CS 2429 - Foundations of Communication Complexity

Lecture #6: 20 October 2009

Lecturer: Toniann Pitassi

Scribe Notes by: Lila Fontes

1 Review of complexity classes

We've discussed P^{cc}, RP^{cc}, BP^{cc}, NP^{cc}, and $coNP^{cc}$. For the regular (not communication complexity) classes, we think that $P = RP = BP$, but for the communication classes, we have seen separations:

- $P^{cc} \subsetneq RP^{cc}$ by EQUALITY.
- $RP^{cc} \subsetneq BPP^{cc}$ by GT.
- $BPP^{cc} \subsetneq NP^{cc}$ by SET-DISJOINTNESS.
- NP^{cc} and $coNP^{cc} \nsubseteq BPP^{cc}$ by GT.
- $NP^{cc} \nsubseteq$ and $\nsupseteq coNP^{cc}$ by SET-DISJOINTNESS.

IP is not in any of these classes.

Definition Function f has an *unbounded error* communication complexity protocol P if P is a randomized protocol such that

$$\forall x, y \ \Pr[P(x, y) = f(x, y) > \frac{1}{2}]$$

$U(f) \overset{\text{def}}{=}$ least cost of an unbounded error protocol for f
$PP^{cc} \overset{\text{def}}{=}$ set of all functions that have unbounded error protocols of cost $\text{poly} \log n$.

Thus RP^{cc}, BPP^{cc}, and NP^{cc} are special cases of PP^{cc}.

Theorem 1 (Paturi-Simon) $U(f) = \log(\text{sign-rank}(M_f)) \pm O(1)$ *where* $\text{sign-rank}(M_f) \overset{def}{=}$ *the minimum rank of any matrix* M' *that sign-agrees (entry-by-entry) with* M.

Notice that the sign-rank of a matrix can be *much* smaller than its rank.

Theorem 2 (Forster) *Linear bounds for* $U(\text{IP})$ *by giving lower bounds on* $\text{sign-rank}(M_{\text{IP}})$*, i.e.* IP $\notin PP^{cc}$.

2 Variable partition models

In the standard model, the inputs x and y are partitioned in a *fixed* way amongst the two players.

Definition *(Best/any partition model)*
Let $f : \{0,1\}^m \mapsto \{0,1\}$. Let S, T be a partition of $\{1,\ldots,m\}$ into two equal-sized sets. Then the minimum deterministic complexity over the "best" partition is:

$$D^{\text{best}}(f) \stackrel{\text{def}}{=} \min_{S,T} D^{S,T}(f)$$

And the maximum deterministic complexity over the "worst" partition is:

$$D^{\text{worst}}(f) \stackrel{\text{def}}{=} \max_{S,T} D^{S,T}(f)$$

It is harder to obtain lower bounds for $D^{\text{best}}(f)$.

Theorem 3 (Hayes) *Fixed partition lower bounds imply best partition lower bounds for a related function.*

2.1 More about IP

Definition Let G be a d-regular bipartite graph (with adjacency matrix A) on $[n] \times [n]$.

$$\text{GRAPH-IP}_A \stackrel{\text{def}}{=} \sum_{i,j} A(i,j) x_i y_j \bmod 2$$

(This is analogous to $\text{IP}(X,Y) = \sum_i x_i y_i \bmod 2$.) M_G is the normalized version of A. λ_2 is the second-largest eigenvalue of M_G.

Theorem 4 (Alor-Boppana(?)) *For all sufficiently large n, a random d-regular bipartite graph has $\lambda_2 = \Theta(\frac{1}{\sqrt{d}})$ with high probability.*

Lemma 5 (Expander mixing lemma) *Let G be a d-regular graph on $V_1 \times V_2$ with $|V_1| = |V_2| = n$. Let $A \subseteq V_1$, $B \subseteq V_2$. Then the number of edges between A and B is $\geq \frac{d|A|\cdot|B|}{n} - \lambda_2 d\sqrt{|A|\cdot|B|}$. When $|A| = |B| > \lambda_2 n$, then there exists at least one edge between A and B.*

Theorem 6 *Let G be a bipartite d-regular graph (with adjacency matrix A) on $V_1 \cup V_2$, with $\lambda_2 = \frac{1}{\sqrt{d}}$ and $d \geq 4$. If the best-partition communication complexity of $\text{GRAPH-IP}_A(x,y) \leq f(n)$, then the fixed partition communication complexity of $\text{IP}(x,y) \leq O(f(n))$.*

Proof Let $V = V_1 \cup V_2$; let $P = P_1, P_2$ be the partition of $V_1 \cup V_2$ such that GRAPH-IP_A is easy ($\leq f(n)$) with respect to this partition. Assume WLOG that P_1 contains $V_1' \subset V_1$ of size $\geq \frac{n}{2}$, and similarly for V_2'.

