LECTURE 3: RANDOMIZED COMMUNICATION COMPLEXITY

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

Communication transcript on (x,y) becomes a random variable

Output also r.v.

Several different flavours

(1) ERROR
- (a) "Las-Vegas protocols" always correct
- (b) "Monte-Carlo protocols" correct with high prob.
- (c) One-sided error: always correct when $f(x,y)=0$, say, and correct w.h.p when $f(x,y)=1$.

(2) RANDOMNESS
- (a) private-coin: Alice has her private random string r_A and Bob has r_B
- (b) public-coin: Alice & Bob see the same common random string

(3) COST on input (x,y) for fixed protocol P
- (a) worst-case over all choices of random strings
- (b) average-case expected cost.

Note that average case is over random strings for fixed input (x,y), not over inputs.

We will focus on:

1(b) Monte-Carlo with two-sided error

2(b) public-coin, although will briefly discuss also private coin.

[Difference not too big – pset 1 will ask you to figure out part of this.]

3(a) worst case cost

[Doesn't really matter – again see pset]

DEF 1

$R_\varepsilon^{pub}(f)$: Worst case cost over all inputs for best two-sided error public-coin randomized communication protocol P for f s.t.

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

with prob taken over random coin tosses only.

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

If we just write R_ε, we mean public coin.

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

Ex 2 EQ with private-coin protocol

Alice gets input $a = a_0 a_1 \dots a_{n-1}$

Bob $\quad\dots\quad b = b_0 b_1 \dots b_{n-1}$

Think of a & b as polynomials over $GF[p] = \mathbb{F}_p$

for $n^2 < p < 2n^2$ prime (must exist)

$$a(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1} \pmod p$$
$$b(x) = b_0 + b_1 x + \dots + b_{n-1} x^{n-1} \pmod p$$

[say smallest such p for concreteness]

Alice picks unif random $c \in \mathbb{F}_p$ and sends Bob c and $a(c)$ $\quad O(\log p) = O(\log n)$ bits

Bob evaluates $b(c)$ and outputs 1 if $a(c) = b(c)$ and 0 otherwise.

If $a = b$, clearly answer always 1 (so one-sided error).

If $a \neq b$ polynomials $a(x)$ & $b(x)$ are different.

$a(x) - b(x)$ is non-zero poly of degree $\leq n-1$

$\Rightarrow$ at most $n-1$ zeros

$\Rightarrow$ $a(c) = b(c)$ on at most $\frac{n-1}{p} \leq \frac{n}{n^2} = 1/n$ points. So error probability $\leq 1/n$

Hence $R^{priv}_{1/n}(EQ) = O(\log n)$

(Recall $D(EQ) = n+1$.)

Ex 3 EQ with public-coin protocol

Joint n-bit random string $r = r_0 r_1 \cdots r_{n-1}$

Alice computes inner product

$$\langle a, r \rangle = \sum_{i=0}^{n-1} a_i r_i \pmod 2$$

and sends to Bob. Bob checks whether

$\langle a, r \rangle = \langle b, r \rangle$ and if so outputs 1 for equal, else 0 for not equal.

2 bits of communication.

If $a = b$, clearly answer always 1 (so again one-sided error).

Suppose $a \neq b$; let j be a position such that $a_j \neq b_j$, say $a_j = 1$ $b_j = 0$

Consider value of $\sum_{i \neq j} a_i r_i$ and $\sum_{i \neq j} b_i r_i$

If they are equal, w prob $= 1/2$ $r_j = 1$ and so will flip $\sum_i a_i r_i$

If they are not equal, w prob $= 1/2$ $r_j = 0$ and they stay distinct. So error prob $= 1/2$

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

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

Some facts

Fact 4 For private-coin Las-Vegas protocols (zero error), using worst-case cost just gives us back deterministic CC.

This is so since a deterministic protocol can just fix r_A and r_B and then run the "randomized" protocol

Prop 5 $R_\varepsilon^{priv}(f) = \Omega(\log \frac{1}{\varepsilon} D(f))$.

Proof Skipped because of time constraints.

For public-coin protocols, we can view them as distributions $\{P_r\}_{r \in \Pi}$ over deterministic protocols, where r is chosen randomly acc to Π and Alice & Bob then follow deterministic P_r. Useful to define it this way

DEF 6 A RANDOMIZED PUBLIC-COIN PROTOCOL is a probability distribution over deterministic protocols. The success probability on input (x,y) is the prob of choosing a (deterministic) protocol P acc to distr Π such that $P(x,y) = f(x,y)$. Then $R_\varepsilon^{pub}(f)$ [or just $R_\varepsilon(f)$] is the min cost of a public-coin protocol that computes f with success prob $\geq 1-\varepsilon$.

DISTRIBUTIONAL COMPLEXITY

Technique to prove lower bounds for randomized protocols with 2-sided error. Consider random inputs to be acc to some probability distribution.

DEF 7 Let μ be a prob distro on $X \times Y$. The (μ, ε)- DISTRIBUTIONAL COMMUNICATION COMPLEXITY of f, denoted $D^{\mu}_{\varepsilon}(f)$ is the cost of the best deterministic protocol P that gives the correct answer for f for at least an $(1-\varepsilon)$ fraction of all inputs in $X \times Y$ weighted by μ, i.e. s.t.

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

EX 8 Let unif denote the uniform distribution.
Recall Let $GT(x,y) = [x > y]$

$$D^{unif}_{1/4}(GT) \leq 2.$$

Alice sends Bob most significant (x_{n-1}) bit of x
Bob compares with y_{n-1}
If $x_{n-1} \neq y_{n-1}$, Bob knows the answer
this happens w prob $1/2$.
If $x_{n-1} = y_{n-1}$, Bob answers 0, say.
This happens for the other half of inputs and is correct w prob $1/2$. Total success prob $3/4$.

