IAS/Park City Mathematics Series
Volume 3, 1993

Circuit and Communication Complexity

Ran Raz

LECTURE 5
Lower bounds for probabilistic communication complexity

In this lecture, we will discuss certain lower bounds for the probabilistic communication complexity of functions. In particular, we will prove a linear lower bound for the probabilistic communication complexity of the Inner Product function, and we will discuss a linear lower bound for the probabilistic communication complexity of the Set Disjointness function.

5.1. Probabilistic Protocols and Unbalanced Rectangles

Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ be a function, and let A be the corresponding input matrix of zeros and ones. Assume that $CC_\epsilon(A) = k$ and let P be a probabilistic protocol for A, with error probability ϵ and communication complexity k. The protocol P depends on a public random string, denoted by R. Denote by $P_R(x,y)$ the output of the protocol P on the input (x,y), when using the random string R.

By the definition, we know that for every (x,y),

$$\Pr_R[P_R(x,y) = A_{x,y}] \geq 1-\epsilon,$$

(where the probability is over the random string R). Hence,

$$\Pr_{(x,y),R}[P_R(x,y) = A_{x,y}] \geq 1-\epsilon,$$

where (x,y) is a uniformly chosen input (note that the same holds for input (x,y) chosen according to any other probability distribution). By changing the order of

[1]Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel.
E-mail address: `ranraz@wisdom.weizmann.ac.il`.

November 15, 2000

Based on lecture notes taken by Amiel Ferman, Noam Sadot and Yoad Lustig.

the probabilities, we obtain:

$$\Pr_{R,(x,y)}[P_R(x,y) = A_{x,y}] \geq 1 - \epsilon.$$

Therefore, there exists a fixed random string r, such that, the deterministic protocol P_r (obtained by fixing the random string of P to be r) satisfies:

$$\Pr_{(x,y)}[P_r(x,y) = A_{x,y}] \geq 1 - \epsilon,$$

where as before the input (x, y) is chosen at random according to the uniform distribution. Thus, we obtained a deterministic protocol P_r that works well in the average case. That is, the deterministic protocol P_r solves the problem with high probability over the inputs (but not necessarily for all the possible inputs). Any lower bound for the communication complexity of the (deterministic) protocol P_r is also a lower bound for the communication complexity of the original (probabilistic) protocol P.

The protocol P_r is a deterministic protocol with communication complexity k. Hence, P_r defines a partition of the input matrix A into at most 2^k rectangles (as described in the previous lecture). In each of these rectangles, the output $P_r(x, y)$ is fixed. Since the output $P_r(x, y)$ is not always correct (i.e., it is not always equal to $A_{x,y}$), these rectangles are not necessarily monochromatic. Nevertheless, since the output tends to be correct, the rectangles tend to be unbalanced, in the sense that in many of these rectangles almost all the entries have the same value. In particular, we can prove the existence of at least one rectangle which is both, large and unbalanced. The following lemma shows the existence of a rectangle $U \times V$ of fraction at least $2^{-(k+1)}$ (in $\{0,1\}^n \times \{0,1\}^n$) and such that either at least $(1 - 2\epsilon)$ fraction of the entries in $U \times V$ are zeros or at least $(1 - 2\epsilon)$ fraction of the entries in $U \times V$ are ones.

Lemma 5.1.1. Assume that $CC_\epsilon(A) = k$. Then, there exists a rectangle $U \times V \subset \{0,1\}^n \times \{0,1\}^n$, such that:

(1) $|U \times V| \geq 2^{-(k+1)} \cdot 2^{2n}$.
(2) Either: at least $(1 - 2\epsilon)$ fraction of the entries in $U \times V$ are zeros, Or: at east $(1 - 2\epsilon)$ fraction of the entries in $U \times V$ are ones.

Proof. We say that a rectangle $U \times V \subset \{0,1\}^n \times \{0,1\}^n$ is large if

$$|U \times V| \geq 2^{-(k+1)} \cdot 2^{2n}$$

(i.e., if the rectangle is of fraction larger than $2^{-(k+1)}$). Otherwise, the rectangle is called small. Recall that the protocol P_r defines a partition of the input matrix A into at most 2^k rectangles. Since the total number of rectangles is 2^k and since each small rectangle is of fraction smaller than $2^{-(k+1)}$, we know that more than half of the entries of A are in large rectangles in the partition.

