CS 2429 - Foundations of Communication Complexity

Lecture #7: 27 October 2009

Lecturer: Toniann Pitassi and Matei David

Scribe Notes by: Lei Huang

1 Applications of 2-Party Communication Complexity Cont'd

1.1 Circuit Depth via Communication Complexity

In order to get circuit lower bounds, we need to extend our notion of 2-party communication complexity so that it can compute relations.

Definition A relation R is a subset $R \subseteq X \times Y \times Z$

Given a relation R the cc problem associated with R follows:
Alice gets $x \in X$
Bob gets $y \in Y$
Alice and Bob must both compute (and output) some z s.t. $(x, y, z) \in R$

A protocol for relations is the same as a protocol for functions, in each step it must specifiy which party sends a message and the value of that message.

Note that for a given relation there may be more than on z satisfying the above property, Alice and Bob only need to give one such z. In general, lower bounds are harder to prove for relations as we need to show it is hard for Alice and Bob to compute *any* z.

Definition For any boolean function $f : \{0,1\}^n \to \{0,1\}$ and $X = f^{-1}(1)$, $Y = f^{-1}(0)$. We define $R_f \subseteq X \times Y \times \{1, 2, ...n\}$ to be the associated relation where,

- $R_f = \{(x, y, i) | x \in X, y \in Y, x_i \neq y_i\}$

R_f is the set of all (x, y, i) where $f(x) = 1$, $f(y) = 0$ and x and y differ on bit i. Similarly if f is monotone then

- $M_f \subseteq X \times Y \times \{1, 2, ...n\}$ is the set of all (x, y, i) such that $x \in X$, $y \in Y$ and $x_i = 1$, $y_i = 0$.

(Recall that for a monotone boolean function f, $f(x) = 1$ implies that for all x' where $x'_i \geq x_i$ on every i, x' is also a 1 of the function.)

Communication complexity lower bounds on M_f give bounds on monotone circuit depth of f and lower bounds on R_f give circuit depth bounds for general circuits.

Let $d(f)$ and $d^{monotone}(f)$ denote the min depth of a circuit computing f over $\wedge$, $\vee$, $\neg$, and the min depth of a monotone circuit computing f over $\wedge$, $\vee$ respectively. In both cases the circuits must have bounded fan-in.

Theorem 1 *(Karchmer and Widerson '80s)*

1. *For every boolean function* $f : \{0,1\}^n \to \{0,1\}$, $cc(R_f) = d(f)$
2. *For* f *monotone,* $cc(M_f) = d^{monotone}(f)$.

For formulas it is known that $2^{d(f)} = \text{formula-size}(f)$ so proving lower bounds on communication complexity of relations is also equivalent to proving formula size lower bounds.

It is a major open problem to get even super log-depth lower bounds for the general case. But for the monotone case the method above has been used to show that $NC^i_{monotone} \neq NC^{i+1}_{monotone}$ for all i [see Theorem 2 and 3].

Proof of Theorem 1 "⇒"

Let C be a circuit for f, depth$(C) = d$. We can assume that all the negations in the circuit are at the leaves. (If not, the negations can be pushed to the leaves without affecting depth in any circuit by repeated application of DeMorgan's laws.)

We want to use the circuit to obtain a protocol for R_f.

The protocol will involve Alice and Bob taking a particular path down the circuit with Alice, deciding the branch to take at OR gates and Bob deciding at AND gates. As long as the two parties maintain the invariant that at each subnode v $C_v(x) = 1$ while $C_v(y) = 0$ then the leaf reached is a bit i where $x_i \neq y_i$.

The protocol follows:

Starting from the top of the circuit, for each each node v with children v_L, v_R

if the gate is an OR Alice says 0 if $C_{v_L}(x) = 1$ and 1 otherwise.

if the gate is an AND Bob says 0 $C_{v_L}(y) = 1$ and 1 otherwise.

At the end of the exchange, both Alice and bob recurse on v_L if the message sent was 0 and v_R if the message sent was 1.

Clearly at the top of the circuit, for any inputs (x, y), $C(x) \neq C(y)$. Suppose at some point during the protocol Alice and Bob are at some inner node v where $C_v(x) \neq C_v(y)$.

Case 1 v is an or node.

