Communication Complexity

Michel de Rougemont
IRIF-CNRS

Abstract

Introduction to Communication Complexity, a framework to better understand the algorithmic computational complexity. We will use it to prove lower bounds for Testers and Streaming algorithms.

Contents

1 Introduction

In the classical setting for the complexity class IP, a Prover and a Verifier interact and the goal is to have the simplest possible Verifier, whereas the Prover is arbitrary. In the classical definition the Verifier is a BPP algorithm. In the PCP setting the Verifier fixes in advance a possible long proof but the Verifier must be simple, in particular ask few queries. In the new *Communication Complexity* setting, we want to minimize the number of bits exchanged between the Prover and the Verifier, renamed Alice and Bob.

There are two parties: Alice and Bob. Each one knows a variable, $x \in X$ for Alice, and $y \in Y$ for Bob. They want to compute a function $f(x, y)$ and we are only interested in the number of bits exchanged by the two players. Let us assume that x, y are of length n. A protocol is a binary tree where each node is labeled A, i or B, j to indicate that Alice sends x_i to Bob or that Bob sends y_j to Alice. If they compute new bits, we label them $x_{n+1}, x_{n+2}, ...$ for Alice and similarly for Bob.

The classical references are: the books [2, 4].

2 Deterministic protocols

A deterministic protocol is a sequence of bits sent by Alice and Bob. Alice sends the bit x_i and we start a decision tree: on the left branch $x_i = 0$ and on the right branch $x_i = 1$. We continue with the next bit send by Alice or Bob, unless the value of the function f is attached to the current node. Let $P(x, y)$ be the output of the protocol and $C(x, y)$ its cost, i.e. the number of bits exchanged by the protocol P or the depth of the tree on input (x, y).

For each level of the tree, there is a function $f(x_{i_1}, ...x_{i_k}, y_{i_1}, ...y_{i_l}, x, y)$ which depends on the previous choices $x_{i_1}, ...x_{i_k}$ of Alice and $y_{i_1}, ...y_{i_l}$ of Bob and determines the choice of Alice or Bob. On each leaf, there is the value of the function.

Definition 1 $C(f) = Min_P \ Max_{x,y} \ C(x, y)$

Let $M(f)$ be the matrix where the lines are the possible x, the columns are the possible y and $M(x, y) = f(x, y)$.

Examples: the equality function $EQ(x, y)$, the Parity function $P(x, y)$ for $x, y \in \{0, 1\}^n$. Suppose $x, y \subseteq \{1, 2, ...n\}$, the $AVG(x, y)$ function computes the average value of the multiset $x \cup y$, the $MAX(x, y)$ function computes the maximum value of x and y.

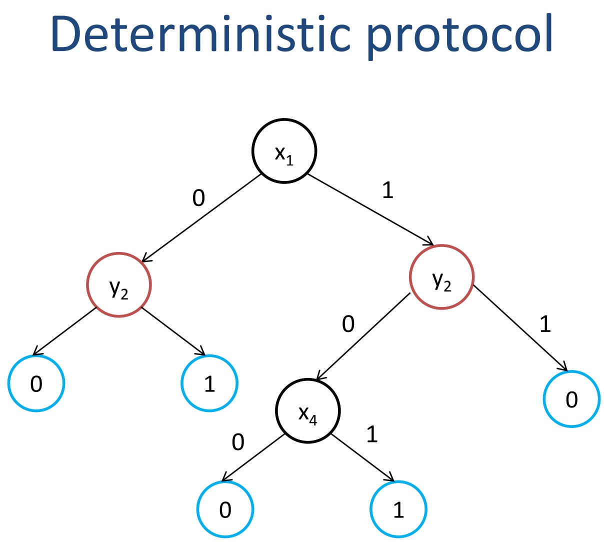

Figure 1: A deterministic protocol

2.1 Rectangles and fooling sets

Suppose we observe the bits exchanged between Alice and Bob, say z. What can we say about the inputs (x, y) which give the same z ? We can view z as a path in the decision tree and all inputs which follow the same path yield the same z.