Distributional comm cplx LB completely determines public-coin comm cplx

Theorem 9 [Yao '83]

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

Proof of $\geq$ direction: Let P protocol achieving cost $R_\varepsilon^{pub}(f)$. P correct for *every* input in prob $> 1-\varepsilon$

Thus $\Pr_{r\sim\Pi,\,(x,y)\sim\mu}\left[P_r(x,y) = f(x,y)\right] > 1-\varepsilon$

This is = [just by switching order of summation]

$$\sum_{r^*} \Pr[r^* = r] \Pr_{(x,y)\sim\mu}\left[P_r(x,y) = f(x,y)\right] = (*)$$

If $(*) > 1-\varepsilon$, then in particular must exist some fixed r^* s.t.

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

This protocol P_{r^*} shows that

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

For $\leq$ direction we need a small detour into zero-sum games.

2-PLAYER ZERO-SUM GAMES

Players 1 & 2 make moves simultaneously, and outcome determined by combination of moves.
Player 1 has m possible moves
Player 2 has n possible moves.

$(m \times n)$-matrix A determines outcome

If P1 chooses i & P2 chooses j,
P2 has to pay P1 A_{ij} (might be positive or negative)

Randomized strategies (column vectors)
for P1 distr $\sigma \in \Delta$ on rows
for P2 distr $\mu \in \Delta$ on columns

With prob $\sigma_i \cdot \mu_j$ pay-off is A_{ij}
Sum over all i,j to get expected pay-off.

This is $\sigma^T A \mu$

P1 wants to maximize this
P2 wants to minimize this

The Minimax thm by [von Neumann '28] says that there are optimal mixed strategies
One of the players can announce his/her strategy and the other one respond adversarially, and it doesn't matter.

THM 10

$$\max_{\sigma} \min_{\mu} \sigma^T A \mu = \min_{\mu} \max_{\sigma} \sigma^T A \mu$$

Proof of $\leq$ direction

Player 1 chooses determ prot P of cost $\leq c$ for $c = \max_\mu D^\mu_\varepsilon(f)$

Player 2 chooses $(x,y) \in X \times Y$.

Pay-off for P1 is 1 if $P(x,y) = f(x,y)$
0 otherwise.

[Want to prove $\exists$ randomized protocol of cost c that works for all $(x,y) \in X \times Y$ with prob $\geq 1-\varepsilon$.]

Whatever the μ chosen by Player 2, Player 1 can find a deterministic protocol that works w prob $\geq 1-\varepsilon$ wrt μ.
Let σ be strategy with coordinate 1 for this row and 0 elsewhere. Shows

$$\min_\mu \max_\sigma \sigma^T A \mu \geq 1-\varepsilon.$$

By Minimax thm $\max_\sigma \min_\mu \sigma^T A \mu \geq 1-\varepsilon$

that is, $\exists$ distr σ over deterministic protocols that works w prob $\geq 1-\varepsilon$ for any distr μ.
In particular, for distr putting all probability mass on a single (x,y). I.e., we found a distr ~~of~~ over determ prot that succeeds w prob $\geq 1-\varepsilon$ in worst case = randomized prot acc to Def 6.

QED

DISCREPANCY

Recall for deterministic comm cplx, we showed that if for function f 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 distr CC, we say that rectangle $S \times T$, $S \subseteq X$, $T \subseteq Y$, is "unbalanced" if it is mostly 1s or mostly 0s wrt μ. To show large distributional comm cplx, sufficient to show that unbalanced rectangles must be small. (If a protocol is mostly correct wrt μ, then intuitively most rectangles should be unbalanced.) More formally:

DEF 11 $f: X \times Y \to \{0,1\}$ function, R rectangle, μ prob distr on $X \times Y$

The DISCREPANCY of f on R under μ is

$$Disc_\mu(R,f) = \left| Pr_\mu[(x,y) \in R \text{ and } f(x,y)=0] - Pr_\mu[(x,y) \in R \text{ and } f(x,y)=1] \right|$$

The DISCREPANCY of f according to μ is

$$Disc_\mu(f) = \max_R Disc_\mu(R,f)$$

where max taken over all rectangles R.

LEMMA 12

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

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

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

Proof Let P protocol of cost c that computes f correctly w prob $\geq 1/2 + \varepsilon$ (wrt μ)
Then we can upper bound the extra advantage ε over "random guessing" as follows.

$$\begin{aligned} 2\varepsilon &= \left(\tfrac{1}{2} + \varepsilon\right) - \left(\tfrac{1}{2} - \varepsilon\right) \\ &\leq Pr_{\mu}[P(x,y) = f(x,y)] - Pr_{\mu}[P(x,y) \neq f(x,y)] \\ &= \sum_{\substack{i \\ \text{leaf } \ell \text{ of} \\ \text{protocol } P}} \Big(Pr_{\mu}[(x,y) \in R_\ell \text{ and } P(x,y) = f(x,y)] \\ &\qquad - Pr_{\mu}[(x,y) \in R_\ell \text{ and } P(x,y) \neq f(x,y)] \Big) = \dagger \end{aligned}$$

For each leaf ℓ P gives a specific output 0 or 1, and clearly picking the more likely answer wrt μ is the best strategy.

Hence

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

$$= (\ddagger)$$

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

At most a total of 2^c leaves = terms

So

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

That is,

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

or taking logarithms

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

No examples today...

Will return to related ideas in future lectures