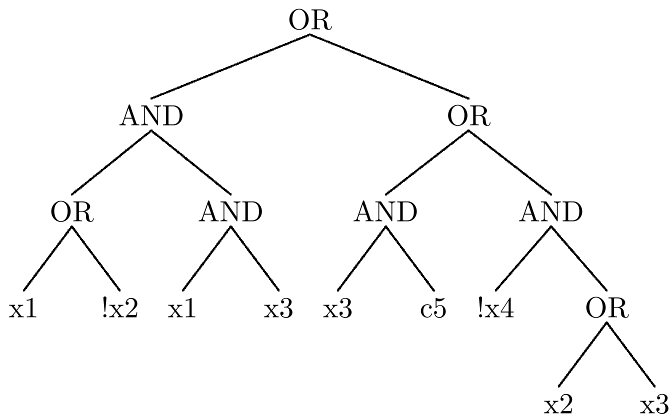
Then $C_v(y) = 0$ implies that both $C_{v_L}(y)$ and $C_{v_R}(y)$ are also 0. By choosing the subcircuit for which her input evaluates to 1, Alice ensures that the recursion continues on a subcircuit where the two inputs differ.

Case 2 v is an and node.

Likewise, $C_v(x) = 1 \Rightarrow C_{v_L}(x) = C_{v_R}(x) = 1$ so by choosing the subcircuit for which his input evaluates to 0 Bob can also maintain the above invariant.

By induction, when the protocol reaches a leaf, both A and B know an i at which their inputs differ. The total number of bits sent is bounded by the depth of the circuit. If C was monotone the same protocol reaches a left where $x_i = 1$.

Example



Suppose Alice and Bob have inputs (01101) and (01010) respectively. Then on the circuit above the sequence of bits sent would be.

Alice : 0 (go right)
Bob : 1(go left)
Alice : 0 (go left)

At which point they reach x_3 a bit on which they differ.

Proof of Theorem 1 "$\Leftarrow$"

Given a protocol for R_f we can construct a circuit computing f of bounded depth.

Consider a protocol tree T for R_f. Convert T into a circuit as follows:

1. For each node where the message is sent by Alice, replace the node with an OR gate
2. For each node where the message is sent by Bob, replace the node with an AND gate
3. At each leaf of the protocol tree, with associated monochromatic rectangle $A \times B$ and input bit i

   Claim Exactly one of the following hold

   (a) $\forall \alpha \in A$, $\alpha_i = 1$ and $\forall \beta \in B$, $\beta_i = 0$
   (b) $\forall \alpha \in A$, $\alpha_i = 0$ and $\forall \beta \in B$, $\beta_i = 1$

   Assign the leaves in case (a) to be z_i and and the leaves in case (b) to be $\bar{z}_i$.

Given the claim we can prove by induction that the circuit thus constructed calculates $f(z)$.

Proof of Claim

Let $\alpha \in A$, $\alpha_i = \sigma$. Then for every $\beta \in B$, $\beta_i = \bar{\sigma}$ which in turn implies that $\forall \alpha \in A$, $\alpha_i = \sigma$.

Theorem 2 *(KW)*

The monotone depth of st-connectivity is $\Omega(\log^2 n)$.

Theorem 2 separates monotone NC^1 from monotone NC^2. A similar lower bound proved for clique separates $monotone - P$ from $monotone - NP$.

Theorem 3 *Theorem(Raz, McKenzie)*
For every i there exists a monotone function in monotone-NC^{i+1} but not in monotone-NC^i.

2 NOF (Number on Forehead) Communication Complexity

Thus far, we have looked at 2-party communication complexity. One extention of this model to a multi-party problem is the Number on Forehead model.

In an NOF cc problem, there are k players where player i receives x_i, $|x_i| = n$. We can imagine each player wearing their input on their forehead. Thus player i can see all inputs except x_i and players communicate on a shared blackbord to compute some function $f(x_1, ...x_k)$. Note that when $k = 2$ this reduces to the 2-party model.

Intuitively, this model can be more powerful than the 2-party model since more players (k-1 to be exact) have access to each bit.

Example The multi party Equality problem $EQ_n^k(x_1, ...x_k) = 1$ iff $x_1 = ... = x_k$
In the first lecture, we showed using Fooling Sets that for $k = 2$ $D(EQ_n^2) = n + 1$.

In contrast, for any $k \geq 3$ it only takes 2 bits under the following protocol
Step 1. Player 1 sends 1 iff $x_2 = ... = x_k$
Step 2. Player 2 sends 1 iff $x_1 = x_3$.

