CS 59000 CTT Current Topics in Theoretical CS	Nov 1, 2012
Lecture 22	
Lecturer: Elena Grigorescu	*Scribe: Jeffrey, Gaowtham, Venkata*

Introduction

In this lecture we'll introduce some basic concepts from communication complexity (CC), which is another model of sublinear computation with numerous applications in practice and theory. Let $f : X \times Y \to \{0, 1\}$. Let A and B be two parties who wish to compute the value of the function $f(x, y)$ where A has the input x, an n-bit binary string ($x \in \{0, 1\}^n$), and B has $y \in \{0, 1\}^n$. Their goal is to minimize the number of bits communicated between them while computing $f(x, y)$.
There are different models of CC: deterministic, nondeterministic, randomized (with public coins, or private coins), quantum. In this lecture we'll only focus on deterministic protocols.
We start with the simple example of computing the parity function defined as $par(x, y) = 0$ if the number of bits set to 0 in (x, y) is even, and 1 otherwise. A simple protocol for computing $par(x, y)$ is that A sends the parity of x to B and B sends the parity of y to A. So, both can then compute $par(x, y) = par(x) + par(y)$ by communicating 2 bits. Notice that there is no protocol with 1 bit of communication since both A and B have to speak.

Communication complexity protocols

Suppose $x \in X$ and $y \in Y$, where X, Y are finite discrete domains. A trivial protocol for any communication problem is that A sends x to B, B computes the value of $f(x, y)$, B send the value of $f(x, y)$ to A. Thus the number of bits communicated $= \min(log(|X|), log(|Y|)) + 1$
This gives a trivial upper bound on the complexity of computing any function $f : X \times Y \to \{0, 1\}$: i.e. $\min(log(|X|), log(|Y|)) + 1$.
We will now formalize the notion of a protocol for a function $f : X \times Y \to \{0, 1\}$ and its cost. A protocol can be represented by a rooted binary decision tree. On input (x, y) the transcript of the exchange of bits between the parties can be represented as the path from the root to one of the leaves whose label should be the value $f(x, y)$. Each internal node v corresponds to a function $a_v : X \to \{0, 1\}$ if it is A's turn to speak or $b_v : Y \to \{0, 1\}$ if it is B's turn to speak. If at an A's node v $a_v(x) = 0$ the protocol proceeds by following the left edge out from v, otherwise it follows the right edge. Similarly, if at a B's node u we have $b_u(y) = 0$ the protocol proceeds by taking the left edge out of v otherwise it takes the right one.
A protocol is correct if $\forall(x, y) \in X \times Y$, the transcript of the protocol on (x, y) leads to leaf with value $f(x, y)$.

Definition 1 *Cost of a protocol is defined as the height of the protocol tree.*

Definition 2 *The deterministic communication complexity of function f (denoted $D(f)$) is the $min_{\forall protocols}$(cost of protocol)*

The main focus in CC is proving lower bounds for $D(f)$ for fixed functions or for functions coming from specific families. Such lower bounds are in some sense information theoretic lower bounds. Proving lower bounds in this model has provided lower bounds to other related areas, primarily for streaming algorithms, data structures, complexity theory (proof complexity), and lately property testing. The CC abstraction allows the use of technique including linear algebra, Fourier analysis and information theory. We will see in this lecture how linear algebraic methods could be useful here.

1 The Communication matrix

The function f can be represented by a matrix with rows labelled with distinct x values and columns labelled with distinct y value. The entry in the matrix of f corresponding to a given x and y value, is the value of the $f(x, y)$.
In a communication protocol both parties A and B know the protocol (i.e. know the "meaning" of each bit sent). One can think of a protocol as a sequential way of selecting possible valid inputs from the communication matrix. Suppose A starts by saying 1; then B knows the set of inputs for which A would have said 1 and every other input will never be considered again. So each bit communicated partitions the communication matrix into "rectangles" and the players seek to find a rectangle of valid inputs where there is no doubt about the value of $f(x, y)$. That is, the value of $f(x, y)$ is determined when the valid rectangle that they arrived at, say on inputs from $X' \times Y'$ (which much contain (x, y)), is s.t. f is constant on $X' \times Y'$. This brings us to some useful definitions.

Definition 3 *$C \subset X \times Y$ is a* combinatorial rectangle *if $C = X' \times Y'$ where $X' \subset X$ and $Y' \subset Y$.*

Definition 4 *C is called* monochromatic *with respect to a function f, if f takes the same value for all (x, y) in C, i.e $f(x, y) = f(x', y')$ where, $\forall (x, y), (x', y') \in C$*

The number of monocromatic rectangles gives us a lower bound on $D(f)$.

Definition 5 *The partition number of f, denoted $C^D(f)$, is the minimum number of monochromatic combinatorial rectangles that can partition the matrix of f.*

Theorem 6 *Let $f : X \times Y \to 0, 1$, then, $D(f) \geq \log(C^D(f))$.*

Proof
We will reason about properties of the protocol tree. Recall that the cost of the protocol is the height of its protocol tree.