GOAL: find $S \subset V_1'$, $T \subset V_2'$ such that the edges between S and T are a perfect matching, and $|S|$ and $|T| = O(n)$.

1. Use expander mixing lemma. Initially, let $A_0 = V_1'$, $B_0 = V_2'$; $|A_0| = |B_0| = \frac{n}{2}$.

2. Choose an edge $e = (i, j)$ from $A \times B$. (Such an edge exists by the expander mixing lemma.) Remove all neighboring vertices of edge: $A_1 = A_0 - \text{nbrs}(j)$ and $B_1 = B_0 - \text{nbrs}(i)$. Repeat this step on A_1 and B_1.

By the expander mixing lemma, we can continue for s steps as long as $|A_s|$ and $|B_s| \geq \lambda_n = \frac{n}{\sqrt{d}}$, i.e., $\frac{n}{2} - ds \geq \frac{n}{\sqrt{d}}$.

Consider an instance of IP where Alice has x of size s and Bob has y of size s. Alice creates x', an instance of GRAPH-IP$_A$ on n bits with partition P, as follows: the bits of x are put in one-to-one correspondence with elements of S; all other bits of x' are zero. Bob does the same, matching his y to corresponding bits of T. This is done nonuniformly; Alice and Bob share graph G ahead-of-time.

For any function that has large complexity under one partition, we construct a function GRAPH-IP that has large complexity under *any* partition.

2.2 Examples

2.2.1 Q-decision trees

Usually we think of a decision tree for a boolean function $f(\vec{x})$ as having one bit x_i of $\vec{x}$ at each node, with the edges to its two children labelled by values 0 and 1 for x_i.

In Q decisions trees, each node is instead labelled with any function from the function class Q (e.g., $x_1 \wedge x_5$, $x_3 \vee x_2$).

Let $\text{worst}_Q \stackrel{\text{def}}{=} \max_{q \in Q} D^{\text{worst}}$. Let $\text{best}_Q \stackrel{\text{def}}{=} \min_{q \in Q} D^{\text{best}}$.

Lemma 7 $T_Q(f) \geq \frac{D^{worst}(f)}{worst_Q}$

Proof Fix an arbitrary partition of the inputs into two disjoint sets. Alice and Bob simulate the tree. At every query, they compute the function $q(x)$ that is being queried at that node. Each step takes $\leq \text{worst}_Q$ bits of communication complexity. They take a total $T_Q(f)$ number of steps. Total communication complexity $= T_Q(f) \cdot \text{worst}_Q$.

2.2.2 Turing machine time-space tradeoff

Consider a multitape Turing machine, with a read-only input tape and several read-write work tapes.

Lemma 8 *Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Let M be a multitape Turing machine running in time $T(n)$ and space $S(n)$ that computes f on inputs of size $m = 3n$, such that*

- *M accepts $\{x0^n y | f(x,y) = 1\}$*
- *M rejects $\{x0^n y | f(x,y) = 0\}$*

Then $D(f) = O\left(\frac{T(n)S(n)}{n}\right)$.

In particular, if f has $O(n)$ communication complexity, then $T(n)S(n) \geq O(n^2)$.

Why?

We'll show that, given a Turing machine, we can construct a communication complexity protocol simulating it.

The input tape of M looks like: $\boxed{x \mid 0^n \mid y}$. Alice simulates M's computation until input tape head moves over to y, then sends the state and the tapes' contents to Bob (for a cost of $O(S(n))$ bits). Bob simulates M's computation until the input tape head moves over to x, then transfers the state and tapes' contents to Alice. They continue in this manner until the simulated computation has halted.

The number of transfers is at most $T(n)/n$, since the middle pad of n zeroes forces Alice and Bob to do a lot of work. (If the zeroes were not there, then we would consider the number of reversals. . .) Each transfer uses $O(S(n))$ bits, for a total deterministic protocol cost of $D(f) = O\left(\frac{T(n)S(n)}{n}\right)$.

2.2.3 Circuit depth lower bounds

This material is from Karchmov and Wigderson (sp?).

Consider a function $f : \{0,1\}^n \to \{0,1\}$ and the relation

$$R_f = \{(x, y, i) \mid x \in f^{-1}(1),\, y \in f^{-1}(0),\, x_i \neq y_i\}$$

Define $cc(R_f)$ as the minimum number of bits of communication required in a protocol in which player 1 gets $x \in f^{-1}(1)$ and player 2 gets $y \in f^{-1}(0)$. The protocol outputs an i such that $x_i = y_i$.

Theorem 9 *For any boolean function $f : \{0,1\}^n \to \{0,1\}$, the minimum depth of a circuit over $\{\wedge, \vee, \neg\}$ is equal to $cc(R_f)$.*