A *rectangle* is a subset of $X.Y$ which can bre written as $A.B$ for $A \subseteq X$ and $B \subseteq Y$.

Lemma 1 *For every transcript z of a deterministic protocol P, the set of (x, y) which generate z is a rectangle.*

Proof: By induction on the length of z. At the beginning the set of (x, y) is $X.Y$, a rectangle. Assume the property true at depth d, i.e. the set of (x, y) at depth d is the rectangle $A.B$. At depth $d+1$, suppose Alice transmits x_i. Let A_1 the subset of A such that $x_i = 1$, and symetrically let A_0 the subset of A such that $x_i = 0$. On the right branch the set of (x, y) is $A_1.B$ and on the left branch the set of (x, y) is $A_0.B$. Both are rectangles. $\square$

Assume P is a correct protocol for a function f, and a rectangle along P is the rectangle of a path in the decision tree.

Lemma 2 *If P computes f, every rectangle along P is monochromatic in $M(f)$.*

Proof: Along each rectangle the protocol gives the same value, which must be $f(x, y)$. $\square$

Theorem 1 *Let f be a function such that every partition of $M(f)$ into monochromatic rectangles requires r rectangles. Then $C(f) \geq \log r$.*

Proof: The number of leaves is greater then r hence the depth is greater than $\log r$. $\square$

A *fooling set* is a set $S \subseteq X.Y$ such that there exists z and:

- for every $(x, y) \in S$, then $f(x, y) = z$,
- for two distinct pairs $(x_1, y_1), (x_2, y_2) \in S$, then either $f(x_1, y_1) \neq z$ or $f(x_2, y_2) \neq z$

Lemma 3 *If f has a fooling set S of size t, then $C(f) \geq \log t$.*

Proof: Each monochromatic rectangle of S is of size at most 1, hence there are at least t monochromatic rectangles and by theorem 1, $C(f) \geq \log t$. $\square$

3 Randomized protocols

Alice and Bob access public or private coins r_A and r_B. The protocol is a tree which depends on x, y, r_A, r_B. A protocol P computes with 0-error if:

$$\mathbb{P}rob_{r_A,r_B}[P(x, y) = f(x, y)] = 1$$

A protocol P computes with ε-error if:

$$\mathbb{P}rob_{r_A,r_B}[P(x, y) = f(x, y)] \geq 1 - \varepsilon$$

There several variations:

- One way communication or 2-ways,
- Private or public coins,
- One sided or two-sided errors.

The random choices r_A and r_B generate different communication costs (number of bits exchanged). For a given (x, y), the *worst-case cost on* (x, y) is the maximum number of bits exchanged over all choices of r_A, r_B. The *worst-case cost* is the maximum worst-case cost over all choices of (x, y). One could also choose an *average cost on* (x, y) as the average number of bits exchanged over all choices of r_A, r_B. The *average cost* is the maximum average cost on (x, y) over all choices of (x, y).

- $R_\varepsilon(f)$ is the minimum worst case cost with error for protocols which computer with ε error, i.e.
$$R_\varepsilon(f) = Min_P \ Max_{x,y} \ C(x, y)$$

- $R^1_\varepsilon(f)$ is the minimum worst case cost with 1-sided error ε,
- $R^{pub}_\varepsilon(f), R^{1\ pub}_\varepsilon(f)$ are the version with public coins.

The complexity class BPP^{cc} is the class of problems such that $R_\varepsilon(f) \in O(poly(\log n))$.