Recall that

$$\Pr_{(x,y)}[P_r(x,y) = A_{x,y}] \geq 1 - \epsilon.$$

In other words, at most ϵ fraction of the entries of A do not satisfy $P_r(x, y) = A_{x,y}$. Even if all these entries are in large rectangles, their fraction in the set of all entries of large rectangles is at most 2ϵ (since more than half of the entries are in large rectangles). Hence, there exists a large rectangle $U \times V$, such that for at least $(1-2\epsilon)$ of the entries in $U \times V$ we have $P_r(x, y) = A_{x,y}$. Recall that for all $(x, y) \in U \times V$ the output $P_r(x, y)$ is the same. Hence, either at least $(1-2\epsilon)$ fraction of the entries in $U \times V$ are zeros, or at east $(1-2\epsilon)$ fraction of the entries in $U \times V$ are ones. ■

5.2. Lower Bound for Inner-Product

Recall that in the Inner Product problem (IP), Player 1 gets an n-bits string x and Player 2 gets an n-bits string y. The goal of the two players is to compute

$$\sum_{i=1}^{n} x_i y_i \bmod 2.$$

Obviously $CC(IP) = O(n)$, since the trivial protocol has communication complexity $n+1$. Using methods from the previous lecture (e.g., the rank lower bound), it is not hard to prove a matching lower bound of n for the deterministic case. That is, no deterministic protocol computes IP with the exchange of less than n bits. We will prove that the same holds also for probabilistic protocols. For simplicity, we state the next theorem for constant probability of error. The same holds also for exponentially small.

Theorem 1. For any constant $\epsilon > 0$,

$$CC_\epsilon(IP) = \Omega(n).$$

Denote by A the $2^n \times 2^n$ input matrix for the Inner Product problem. That is,

$$A = \Big\{ x \cdot y \bmod 2 \Big\}_{(x,y)\in\{0,1\}^n\times\{0,1\}^n}$$

where $x \cdot y \bmod 2$ denotes the inner product of x and y modulo 2. The matrix A is a matrix of zeros and ones. It will be convenient to switch 1 to -1 and 0 to 1, and obtain a new matrix H of ones and minus-ones. The matrix H is the Hadamard matrix of order n (for simplicity we omit the n from the notation). H is given by

$$H = \Big\{ (-1)^{x\cdot y \bmod 2} \Big\}_{(x,y)\in\{0,1\}^n\times\{0,1\}^n}$$

where (as before) $x \cdot y \bmod 2$ denotes the inner product of x and y modulo 2. We will prove that the matrix H satisfies the following property.

Lemma 5.2.1. For every rectangle $U \times V \subset \{0,1\}^n \times \{0,1\}^n$,

$$\left| \sum_{(x,y)\in U\times V} H_{x,y} \right| \leq 2^{1.5\cdot n}.$$

Theorem 1 follows immediately from Lemma 5.2.1 and Lemma 5.1.1 as follows. By Lemma 5.2.1, the matrix H doesn't have a rectangle which is both, large and unbalanced. At the other hand, by Lemma 5.1.1 if the probabilistic communication complexity of H is small then such a rectangle does exist. More specifically, by Lemma 5.2.1, there is no rectangle $U \times V$ of size at least $2^{1.5\cdot n+1}$ and such that at least 3/4 of the entries in $U \times V$ have the same value (either 1 or -1). Hence, by Lemma 5.1.1 there is no probabilistic protocol for IP, with error probability

$\epsilon = 1/8$ and communication complexity smaller than $n/2 - 1$. (The bound for other values of ϵ is proven in the same way).

Before proving Lemma 5.2.1, let us introduce some notations. For two functions g, h from $\{0,1\}^n$ to the reals, denote by $\langle g, h \rangle$ their scalar product over the reals. That is,

$$\langle g, h \rangle = \sum_{y \in \{0,1\}^n} g(y)h(y).$$

For a subset $U \subset \{0,1\}^n$, denote by I_U the characteristic function of U. That is,

$$I_U(x) = \begin{cases} 1 & \text{if } x \in U \\ 0 & \text{if } x \notin U \end{cases}$$

We think of I_U as a function from $\{0,1\}^n$ to the reals. Denote by e_i the i^{th} column of the matrix H (for $i = 0, ..., 2^n - 1$). We think of each e_i as a function from $\{0,1\}^n$ to the reals. For the proof of Lemma 5.2.1, we will need the well known fact that the matrix H is an orthogonal matrix over the reals, that is, every two different columns (or two different rows) of H are orthogonal over the reals.

Claim 5.2.2. H is an orthogonal matrix over the reals.

Proof. It is enough to show that every two different rows of the matrix H are orthogonal. The scalar product (over the reals) of row x and row z is

$$\sum_{y \in \{0,1\}^n} (-1)^{x \cdot y \bmod 2} \cdot (-1)^{z \cdot y \bmod 2} = \sum_{y \in \{0,1\}^n} (-1)^{(x \oplus z) \cdot y \bmod 2},$$

where $x \oplus z$ is the bitwise XOR of the n-bits strings x and z. If $x \neq z$ then $(x \oplus z) \cdot y \bmod 2$ gets the value 0 for exactly half of the values of y and gets the value 1 for exactly half of the values of y. Hence, the above sum is 0. ■

By the previous claim and by the fact that every entry of H is either 1 or -1, we have the following two properties. These will be the only two properties needed for the proof of Lemma 5.2.1.