Claim 7 *If transcript of (x, y) and (x', y') end up at leaf l, then $f(x, y) = f(x', y')$.*

Proof This can be inferred from the correctness of the deterministic protocol. ■

Let $C = D(f)$ be the height of the best protocol tree and C_l be defined as the set of (x, y) whose communication transcript end at leaf l in that protocol. Therefore, $\#C_l \leq 2^C$.
Then we claim the following:

Claim 8 *C_l is a combinatorial rectangle i.e. if $(x, y) \in C_l$ and $(x', y') \in C_l$ then $(x, y') \in C_l$ and $(x', y) \in C_l$ and $(x, y') \in C_l$.*

Proof It is sufficient to prove that if $(x, y) \in C_l$ and $(x', y') \in C_l$, then $(x, y') \in C_l$ (the rest follows by the same argument).
We prove this by induction on the transcript at each level. First note that since there is only one path from the root to the leaf l it must be that the trasncript for (x, y) is the same as the transcript for (x', y'). Let $C_l^{(i)}$ be the set of all (x, y) such that the first i characters of (x, y)'s transcript agree with the transcript associated with leaf l. All transcripts are empty at the root, so $C_l^{(0)}$ contains all $(x, y) \in X \times Y$ and it is trivially true that $(x, y), (x', y') \in C_l^{(0)} \Rightarrow (x, y') \in C_l^{(0)}$. Suppose for induction that for some j, $0 \leq j <$ depth of l, the desired implication holds for each $C_l^{(i)}$, $0 \leq i \leq j$; and take $(x, y), (x', y') \in C_l^{(j+1)}$. Then by our induction hypothesis we have $(x, y') \in C_l^{(j)}$. If at this query-level A is sending the data, then because the first argument (x, y') is the same as the fisrt argument of (x, y) the next transcript-values for (x, y') will coincide with that of (x, y) (recall that at this node v the protocol proceeds according to the function $a_v : X \to \{0, 1\}$ which does not depend on y.), so we have $(x, y') \in C_l^{(j+1)}$. If B is sending the data then the protocol proceeds according to the function $b_v : Y \to \{0, 1\}$ and so the next transcript of (x, y') is the same as that for (x', y'). By induction, then, $(x, y), (x', y') \in C_l^{(\text{depth of } l)} \Rightarrow (x, y') \in C_l^{(\text{depth of } l)}$, which is exactly what we wanted. ■

Therefore, from the above theorem, we can conclude that $\#C_l = C^D(f)$ and so $\#C_l \leq 2^C \Rightarrow C = D(f) \geq \log C^D(f)$ ■

To algorithmically compute $D(f)$ is NP hard. Also, computing a bound for $D(f)$ using the number of combinatorial rectangles is difficult in general mainly because of a lack of combinatorial techniques to approach the problem.
There are however some more algebraic techniques that we will describe next.

2 Lower bounds using the communication matrix

Let $f : X \times Y \to \{0, 1\}$. Define, $M_f(x, y) = f(x, y)$ and let $\overline{M_f}(x, y) = \neg f(x, y)$. The following theorem by Melhorn and Schmidt sets an important lower bound on the value of $D(f)$.

Theorem 9 *$D(f) \geq \log(rank(M_f) + rank(\overline{M_f}))$.*

Proof
Let $D(f) = C$, so by theorem 6 we can partition M_f in at most 2^C monochromatic rectangles. Let O be the number of monochromatic rectangles of value 1. Let Z be the number of monochromatic rectangles of value 0. Therefore, $O + Z \leq 2^C$, or, $\log(O + Z) \leq C$. Let $R_1, R_2, \cdots, R_O$ be the monochromatic rectangles of value 1. Therefore, $rank(R_i) = 1 \; for \; i = 1, 2, \cdots O$ (since there is only one linearly independent rows/column in these matrices.) So $M_f = R_1 + R_2 + \cdots + R_O$. Recall that for any matrices A, B: $rank(A + B) \leq rank(A) + rank(B)$. Therefore, $rank(M_f) \leq \sum_{i=1}^{O} rank(R_i) \Rightarrow rank(M_f) \leq O$. Similarly, $rank(\overline{M_f}) \leq \sum_{i=1}^{Z} rank(\overline{R_i}) \Rightarrow rank(\overline{M_f}) \leq Z$.

From the above two inequalities, we can conclude that $rank(M_f) + rank(\overline{M_f}) \leq O + Z \Rightarrow rank(M_f) + rank(\overline{M_f}) \leq 2^C$. ■

We next show a few examples of applications of this theorem.

Example 1: The Equality function: $EQ(x,y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{otherwise} \end{cases}$. Note that the communication matrix is $EQ(x,y) = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$. So $rank(M_f) = 2^n$ and $rank(\overline{M_f}) = 1$.

Therefore, $D(f) \geq \log(2^n + 1) \Rightarrow D(f) \geq n + 1$ and so the trivial communication protocol is optimal.

Example 2: Set Disjointness $Disj(x,y) = \begin{cases} 1, & \text{if } x \cap y = \phi \\ 0, & \text{otherwise} \end{cases}$. In particular, if $x, y \in \{0,1\}$ then $Disj_1(x,y) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and so $rank(Disj_1) = 2$. One can check that if $x, y \in \{0,1\}^2$ then the communication matrix is $Disj_2(x,y) = Disj_1^{\otimes 2}(x,y)$, and more generally, if $x, y \in \{0,1\}^n$ then $Disj_n(x,y) = Disj_1^{\otimes n}(x,y)$.

Therefore, $rank(Disj_n(x,y)) = 2^n$ and as before $D(f) \geq n + 1$, saying that here too the trivial protocol is optimal.