3.1 Private vs. public coins

Consider the equality function $EQ(x, y)$. With a public coin r, Alice computes $x.r = \sum_i x_i.r_i$, Bob computes $y.r = \sum_i x_i.r_i$ and Alice sends the result to Bob who compares the two results. The result is 1 if the two results ar equal and 0 otherwise. If $x = y$ then $I\!Prob_r[P(x, y) = f(x, y)] = 1$. If $x \neq y$, then $I\!Prob_r[P(x, y) = f(x, y)] = 1/2$ and the error is $1/2$. If we repeat the test k times and accept if all the results coincide and reject otherwise, the error probability decreases to $(1/2)^k$.

With private coins, the previous protocol would require to communicate r. Another protocol is as follows:

Let $x = a$ be Alice's input and $y = b$ be Bob's input. Let Let p be a prime number $n^2 \leq p \leq 2.n^2$ and:

$$A(x) = a_0 + a_1.x + ...a_n.x^n \ (p)$$

$$B(x) = b_0 + b_1.x + ...b_n.x^n \ (p)$$

be two polynomials of degree $n - 1$. Alice picks a random number $t < p$ and sends t and $A(t)$ to Bob. Bob accepts (output is 1) if $A(t) = B(t)$, otherwise Bob rejects (output is 0). If $x = y$, it is correct. If $x \neq y$ the error probability is when t is a root of $A(t) - B(t)$ and there are at most $n - 1$ roots. The error probability is $(n - 1)/p \leq (n - 1)/n^2 \leq 1/n$.

3.2 Distributional Complexity

Let $D^\mu_\varepsilon(f)$ be the cost of the best deterministic protocol which gives the right answer to a fraction at least $1 - \varepsilon$ of the input (x, y) weighted by μ.

$$D^\mu_\varepsilon(f) = Min_P \ \ Max_{x,y} \ \ C(x, y)$$

for deterministic protocols such that $I\!Prob_{x,y\in\mu}[P(x, y) = f(x, y)] \geq 1 - \varepsilon$.

Lemma 4 $R^{pub}_\varepsilon(f) \geq Max_\mu \ D^\mu_\varepsilon(f)$.

Proof: Let P be a randomized public coin protocol which computes f with error at most ε on all inputs (x, y). In particular for any μ:

$$I\!Prob_{r.\mu}[P(x, y) = f(x, y)] \geq 1 - \varepsilon$$

Therefore there exists some fixed r' and a deterministic $P_{r'}$ such that:

$$I\!Prob_\mu[P_{r'}(x, y) = f(x, y)] \geq 1 - \varepsilon$$

Then $R^{pub}_{\varepsilon}(f) \geq cost(P_{r'})$ if $cost(P_{r'})$ is the cost weighted by μ, as $R^{pub}_{\varepsilon}(f)$ considers the worst-case cost. Notice that $cost(P_{r'}) \geq Max_{\mu}\ D^{\mu}_{\varepsilon}(f)$, as a specific $P_{r'}$ has a cost larger than the cost for of the optimal deterministic protocol on μ, hence $R^{pub}_{\varepsilon}(f) \geq Max_{\mu}\ D^{\mu}_{\varepsilon}(f)$. $\square$

We now consider the reverse inequality, which requires the use of the MinMax theorem.

3.2.1 Zero sum Games

A zero-sum game between two players is defined by a matrix A of dimension (n, m), when the strategies of I are $1, 2, ...n$ and the strategies of II are $1, 2, ...m$. The utility of player I when he plays strategy i and player II plays strategies j, is $A(i, j)$. Similarly the utility of player II is $B(i, j)$ and $A + B = 0$. The two players are opposed: the gain of one of the player is the loss of the other.

