COMS 4995: Introduction to Communication Complexity — Sep 7, 2011

Lecture 1 – Deterministic Communication Complexity

Instructor: *Xi Chen* — Scribes: *Xiaorui Sun*

1 Two Party Communication Complexity Model

The model assumes following facts:

1. There are two players, Alice and Bob. Alice has an input $x \in X$, and Bob has an input $y \in Y$. At the beginning, Alice has no information about y and Bob has no information about x.

2. The goal of Alice and Bob is to compute $f(x, y)$ by communication, where the function f is a predefined function $f : X \times Y \to Z$. Alice and Bob should follow a protocol to communicate.

3. Alice and Bob want to minimize their communications. There is no computational limits for Alice and Bob.

Definition 1 (Protocol)**.** *A protocol P to compute function $f : X \times Y \to Z$ is a binary tree satisfies following properties:*

- *Each non-leaf node v corresponds to a function of form $a_v : X \to \{0, 1\}$ if Alice sends a bit or $b_v : Y \to \{0, 1\}$ if Bob sends a bit.*

- *Each leaf corresponds to a value $z \in Z$.*

- *For any $(x, y) \in X \times Y$, the protocol performs a walk from the root to a leaf. At each step, the walk goes to left if $a_v(x) = 0$(or $b_v(y) = 0$ if Bob sends a bit), or right if $a_v(x) = 1$(or $b_v(y) = 1$ if Bob sends a bit). If the walk goes to a leaf, then the protocol computes z as the output of the communication, where z is the value corresponding to the leaf.*

- *The communication complexity of the protocol is the height of the tree.*

Example 2 (OR)**.** *Assume $X = Y = \{0, 1\}^n$. Alice and Bob wish to compute $x_1 \wedge x_2 \wedge \cdots \wedge x_n \wedge y_1 \wedge y_2 \wedge \cdots \wedge y_n$ for input $x \in X$ and $y \in Y$.*

The computation of OR function needs only two bits: Alice sends $x_1 \wedge x_2 \wedge \cdots \wedge x_n$ to Bob, then Bob sends $x_1 \wedge x_2 \wedge \cdots \wedge x_n \wedge y_1 \wedge y_2 \wedge \cdots \wedge y_n$ back to Alice.

Example 3 (Equality)**.** *Assume $X = Y = \{0, 1\}^n$. Alice and Bob wish to compute whether $x = y$ for input $x \in X$ and $y \in Y$, a.e. $f(x, y) = 1$ if and only if $x = y$.*

A naive protocol is as follow: Alice sends n-bit x to Bob, and Bob checks whether $x = y$ then sends the result back to Alice. This protocol requires $n + 1$-bit communication.

2 Lower Bound Technique: Rectangle

In this section, we use rectangle technique to show the lower bound of the communication complexity of the equality problem.

Theorem 4. *Any protocol computing the equality problem requires at least n bits communication.*

To show Theorem 4, we first define rectangle.

Definition 5. *A rectangle $R \subseteq X \times Y$ is a subset which can be expressed as $R = A \times B$ for sets $A \subseteq X$ and $B \subseteq Y$.*

Lemma 6. *For every protocol P and a node v in the protocol tree, let $R_v \subseteq X \times Y$ denote the set containing all the $(x, y) \in X \times Y$ which reach the node v. Then R_v is a rectangle.*

Proof. We prove this lemma by induction. First the set corresponds to the root is $X \times Y$, which is a rectangle.

Consider a node v and its parent t. Without loss of generality, we assume v is t's left child, and Alice sends a bit at t. By induction, R_t is a rectangle implies R_t can be expressed as $R_t = X_t \times Y_t$ for some $X_t \subseteq X$ and $Y_t \subseteq Y$. Then $R_v = \{(x, y)|(x, y) \in R_t \text{ and } a_t(x) = 0\} = [X_t \cap \{x : a_t(x) = 0\}] \times Y_t$, which is also a rectangle. □

Lemma 7. *A protocol of height h can partition $X \times Y$ into at most 2^h rectangles.*

Above lemma can be easily proved using the fact that a binary tree of height h has at most 2^h leaves. Now we are ready to prove Lemma 4.

Proof of Theorem 4. To prove Theorem 4, we only need to prove following argument: Let $S_1, S_2, \ldots S_k$ be k rectangles such that

1. $S_1 \cap S_2 \cap \cdots \cap S_k = \emptyset$,
2. $S_1 \cup S_2 \cup \cdots \cup S_k = \{0, 1\}^n \times \{0, 1\}^n$,
3. For each S_i, there exists a $z \in \{0, 1\}$ such that $f(x, y) = z$ for all $(x, y) \in S_i$, where f is the equality function.

Then $k \geq 2^n$.

Now we prove above argument. For any protocol and any $x \in \{0, 1\}^n$, we consider the leaf contains (x, x). By Lemma 6, the protocol defines a set of rectangles $S_1, S_2, \ldots, S_k$ which satisfy above three properties. It is possible to find a rectangle S_i such that $(x, x) \in S_i$.

Assume there exits a $(x', y') \in S_i$ satisfying $(x', y') \neq (x, x)$. If $x' \neq x$, then (x', x) is also an element in S_i by the definition of rectangle. But we have $f(x, x) = 1$ and $f(x', x) = 0$ because $x' \neq x$. This contradicts with the third property of $S_1, S_2, \ldots, S_k$. Similarly, if $y' \neq y$, then $(x, y') \in S_i$ and $f(x, y') = 0$. This means $S_i = \{(x, x)\}$. Since x can be any value within $\{0, 1\}^n$, there are at least 2^n rectangles for any protocol. □