Lecture 15: 2-Party Communication Complexity

Lecturer: *Sanjeev Arora* Scribe: *Don Sheehy*

1 Background

According to the formulation of Yao (1979), 2-party communication complexity is defined as follows. Two parties, A and B are each given a string in $\{0,1\}^n$ and a function $f : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$. A protocol is devised in which at the ith round, the ith bit is communicated between A and B. The computation is complete after both A and B know the value of $f(x,y)$. The only complexity measure we are interested in is the number of bits communicated. The individual parties may be assumed to have unlimited computational resources.

We define a protocol more formally as follows.

DEFINITION 1 *A protocol P is a family of functions $P_1, \ldots, P_r$ with the property that if $b_1, \ldots, b_i$ are the first i bits communicated then*

$$P_i : \{x, y\} \times (b_1, \ldots, b_{i-1}) \rightarrow b_i. \tag{1}$$

At the end of R rounds both A and B have enough information to compute $f(x,y)$.

For a given function f, we denote the communication complexity as

$$C(f) = \min_{protocols\ P} \max_{x,y} \{\text{Number of bits communicated by } P \text{ on } (x,y)\} \tag{2}$$

For any function f there is a trivial upperbound on the communication complexity, $C(f) \leq n+1$. This is achieved by the protocol that sends all of A's bits to B in the first n rounds and sends the value of $f(x,y)$ to A in the $(n+1)$st round.

2 Lowerbounding C(f)

The early motivation for studying communication complexity was in the analysis of distributed computation. The goal was to find good lowerbounds on the time necessary to compute functions on distributed systems. We will be looking at ways of proving lowerbounds on the communication complexity of some different functions. Note that we are looking for **coNP** type statements.

Consider $M(f)$, the $2^n \times 2^n$ matrix defined by a function $f : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$. A comunication protocol gives a method for partitioning $M(f)$ into *monochromatic rectangles*. That is we can find subsets $X, Y \in \{0,1\}^n$ where $f(x,y) = f(x',y')$ for all $x, x' \in X$ and all $y, y' \in Y$. Here, the *rectangle* refers to the set $X \times Y$. If $C(f) \leq k$ then there exists a way to partition $M(f)$ into 2^k or fewer monochromatic rectangles.

EXAMPLE 1 Equality testing.

$$f(x,y) = \begin{cases} 1 & \text{if x=y} \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Note that no monochromatic rectangle can contain two 1s. So, we see that $C(f) \geq n$.

EXAMPLE 2 The inner product function.

$$f(x,y) = \sum_{i=1}^{n} x_i y_i \pmod 2 \tag{4}$$

For our analysis of the inner product function it will be convenient to consider it as taking values in $\{-1,1\}$ by mapping $0 \to 1$ and $1 \to -1$ in the matrix of values of f. Let M denote the matrix $M(f) \in \{-1,1\}^{2^n \times 2^n}$. For any subsets $A, B \subseteq \{0,1\}^n$, let $M(A,B)$ denote $1_A M 1_B$ where 1_A and 1_B denote the characteristic vectors of A and B respectively. That is $M(A,B) = \left|\sum_{x\in A, y\in B} f(x,y)\right|$.

THEOREM 1
For all $A, B \in \{0,1\}^n$,

$$M(A,B) \leq 2^{\frac{n}{2}} \sqrt{|A|\,|B|} \leq 2^{\frac{3}{2}} \tag{5}$$

PROOF: Observe that for all $x \in \{0,1\}^n$,

$$|M^T x|_2 \leq \lambda_{max} ||x||_2, \tag{6}$$

where λ_{max} is th elargest eigenvalue of M. All rows of M are orthogonal and have length $2^{\frac{n}{2}}$. Thus, $MM^T = 2^n I$ and so all eigenvalues of M are $\pm 2^{\frac{n}{2}}$. So it follows that

$$1_A M 1_B \leq |\lambda_{max}| |1_A|_2 |1_B|_2 \leq 2^{\frac{n}{2}} \sqrt{|A||B|}. \tag{7}$$

□

So, the theorem implies that every monochromatic rectangle has at least $2^{\frac{n}{2}}$ rectangles and thus $C(f) \geq \frac{n}{2}$.