The players play mixed strategies x and y and the expected gain for I is:

$$x^t.A.y = \sum_{i,j} x_i.A(i, j).y_j$$

I tries to maximize his gain and II tries to minimize it. Hence x satisfies:

$$Max_x\ Min_y\ x^t.A.y$$

Let y_j be a pure strategy for II and let be $t = Min_{y_j}\ x^t.A.y = Min_j\ \sum_i x_i.A(i, j)$.
Observe that $t \geq Min_y\ x^t.A.y$ as there are infinitely many possible y.
Let us show that $t \leq Min_y\ x^t.A.y$. Observe that

$$x^t.A.y = \sum_{i,j} x_i.A(i, j).y_j = \sum_j y_j. \sum_i x_i.A(i, j) \geq \sum_j y_j.t = t$$

We conclude that $Min_y\ x^t.A.y = Min_{y_j}\ x^t.A.y = Min_j\ \sum_i x_i.A(i, j)$.
We can therefore find the best x with a linear program over the variables $t, x_1, ...x_n$.

$$Max\ t$$

$$t \leq \sum_i x_i.A(i, j), \quad for\ j = 1, ...m$$

$$\sum_i x_i = 1$$

We can argue similarly, if we start with the expression $Min_y\ Max_x\ x^t.A.y = Min_y\ Max_i\ y^t.A^t.x_i$, which is the goal of player II.

$$Min\ z$$

$$z \geq \sum_j y_j.A^t(j, i), \quad for\ i = 1, ...n$$

$$\sum_j y_j = 1$$

The gain of player II is just $-z$ for the optimal z. These two programs are dual. By duality:

$$Max_x\ Min_y\ x^t.A.y = Min_y\ Max_x\ x^t.A.y$$

The optimal value is called the *value of the game*. It is also a Nash equilibrium.

Examples of classical games such as Rock-Paper-Scissors- where $n = m = 3$ and the matrix defining the gain of player I is:

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}$$

In this case the value of the game is 0 and the optimal strategy of I and II is:

$$x_* = y_* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$$

Consider now the matrix

$$A = \begin{bmatrix} -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}$$

In this case, the value is 0.33 for player I, hence -0.33 for player II. The optimal strategy of player I is:

$$x_* = \begin{bmatrix} 0,0833 \\ 0,1250 \\ 0,3333 \\ 0,1250 \\ 0,3333 \end{bmatrix}$$

If we take the dual, the optimal strategy of player II is:

$$y_* = \begin{bmatrix} 0,3333335 \\ 0,333333 \\ 0 \\ 0,3333335 \end{bmatrix}$$

3.2.2 Protocols versus data

Consider the 0-sum game between 2 players. Player 1 is the protocol designer: his pure strategies are the deterministic protocols exchanging c-bits. Player 2 chooses the input x, y. We label the utility matrix $M(P, (x, y))$ by 1 if $P(x, y) = f(x, y)$ and by -1 otherwise.

Theorem 2

$$R_\varepsilon^{pub}(f) = Max_\mu \; D_\varepsilon^\mu(f)$$

Proof: Let us show that $R_\varepsilon^{pub}(f) \leq Max_\mu \; D_\varepsilon^\mu(f)$. Let $c = Max_\mu \; D_\varepsilon^\mu(f)$.

The mixed strategies of player 1 are the randomized strategies with public coins. The mixed strategies of player 2 are the distributions μ on the inputs.

We assume $D_\varepsilon^\mu(f) \leq c$. In this setting:

$$Max_P \; Min_\mu \; M(x,y) \geq 1 - \varepsilon$$

By the Minimax theorem, we conclude that:

$$Min_\mu \; Max_P \; M(x,y) \geq 1 - \varepsilon$$

Consider the distribution which achieves the minimum. There is a deterministic protocol P which achieves the bound. A probabilistic protocol can only be better. Hence $Max_\mu \; D_\varepsilon^\mu(f) \geq R_\varepsilon^{pub}(f)$. □

4 Lower bounds

There are at least 3 methods to show lower bounds on the Communication Complexity. Most methods use the previous result and look for a distribution μ where $D_\varepsilon^\mu(f)$ is larger than some bound.

- the Corruption bound method, where we look at $\varepsilon-$monochromatic rectangles. We show that if there are large, they must be corrupted.
- the Discrepancy based method,
- the method based on the Information Complexity.

4.1 Lower bound on Disjointness via the corruption bounds

We describe an $\Omega(\sqrt{n})$ lower bound, based on [1].