(1) for $i \neq j$, we have $\langle e_i, e_j \rangle = 0$.
(2) for any i, we have $\langle e_i, e_i \rangle = 2^n$.

That is, $\{e_i\}$ is an orthogonal basis for the space of all functions from $\{0,1\}^n$ to the reals. To obtain an orthonormal basis, we need to normalize the vectors e_i. For every i, we define

$$\hat{e}_i = \frac{1}{\sqrt{2^n}} e_i.$$

We can now rewrite the above two properties as:

(1) for $i \neq j$, we have $\langle \hat{e}_i, \hat{e}_j \rangle = 0$.
(2) for any i, we have $\langle \hat{e}_i, \hat{e}_i \rangle = 1$.

That is, $\{\hat{e}_i\}$ is an orthonormal basis for the space of all functions from $\{0,1\}^n$ to the reals. Hence, for every function g from $\{0,1\}^n$ to the reals,

$$\langle g, g \rangle = \sum_{i \in \{0,1\}^n} \langle g, \hat{e}_i \rangle^2,$$

(since the L_2 norm doesn't change when switching to a different orthonormal basis).

Let us now present the proof of Lemma 5.2.1. Let $U \times V$ be a rectangle. We can bound

$$\left| \sum_{(x,y)\in U\times V} H_{x,y} \right| = \left| \sum_{i\in V} \langle I_U, e_i \rangle \right| \leq \sum_{i\in\{0,1\}^n} \left| \langle I_U, e_i \rangle \right|.$$

By the Cauchy-Schwartz inequality, the last expression is at most

$$\sqrt{2^n \cdot \sum_{i\in\{0,1\}^n} \left| \langle I_U, e_i \rangle \right|^2} = \sqrt{2^n \cdot 2^n \cdot \sum_{i\in\{0,1\}^n} \langle I_U, \hat{e}_i \rangle^2} =$$

$$= 2^n \cdot \sqrt{\sum_{i\in\{0,1\}^n} \langle I_U, \hat{e}_i \rangle^2} = 2^n \cdot \sqrt{\langle I_U, I_U \rangle} \leq 2^n \cdot \sqrt{2^n}.$$

5.3. Lower Bound for Set Disjointness

The Set Disjointness problem ($Disj$) can be redefined as follows. Player 1 gets an n-bits string x. Player 2 gets an n-bits string y. The goal is to find whether the intersection of x and y is non-empty, that is, whether there exists an i, such that, $x_i = y_i = 1$. The output should be 1 iff there exists such an i. Obviously, $CC(Disj) = O(n)$, as the trivial protocol solves the problem with the exchange of $n+1$ bits. Using methods from the previous lecture (e.g., the rank lower bound), it is not hard to prove a matching lower bound of n for the deterministic case. For the probabilistic case, a lower bound of $\Omega(n)$ was proved by Kalyanasundaram and Schnitger (for any constant probability of error ϵ). This lower bound was found out to be very useful. We will not give here the proof for the lower bound. Nevertheless, we will give some idea of how the proof goes.

Unlike the Hadamard matrix (the input matrix for IP), the input matrix of the Set Disjointness problem is very unbalanced. Most of the entries are 1 (as most sets intersect). Moreover, there are huge rectangles of ones (e.g., all pairs (x, y), s.t., $x_1 = y_1 = 1$). However, there are no big rectangles of zeros.

Denote,

$$T_0 = \{(x, y) |\ |x \cap y| = 0\},$$
$$T_1 = \{(x, y) |\ |x \cap y| = 1\}.$$

That is, T_0 is the set of non-intersecting pairs, and T_1 is the set of pairs that intersect in exactly one element. Note that for every $(x, y) \in T_0$ we have $Disj(x, y) = 0$, and for every $(x, y) \in T_1$ we have $Disj(x, y) = 1$. The lower bound is based on the fact that every (large enough) rectangle that contains many elements of T_0 also contains many elements of T_1. To formalize this statement, we define for any rectangle $U \times V \subset \{0,1\}^n \times \{0,1\}^n$,

$$\alpha_0(U \times V) = \frac{|T_0 \cap U \times V|}{|T_0|},$$

$$\alpha_1(U \times V) = \frac{|T_1 \cap U \times V|}{|T_1|}.$$

That is, $\alpha_0(U \times V)$ is the fraction of T_0 covered by the rectangle $U \times V$, and $\alpha_1(U \times V)$ is the fraction of T_1 covered by the rectangle $U \times V$. The main lemma needed for the lower bound is the following lemma. Roughly speaking, the lemma states that unless $\alpha_0(U \times V)$ is negligible, $\alpha_1(U \times V)$ is not much smaller than $\alpha_0(U \times V)$. We will not give here the proof of the lemma.