3 Discrepancy

DEFINITION 2 *The* Discrepancy *of a function* f *is denoted by* $Disc(f)$ *and is defined as*

$$Disc(f) = \max_{A,B\subseteq\{0,1\}^n} \frac{1}{2^{2n}} \sum_{a\in A b\in B} f(a,b) \tag{8}$$

The discrepancy is obviously an upperbound on the size of the largest monocromatic rectangle in $M(f)$, so it follows that $Disc(f)$ gives an immediate lowerbound on $C(f)$.

$$C(f) \geq \log_2 \left(\frac{1}{Disc(f)}\right). \tag{9}$$

To prove lowerbounds on $C(f)$, it suffices to upperbound $Disc(f)$. In general, the problem of upperbounding the discrepancy is a **coNP**-Complete problem.

We restrict our attention to 2×2 rectangles, $\{a_1, a_2\}, \{b_1, b_2\}$. Recall that we are considering functions of the form $f : \{0,1\}^n \times \{0,1\}^n \to \{-1,1\}$. Define the following rectangle product.

$$\Pi_{f,(a_1,a_2),(b_1,b_2)} = \prod_{i=1}^{2} \prod_{j=1}^{2} f(a_i, b_j) \tag{10}$$

We define the average $\mathcal{E}(f)$ of this product over all rectangles.

$$\mathcal{E}(f) = \mathbf{E}_{(a_1,a_2),(b_1,b_2)} \left[\Pi_{f,(a_1,a_2),(b_1,b_2)}\right] \tag{11}$$

THEOREM 2
$|\mathcal{E}(f)| \geq (Disc(f))^4$.

PROOF: The proof will have two main steps.

1. For all functions h, $\mathcal{E}(h) \geq (\mathbf{E}_{a,b}[h(a,b)])^4$.
2. There exists h such that $\mathbf{E}_{a,b}[h(a,b)] \geq Disc(f)$ and $\mathcal{E}(f) = \mathcal{E}(h)$.

Together, these prove the Theorem.
Proof of step 1.

$$\begin{aligned}
(\mathbf{E}_x[g(x)])^2 &= \mathbf{E}_{x_1,x_2}[g(x_1)g(x_2)] && (12)\\
\mathbf{E}[z^2] &\geq (\mathbf{E}[z])^2 \text{ (by Cauchy-Schwartz)} && (13)\\
\mathcal{E}(h) &= \mathbf{E}_{a_1,a_2}\left(\mathbf{E}_{b_1,b_2}[\Pi_{h,(a_1,a_2),(b_1,b_2)}]\right) && (14)\\
&= \mathbf{E}_{a_1,a_2}\left(\mathbf{E}_{b_1,b_2}[h(a_1,b_1)h(a_1,b_2)h(a_2,b_1)h(a_2,b_2)]\right) && (15)\\
&= \mathbf{E}_{a_1,a_2}\left[(\mathbf{E}_b[h(a_1,b)h(a_2,b)])^2\right] && (16)\\
&\geq (\mathbf{E}_{a_1,a_2}\mathbf{E}_b[h(a_1,b)h(a_2,b)])^2 && (17)\\
&= \left(\mathbf{E}_b(\mathbf{E}_a[h(a,b)]^2)\right)^2 && (18)\\
&\geq (\mathbf{E}_{a,b}[h(a,b)])^4 && (19)
\end{aligned}$$

Proof of step 2. Let $A \times B$ be the rectangle for which $Disc(f)$ is attained. We prove the existence of h by the probabilistic method. Define the following two functions.

$$g_1(a,b) = \begin{cases} 1 & \text{if } a \in A \\ \text{set randomly to } \pm 1 & \text{otherwise} \end{cases} \tag{20}$$

$$g_2(a,b) = \begin{cases} 1 & \text{if } b \in B \\ \text{set randomly to } \pm 1 & \text{otherwise} \end{cases} \tag{21}$$

These functions have the property that they depend only on the rows or the columns respectively. That is, $g_1(a,b) = g_1(a,b')$ and $g_2(a,b) = g_2(a',b)$ for all a, b, a', b'. Let $h = fg_1g_2$. That is

$$h(a,b) = f(a,b)g_1(a,b)g_2(a,b). \tag{22}$$

We can now average over the choices for g_1, g_2 and also the pair (a,b) to see that

$$\mathbf{E}_{g_1,g_2}\mathbf{E}_{a,b}[h(a,b)] = Disc(f). \tag{23}$$

So there exists some choice of g_1, g_2 such that $\mathbf{E}_{a,b}[h(a,b)] \geq Disc(f)$. Finally, $\mathcal{E}(f) = \mathcal{E}(h)$ for all such g_1, g_2 because g_1 is constant on the rows and g_2 is constant on the columns so the products cancel to 1.

□