The distribution: uniform distribution on the $\binom{n}{\sqrt{n}}$ binary words with $\sqrt{n}$ bits at 1. We consider almost monochromatic rectangles $R = S \times T$, i.e.

$$\mathbb{P}rob_\mu[Disjointness(x,y) = 0 | (x,y) \in R] \leq \varepsilon$$

and show that either $|S|$ or $|T|$ are small, i.e. less than

$$2^{-c.\sqrt{n}}. \binom{n}{\sqrt{n}}$$

Consequently, there are many almost monochromatic rectangles and by theorem 1, the communication complexity is large.

We first prove some useful lemmas. Let:

$$S' = \{x : x \in S, |x \cap y| \leq 2\varepsilon.|T|\}$$

i.e. the x which intersect with at most $2\varepsilon.|T|$ of the $y \in T$. As $S \times T$ is ε-monochromatic, $|S'| \geq |S|/2$.

Lemma 5 *Assume S is large, then there exists $x_1, ...x_k \in S'$ where $k = \sqrt{n}/3$ such that:*

$$|x_i \cap \bigcup_{j<i} x_i| \leq \sqrt{n}/2$$

Proof: Each new x_i brings at least $\sqrt{n}/2$ new points. Let us prove the property by induction on i. Let $z_i = \bigcup_{j<i} x_i$, which is of size less than $\sqrt{n}/.\sqrt{n}/3 = n/3$ and we selected $\{x_1, ...x_{i-1}\}$. Let us evaluate how many $x's$ are such that $|x \cap z_i| \geq \sqrt{n}/2$, i.e. bad points. We show that are only few and therefore we can always find a new x_i which satisfies the condition.

We select $j \geq \sqrt{n}/2 \in z_i$ and $\sqrt{n} - j \notin z_i$, hence the possible number of x_i is less than:

$$\sum_{j=\sqrt{n}/2}^{\sqrt{n}} \binom{n/3}{j} . \binom{2n/3}{\sqrt{n} - j}$$

$$\leq \sqrt{n}/2. \binom{n/3}{\sqrt{n}/2} . \binom{2n/3}{\sqrt{n}}$$

$$\leq \sqrt{n}/2. \binom{2n/3}{\sqrt{n}}$$

$$\leq \sqrt{n}/2.(2/3)^{\sqrt{n}} \binom{n}{\sqrt{n}}$$

$$\leq 2^{-c\sqrt{n}}. \binom{n}{\sqrt{n}}, \quad c = (\log 3 - \log 2)/2. \log 2$$

$$< |S|$$

Hence there is an x_i such that $|x_i \cap \bigcup_{j<i} x_i| \leq \sqrt{n}/2$. □

Let S'' be the set $x_1, ...x_k$ and $T'' = \{y : x \in S'', \ \ |x \cap y| \leq 4\varepsilon.|S''|\}$. Each y intersect at most 4ε fraction of S''.

Lemma 6 *If S is large, then T is small*

Proof: Notice that $T'' \geq |T|/2$, using the same averaging argument as for S'. Let us bound the size of T''. Each y does not intersect with at least $(1-4\varepsilon).k$ of the x_i. The size of these x_i is greater than $(1-4\varepsilon).k.\sqrt{n}/2 \geq k.\sqrt{n}/3 = n/9$. The number of possible y is then less than:

$$\binom{k}{4\varepsilon.k}\binom{8n/9}{\sqrt{n}} = \binom{\sqrt{n}/3}{4\varepsilon.\sqrt{n}/3}\binom{8n/9}{\sqrt{n}} \leq 2^{-c'\sqrt{n}}.\binom{n}{\sqrt{n}}$$

□

Theorem 3 *There exists a distribution μ such that $D^{\mu}_{\varepsilon}(Disjointness) = \Omega(\sqrt{n})$.*

