P_r \text{ is correct on } (x,y)\end{array}\right]$$

So maybe it can help to move a step back in the direction of deterministic communication. One more step:

Distributional communication complexity:

Protocols deterministic, inputs random

μ probability distribution on $\mathcal{X} \times \mathcal{Y}$

(μ, ε) - distributional communication complexity $\mathcal{D}^{\mu}_{\varepsilon}(f)$ is cost of optimal protocol s.t.

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

Distributional communication complexity <u>exactly characterizes</u> public-coin randomized communication complexity

Numbering from last lecture

THM 6 [Yao '83]

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

We will only need $\geq$-direction (which we proved).

So now we're back (almost) to deterministic setting!

- Protocol <u>partitions</u> input space $\mathcal{X} \times \mathcal{Y}$
- Partition into <u>rectangles</u> = for each leaf ℓ in protocol tree P, the inputs (x,y) that end up there
- <u>No monochromaticity</u>, but almost... Most rectangles had better be "almost monochrome" (where inputs (x,y) weighted according to μ), or else the protocol is wrong on more than ε-fraction of inputs.

Leads to notion of <u>discrepancy</u>

How large and almost monochrome rectangles can f have (when we view f as an $\mathcal{X} \times \mathcal{Y}$-matrix M_f with (x,y)-entry containing $f(x,y)$)?

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

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

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

LEMMA 8
For every function $f : \mathcal{X} \times Y \to \{0,1\}$, every probability distribution μ on $\mathcal{X} \times Y$, and every $\varepsilon \in (0, 1/2)$ it holds that

$$D^\mu_{\frac{1}{2}-\varepsilon}(f) \geq \log_2 \left(\frac{2\varepsilon}{\mathcal{D}isc_\mu(f)} \right)$$

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.

Would like to do this for set disjointness $\mathrm{DISJ}(x,y) = [x \cap y = \emptyset]$
Unfortunately, this doesn't work. We'll prove a lower bound for a somewhat related function instead.

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

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

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

For each leaf ℓ the deterministic protocol $\mathcal{P}$ 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 (sets in partition of $X \times Y$). 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

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

Rectangle with high discrepancy

Inner product

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

THM 9 [Chor & Goldreich '85]

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

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

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

LEMMA 10

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

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

$d := \mathrm{Disc}_{\mu}(R', IP) =$

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

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

$$= \left| E_{\mu}\left[(-1)^{x_1y_1 + \dots + x_ny_n} X'(x)\, Y'(y) \right] \right|$$

$$\leq E_x\Big[\underbrace{|X'(x)|}_{\leq 1} \cdot \left| E_y\left[(-1)^{x_1y_1+\dots+x_ny_n} \cdot Y'(y) \right] \right| \Big]$$

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

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 $0 \leq E[(Z - E[Z])^2]$
$= E[Z^2] - (E[Z])^2$)

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 $u \neq v$ 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

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

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

Hence, from (**) we get

$$d^2 = \left[\mathrm{Disc}_u(R', IP)\right]^2 \leq 2^{-n}$$

~~and~~ for <u>any</u> R' and hence

$$\mathrm{Disc}_u(IP) \leq 2^{-n/2}$$

which proves the lemma.

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

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

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

Now we have

$$R^{pub}_{\varepsilon}(IP) \geq D^{u}_{\varepsilon}(IP)$$

$$\geq \log\left(\frac{1-2\varepsilon}{\mathrm{Disc}_u(IP)}\right)$$

$$\geq \log\left(\frac{1-2\varepsilon}{2^{-n/2}}\right)$$

$$= \Omega(n)$$

Computational complexity theory:
How hard are different problems
(1) to solve
(2) in a computational model
(3) wrt some specified resources

(1) SOLVE
We looked mostly <u>not</u> at computing general functions but at <u>deciding problems/languages</u> = computing Boolean functions.
No big deal.

(a) Exact, ~~solution~~ correct
(b) Correct solution with high probability
(c) Approximately correct solution
(d) Combination of (b) & (c)

(2) MODEL
(a) Turing machines
(i) vanilla
(ii) + randomness
(iii) + nondeterminism
(iv) + oracles
(v) + advice

(b) Circuits
(c) Interactive proofs
No trust between computationally bounded verifier and all-powerful prover

(d) Proof verification
(i) randomized ⇝ PCP
(ii) deterministic ⇝ proof cplx

(3) RESOURCES

(a) time

(b) space (memory)

(c) randomness

(d) depth (of circuits)

(e) queries to input

(f) communication (i) total # bits
(ii) # rounds

(g) secrecy / security

WHAT WE DIDN'T COVER?

Most of computational complexity theory...
(Well, not really)

Notable omissions (incomplete list)

- Quantum computation
- Algebraic computation models (circuits & others)
- Complexity of counting (how many satisfying assignments does F have)
- Average-case complexity (not worst-case)
- Derandomization
- Pseudorandomness
- Learning theory
- Kolmogorov complexity
- Limitations to lower bound techniques (Natural proofs Razborov Rudich)
- Streaming algorithms
- "Big data" computation (map-reduce and similar)
- And the list just goes on...

But I hope you liked our selection of topics ... :-)

PLEASE LET US KNOW WHAT YOU THINK by filling in the web-based COURSE EVALUATION! (Should be out next week.)

AND GOOD LUCK with problem sets 3 and 4 !