2.1 NOF Complexity Classes

As with two party communication complexity, NOF communication complexity has the following analogs to the usual complexity classes.

$$P^{k,cc} \quad NP^{k,cc} \quad RP^{k,cc} \quad BPP^{k,cc}$$

In recent years the following two facts have been shown

1. $P^{k,cc} \not\subseteq RP^{k,cc}$
2. $BPP^{k,cc} \not\subseteq NP^{k,cc}$

2.2 Importance of NOF model (Connection to ACC)

Definition ACC is the family of unbounded fan-in circuits of constant depth and polysize over $\vee$, $\wedge$, $\neg$, MOD_m for some fixed m. ($\text{MOD}_m(x_1, ...x_n) = 1 \iff \sum_i x_i = 0 \bmod m$)

When m is a power of a prime, we know lower bounds for ACC but otherwise very little is known even for m as small as 6. We will show that lower bounds for any explicit function $f(x_1, ...x_k)$ for polylog$n$ values of k imply super polynomial lower bounds for the class ACC.

Definition SYM^+ is the family of depth 2 circuits where the top gate is a symmetric function and the bottom level consists of AND gates with fan-in $d = \text{polylog}n$. The overall size is $2^{\text{polylog}n}$.

Theorem 4 *(Yao, Beigel, Tarui)*
$ACC \subseteq SYM^+$

Lemma 5 *Let f be a boolean function computed by $C \in SYM^+$, where C has size S and bottom fan-in d.*

Then there exists a $d+1$ player NOF protocol for computing f (under any partition of the inputs) that sends $O(d \log S)$ bits.

Given the lemma, if we can prove that a function f requires super polylog computational complexity for polylog players then $f \notin SYM^+ \Rightarrow f \notin ACC$. Furthermore, finding such an $f \in P$ would imply that $P \neq ACC$.

Proof of Lemma

Each AND gate can have fan-in at most d so there must be at least one x_i in a $d+1$ partion of the input such that the AND gate does not depend on inputs from x_i by the pigeon hole principal. The i^{th} player can compute the this and gate without any communication from other players simply by looking at the bits available to him.

A priori, we can agree on a partition of the ANDs into $d+1$ groups where group j are AND functions that can be evaluated by j.

During the protocol, each party i need only send the number of AND gates in group i that evaluate to true. In fact, the evaluation of the ANDs and sending of this number can be done in parallel. Since the top gate is symmetric, this information is sufficient for each i to know the value of the circuit.

Each number sent has at most $\log S$ bits so we get the $O(d \log S)$ bound immediately.

2.3 Cylinder Intersections

The analog of combinatorial rectangles in the NOF model is the *cylinder intersection*. Each node of a protocol tree is consistent with a particular cylinder intersection. These are also the basic objects under consideration for obtaining lower bounds in the NOF model.

Definition

Let X_i be the set of all possible values for x_i (usually $\{0,1\}^n$). A cylinder in the i^{th} dimension is a subset $S_i \subset X_1 \times ... \times X_k$ where

$$(x_1, x_2, ...x_i, x_{i+1}, ...x_k) \in S_i \iff (x_1, x_2, ...x_i', x_{i+1}, ...x_k)\ \forall x_1, ...x_n, x_i'$$

In otherwords, membership in S_i does not depend on the i^{th} coordinate.

S_i looks like $(a_1, ...a_{i-1}, *, a_{i+1}, ...a_k)$ where $(a_1, ...a_{i-1}, a_{i+1}, ...a_k) \in B^{[k]/i}$.

Definition

A subset $S \subset X_i \times ... \times X_k$ is a cylinder intersection if $S = \bigcap_{i=1}^k S_i$ where S_i is a cylinder in the i^{th} dimension.

Example

The entire space $X_1 \times ... \times X_k$ is a cylinder intersection

Example

The main diagonal of a cube is a cylinder intersection. The three i^{th} dimensional cylinders are the three planes intersecting main diagonals of oposing faces of the cube.

Example

Again in a cube, the points $(1,1,1)$, $(0,1,0)$ and $(1,0,0)$ are not a cylinder intersection. These points are known as a "star". One technique of proving lower bounds is exploiting the structure of stars (see Chandra Furst Lipton '86).

Lemma 6 *Let P be a $k-$party deterministic protocol and v be a node in the protocol tree. Then R_v, the set of inputs that reach v is a cylinder intersection. In particular P partitions $X_1 \times ... \times X_k$ into 2^L (disjoint) monochromatic cylinder intersections, where $L =$number of leaves of P.*

Proof (by induction on tree height)

Suppose $R_v = S_{v_1} \cap S_{v_2} \cap ... \cap S_{v_k}$, without loss we can assume that player 1 speaks at node v. This partitions S_{v_1} into two halves, $S^1_{v_1}$ and $S^0_{v_1}$. Then then left-right children of R_v are equal to $S^1_{v_1} \cap ... \cap S_{v_k}$ and $S^0_{v_1} \cap ... \cap S_{v_k}$ both of which are cylinder intersections.