Lemma 5.3.1. For every rectangle $U \times V$, if

$$\alpha_0(U \times V) > 2^{-n/100}$$

then

$$\alpha_1(U \times V) > \frac{1}{100}\alpha_0(U \times V)$$

(for large enough n).

The lemma implies that for every rectangle $U \times V$,

$$\alpha_1(U \times V) > \frac{1}{100} \cdot \alpha_0(U \times V) - 2^{-n/100}$$

(for large enough n).

Using Lemma 5.3.1, one can prove a lower bound for the probabilistic communication complexity of the Set Disjointness problem. The proof uses a similar argument to the one in Section 5.1. Let us give some more details. Fix ϵ to be a very small constant (say, $\epsilon = 1/1000$). A probabilistic protocol for $Disj$ outputs the right answer for every input (x, y) with probability of at least $1 - \epsilon$. Hence, it outputs the right answer with probability $1-\epsilon$, also if the input is chosen randomly (according to any distribution). Averaging over inputs in $T_0 \cup T_1$ (using a similar argument to the one used in Section 5.1), we can prove the existence of a deterministic protocol P_r that outputs the right answer for all but a 2ϵ fraction of the inputs in T_0 and all but a 2ϵ fraction of the inputs in T_1. If the communication complexity of P_r is k then P_r defines a partition of the input matrix A into 2^k rectangles. In each of these rectangles, the output $P_r(x, y)$ is fixed. Since P_r outputs the right answer for almost all inputs in T_0, we know that almost all the inputs in T_0 are in rectangles where $P_r(x, y) = 0$. Hence, we know by Lemma 5.3.1 that a considerable fraction of T_1 is also in rectangles where $P_r(x, y) = 0$. That is, the protocol P_r is wrong for a considerable fraction of entries in T_1. This is a contradiction, since we know that P_r is right for at least $(1 - 2\epsilon)$ fraction of entries in T_1.

Formally, denote the rectangles where $P_r(x, y) = 0$ by $\{U_i \times V_i\}_{i=1}^m$. Since P_r is right on at least $(1 - 2\epsilon)$ fraction of T_0, we know that

$$\sum_{i=1}^{m} \alpha_0(A_i \times B_i) \geq (1 - 2\epsilon).$$

Hence, by applying Lemma 5.3.1 for each one of the rectangles $U_i \times V_i$, (and using the fact that $m \leq 2^k$),

$$\sum_{i=1}^{m} \alpha_1(A_i \times B_i) \geq \frac{1}{100}(1 - 2\epsilon) - 2^k \cdot 2^{-n/100}.$$

If the communication complexity k is smaller than (say) $n/200$ we get (for large enough n),

$$\sum_{i=1}^{m} \alpha_1(A_i \times B_i) \geq \frac{1}{200}.$$

That is, for at least $1/200$ fraction of T_1 the protocol P_r outputs 0. This is a contradiction to the fact that the protocol P_r outputs the right answer for at least $(1 - 2\epsilon)$ fraction of entries in T_1.

Note that the above lower bound is proved not only for the Set Disjointness function but also for any other function that separates T_0 from T_1.

5.4. The 3-Distinctness Problem

The 3-Distinctness problem is the following: $X, Y = \{a, b, c\}^n$. That is, each player gets a string of n letters from $\{a, b, c\}$. The goal is to decide whether or not there is a coordinate i, such that $x_i = y_i$. As before, it is easy to give tight upper and lower bounds of $(\log_2 3) \cdot n + O(1)$ for the deterministic communication complexity of the problem. We will show a lower bound of $\Omega(n)$ for the probabilistic case, by a reduction from the Set Disjointness problem.

Assume that we have a (probabilistic) communication protocol P for the 3-Distinctness problem. We will show how to use this protocol to solve the Set Disjointness problem. Given an input pair (x, y) for the Set Disjointness problem, the two players will generate inputs x', y' for the 3-Distinctness problem in the following way. To generate x', Player 1 starts from x and translates 0 to b and 1 to a. That is, for every i, if $x_i = 0$ then $x'_i = b$ and if $x_i = 1$ then $x'_i = a$. To generate y', Player 2 starts from y and translates 0 to c and 1 to a. That is, for every i, if $y_i = 0$ then $y'_i = c$ and if $y_i = 1$ then $y'_i = a$. Obviously, $x_i = y_i = 1$ iff $x'_i = y'_i$. Hence, the two players can use the protocol P (applied on (x', y')) to solve the Set Disjointness problem for (x, y). Since we already know that any protocol for $Disj$ is of communication complexity $\Omega(n)$, we conclude that the communication complexity of P is $\Omega(n)$.