Proof: Let μ be the distribution where x, y are choosen with $\sqrt{n}$ elements uniformly among n. All the ε-monochromatic rectangles have size smaller than $2^{-c.\sqrt{n}}$. $|X|.|Y|$. We need more than $2^{c.\sqrt{n}}$ of these rectangles and hence $D^{\mu}_{\varepsilon}(Disjointness) = \log 2^{c.\sqrt{n}} = \Omega(\sqrt{n})$. □

A more sophisticated μ [3] leads to the central result:

Theorem 4 *There exists a distribution μ such that $D^{\mu}_{\varepsilon}(Disjointness) = \Omega(n)$.*

4.2 Lower bound via the Discrepancy

For a rectangle R and a distribution μ, let:

$$Disc_{\mu}(R, f) = |Prob_{\mu}[f(x,y) = 0 \wedge (x,y) \in R] - Prob_{\mu}[f(x,y) = 1 \wedge (x,y) \in R]|$$

For a function f and a distribution μ, let:

$$Disc_{\mu}(f) = Max_R Disc_{\mu}(R, f)$$

There is a direct link between $Disc_{\mu}(f)$ and $D^{\mu}_{\varepsilon}(f)$.

Lemma 7 *For every f, μ, ε,*

$$D^{\mu}_{1/2-\varepsilon}(f) \geq \log(2\varepsilon/Disc_{\mu}(f)$$

Proof: Consider a deterministic protocol π, with error $1/2-\varepsilon$ and communication cost c.

$$Prob_{\mu}[\pi(x,y) = f(x,y)] \geq 1/2 + \varepsilon$$

$$Prob_{\mu}[\pi(x,y) \neq f(x,y)] \geq 1/2 - \varepsilon$$

Hence:

$$Prob_{\mu}[\pi(x,y) = f(x,y)] - Prob_{\mu}[\pi(x,y) \neq f(x,y)] \geq 2\varepsilon$$

For all the leaves l of the protocol, corresponding to a rectangle R_l:

$$= \sum_l Prob_{\mu}[\pi(x,y) = f(x,y) \wedge (x,y) \in R_l] - Prob_{\mu}[\pi(x,y) \neq f(x,y) \wedge (x,y) \in R_l] 2\varepsilon$$

On each leave, there is either 0, i.e. $\pi(x,y)=0$ or 1, i.e. $\pi(x,y)=1$. Hence:

$$Prob_\mu[\pi(x,y)=f(x,y)\wedge(x,y)\in R_l]-Prob_\mu[\pi(x,y)\neq f(x,y)\wedge(x,y)\in R_l]=$$

$$|Prob_\mu[f(x,y)=0\wedge(x,y)\in R_l]-Prob_\mu[f(x,y)=1\wedge(x,y)\in R_l]|$$

There are at most 2^c leaves if $D^\mu_{1/2-\varepsilon}(f)=c$. Hence:

$$2^c.Disc_\mu(f)\geq 2\varepsilon$$

$$c\geq\log(2\varepsilon/Disc_\mu(f)$$

□

Let us apply the method to show an $\Omega(n)$ lower bound for the function $IP(x,y)=\sum_i x_i.y_i$.

Lemma 8 *$DISC_{uniform}(IP)\leq 2^{-n/2}$*

Proof: Let H be the matrix such that $H(x,y)=1$ if $IP(x,y)=0$ and $H(x,y)=-1$ if $IP(x,y)=1$. Notice that $H.H^t=2^n.I$, i.e. is the identity matrix multiplied by 2^n.

Let us show that $H.H^t(x,y)=2^n$ if $x=y$, otherwise $H.H^t(x,y)=\sum_z H(x,z).H^t(z,y)=0$. $H(x,z)$ is 1 for half of the $z's$ and $H(x,z)$ is -1 for the other half. $\sum_z H(x,z).H^t(z,y)=\sum_z H(x,z).H(y,z)$, and $H(x,z)=H(y,z)$ for half of the $z's$ and $H(x,z)\neq H(y,z)$ for the other half.