2.3.1 Discrepancy in NOF

Most NOF Lower bounds come from discrepancy. The following two Lemmas from previous lectures still hold for NOF models

1. $D^{\mu,\epsilon}(f) \geq \log \frac{1-2\epsilon}{disc_\mu(f)}$, $\epsilon < 1/2$ and $disc_\mu(f) = \max_T disc\mu(f,T)$
2. $R^\epsilon \geq D^{\mu,\epsilon}(f)\ \forall \mu$

It is useful to consider the definition of discrepancy in a different form.

Definition For a function $f : \{0,1\}^{nk} \to \{-1,1\}$, a distribtution μ on $\{0,1\}^{nk}$,and a set $T \subseteq \{0,1\}^{nk}$

$$\begin{aligned} disc_\mu(f,T) &= \left|\mu\left(f^{-1}(1) \cap T\right) - \mu\left(f^{-1}(-1) \cap T\right)\right| \\ &= \left|E_{\mathbf{x}\sim\mu}\left[f(\mathbf{x}) * 1_T(\mathbf{x})\right]\right| \end{aligned}$$

where $1_T(\mathbf{x}) = 1$ iff $\mathbf{x} \in T$.

Theorem 7 *(Babai, Nisen, Szegedy '92)*

$$E_{\mathbf{x}}\left(f(\mathbf{x}) \cdot 1_T(\mathbf{x})\right)^{2^k} \quad \leq \quad E_{\mathbf{x}^0,\mathbf{x}^1}\left[\Pi_{\mathbf{u}\in\{0,1\}^k} f(\mathbf{x}^{\mathbf{u}})\right]$$

Proof for $k = 3$

Let T be the cylinder intersection such that $disc_\mu(f,T) = disc_\mu(f)$ i.e. T witnessses the max discrepancy.

Consider $E(f(\mathbf{x})1_T(\mathbf{x}))$ for the case where $k = 3$. Writing out the $\mathbf{x}$, we get

$$E_{x_1,x_2,x_3}[f(x_1,x_2x_3) \cdot 1_T(x_1,x_2,x_3)]$$

Because T is a cylinder intersection there exists ψ_1, ψ_2, ψ_3 functions from $\{0,1\}^{2n} \to \{0,1\}$ such that

$$1_T(x_1, x_2, x_3) = \psi_1(x_2, x_3) \cdot \psi_2(x_1, x_3) \cdot \psi_3(x_1, x_2)$$

ψ_1, ψ_2, ψ_3 are characteristic functions of the basis for T. Substituting in the expectation above

$$E_{x_2,x_3}\left[\psi_1(x_2, x_3) \cdot E_{x_1}\left[f(x_1, x_2, x_3\right]\psi_2, \psi_3\right] \qquad (1)$$

From Cauchy-Schwartz we know that $E[z]^2 \leq E[z^2]$

$$\begin{aligned}
(1)^2 &\leq E_{x_2,x_3}\left[\psi_1^2(x_2, x_3)\left(E_{x_1}\left[f(x_1, x_2, x_3)\psi_2\psi_3\right]\right)^2\right] \leq \\
&\quad \text{[expanding out the square and dropping } \psi_1^2 > 0\text{]} \\
&\leq E_{x_2,x_3}\left[E_{x_1^0 x_1^1}\left(\Pi_{u_1 \in \{0,1\}} f(x_1^{u_1}, x_2, x_3)\psi_2\psi_3\right)\right] \\
&\leq E_{x_1^0 x_1^1 x_3}\left[\psi_2\psi_2^1 E_{x_2}\left[\Pi_{u_1} f \cdot \psi_3\right]\right] \\
&\quad \text{another application of Cauchy-Swartz gives} \\
(1)^4 &\leq E_{x_1^0 x_1^1 x_3}\left[E_{x_2^0 x_2^1}\left[\Pi_{u_1, u_2 \in \{0,1\}} f \cdot \psi_3\right]\right] \\
&\vdots
\end{aligned}$$

In general, each application of Cauchy-Swartz eliminates one ψ after k applications we get

$$E_{\mathbf{x}}\left(f(\mathbf{x}) \cdot 1_T(\mathbf{x})\right)^{2^k} \leq E_{\mathbf{x}^0, \mathbf{x}^1}\left[\Pi_{\mathbf{u} \in \{0,1\}^k} f(\mathbf{x}^{\mathbf{u}})\right]$$

where $\mathbf{x} = (x_1, x_2, \ldots x_k)$ and $\mathbf{x}^{\mathbf{u}} = (x_1^{u_1}, x_2^{u_2}, \ldots x_k^{u_k})$.
The product $\Pi_{\mathbf{u} \in \{0,1\}^k}$ takes the products of f over the vertices of a hyper cube.