We compare $\sum_i x_i.z_i$ with $\sum_i y_i.z_i$. Suppose x,y differ in positions j_1,j_2: just consider the 4 possibilities where $z_{j_1}=0$ or 1, and $z_{j_2}=0$ or 1. For $z_{j_1}=z_{j_2}$, then $\sum_i x_i.z_i=\sum_i y_i.z_i$ and $H(x,z)=H(y,z)$. For $z_{j_1}\neq z_{j_2}$, then $\sum_i x_i.z_i\neq\sum_i y_i.z_i$ and $H(x,z)\neq H(y,z)$. There are as many z in each case.

Observe that the Froebinius norm $||H||=\sqrt{2^n}$. For a rectangle $S.T$, let 1_S (resp. 1_T) be the characteristic vector of S (resp. T). For the uniform distribution, observe that:

$$DISC_{uniform}(S.T,IP)=\frac{|\sum_{x\in S,y\in T}H(x,y)|}{2^{2n}}=\frac{|1_S.H(x,y).1_T|}{2^{2n}}$$

As $|S|\leq 2^n$, $||S||\leq\sqrt{2^n}$ By the property of the norm,

$$|1_S.H(x,y).1_T|\leq||1_S||.||H(x,y)||.||1_T||=\sqrt{2^{3n}}$$

Hence:

$$DISC_{uniform}(IP)=Max_{S,T}\leq DISC_{uniform}(S.T,IP)\leq\frac{\sqrt{2^{3n}}}{2^{2n}}=2^{-n/2}$$

□

Corollary 1 *$D^{uniform}_{1/2-\varepsilon}(IP)\geq n/2-\log(1/\varepsilon)$ and $R^{pub}_\varepsilon(IP)\geq n/2-\log(2/\varepsilon)$*

Proof: By lemma 7

$$D^{uniform}_{1/2-\varepsilon}(f) \geq \log(2\varepsilon/Disc_{uniform}(f)$$

By lemma 8, $DISC_{uniform}(IP) \leq 2^{-n/2}$, hence

$$D^{uniform}_{1/2-\varepsilon}(IP) \geq \log(2\varepsilon/Disc_{uniform}(IP) \geq n/2 + \log(2\varepsilon)$$

$$R^{pub}_{\varepsilon}(IP) \geq D^{uniform}_{\varepsilon}(f) \geq n/2 - \log(2/\varepsilon)$$

□

4.3 Lower bound via the Information Complexity

5 Reductions

In the classical setting, problem A reduces to problem B, writen $A \leq B$ if there is a function f such that $x \in A$ iff $f(x) \in B$. in the communication complexity setting $(x, y) \in A$ iff there is an algorithm for Alice (or Bob) which decides if $(f(x), f(y)) \in B$.

- $INDEX \leq DISJOINTNESS$

  INDEX $(x, i) = x_i$ and *DISJOINTNESS*$(x, y) = 1$ if $\forall i x_i \neq y_i$.

  Alice has an n-bit vector and Bob has a position i ($\log n$ bits) and the output is x_i. Let $f(x) = x$ and $f(i) = e_i$ the word with a 1 in position i and 0 elsewhere. If *INDEX* $(x, i) = 1$ then *DISJOINTNESS*$(f(x), f(i)) = 0$ and if *INDEX* $(x, i) = 0$ then *DISJOINTNESS*$(f(x), f(i)) = 1$.

References

[1] Laszlo Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, SFCS '86, pages 337–347, 1986.

[2] Eyal Kushilevitz and Noam Nisan. *Communication complexity.* Cambridge University Press, New York, 1997.

[3] A. A. Razborov. On the distributional complexity of disjointness. *Theor. Comput. Sci.*, 106(2):385–390, December 1992.

[4] Tim Roughgarden. Communication complexity (for algorithm designers). *arXiv*, abs/1509.06257, 2015.