Communication Complexity

Boaz Barak

October 2, 2012

Definition There is a function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, Alice gets input $x \in \{0,1\}^n$ and Bob gets input $y \in \{0,1\}^n$. They exchange messages with one another according to some pre-coordinated protocol (if we don't care about constant factors, without loss of generality the protocol can be that they have $k(n)$ rounds, where in odd rounds Alice sends a bit to Bob and in even rounds Bob sends a bit to Alice). The last message in the protocol should be equal to $f(x,y)$.

Define $C(f)$ to be the amount of bits sent by the protocol with minimal communication that computes f.

Note that $C(f) \leq O(n)$ for every f.

Running Examples **XOR** $f(x,y) = x_1 \oplus \cdots \oplus x_n \oplus y_1 \oplus \cdots \oplus y_n$

EQUALITY $f(x,y) = 1$ iff $x = y$.

INNER PRODUCT $f(x,y) = \oplus_{i=1}^n (x_i y_i)$

DISJOINTNESS $f(x,y) = \neg \vee_{i=1}^n (x_i y_i)$

Motivation, Applications Hits "sweet spot" of simplicity vs. richness. Great many applications.

Equality requires linear communication

Theorem 1. $C(EQUALITY) \geq n$

Proof. Suppose $C(EQUALITY) < n$ and let Π be a protocol computing this function with communication at most $n-1$ bits. Then, there are inputs $x \neq x'$ such that $\Pi(x,x) = \Pi(x',x')$. We claim that $\Pi(x,x') = \Pi(x,x)$. This would yield a contradiction because $EQUALITY(x,x) = 1$ but $EQUALITY(x,x') = 0$.

Indeed, let $\vec{m} = (m_1, ..., m_k)$ be the messages $\Pi(x,x) = \Pi(x',x')$ (we assume as above w.l.o.g that Alice sends odd messages and Bob sends even ones). Clearly m_1 is also the first message in $\Pi(x,x')$ since its only based on Alice's input x. Now, $\vec{m} = \Pi(x',x')$ means that on input x' and after seeing message m_1, Bob sends the message m_2, and hence m_2 is the second message of $\Pi(x,x')$. We can continue this way for all messages. □

Application to lower bounds for one tape TM's Here is one example how communication complexity results can yield complexity lower bounds. Recall that as far as we know, a 2-tape Turing machine can solve SAT in $\tilde{O}(n)$ time. However, it turns out that the simple 1 tape TM is much more restricted, and can't even solve the much lowlier language of Palindromes:

Theorem 2. *Let $PALIN = \{xx^R : x \in \{0,1\}^k\}$ (where $x^R = x_k ... x_1$) then any 1-tape TM M takes $\Omega(n^2)$ time to solve $PALIN$.*

Proof. Suppose for the sake of contradiction that there exists such a TM M, and so in particular it runs in time $o(n^2)$ on inputs of the form $x0^nx^R$ for $x \in \{0,1\}^n$. For every $x \in \{0,1\}^n$ and $i \in [n]$, let $f_i(x)$ the number of steps in the computation that the head of M touches the $n+i^{th}$ cell of the tape on input $x0^nx^R$ (i.e., the location which initially contains the i^{th} zero). Note that for every x, $f_1(x) + \cdots + f_n(x) = o(n^2)$ and so there is some $i(x)$ such that $i(x) = o(n)$.

Now consider the following protocol for solving the equality function:

On inputs x, x' the two players will simulate a run of the TM on the input $x0x'^R$.

- Alice computes $i(x)$, $f_i(x)$ and sends these $2\log n$ bits to Bob, Bob computes $i(x'), f_i(x')$ and sends to Alice. If the numbers disagree then they output zero.
- They start simulating the machine, with Alice responsible for simulating the machine when the head is in locations $1 \ldots n+i$ and Bob responsible to simulating it when the head is in locations $i+n+1 \ldots$.. Any time the head moves from the $n+i^{th}$ position to the right then Alice sends to Bob the (constant sized) state of the TM so he can proceed with the simulation. Similarly Bob sends to Alice the state every time the head moves left into the $n+i^{th}$ position.
- Their total communication is a constant times $f_i(x) = f_i(x') = o(n)$, and the TM outputs 1 iff $x = x'$ hence we get a contradiction.

□

Rank method Here is another proof that $C(EQUALITY) \geq \Omega(n)$

Theorem 3. *$C(f) \geq \log \mathsf{rank}(f)$, where $\mathsf{rank}(f)$ denotes the rank of f when considered as a $2^n \times 2^n$ matrix over the reals.*

Proof. Suppose that $C(f) = k$. Let $\vec{m}$ is one of the $\leq 2^k$ possible transcripts that can arise when the output of f is 1. Now let $S_{\vec{m}} \subseteq \{0,1\}^n \times \{0,1\}^n$ be the set of inputs that result in the transcript $\vec{m}$. Do the same considerations as before, $S_{\vec{m}} = X_{\vec{m}} \times Y_{\vec{m}}$ for some subsets $X_{\vec{m}}, Y_{\vec{m}} \subseteq \{0,1\}^n$.

Thus, we get that

$$f = \sum_{\vec{m}} 1_{X_{\vec{m}}} \times 1_{Y_{\vec{m}}}$$

showing that $\mathsf{rank}(f) \leq 2^k$. □

Other functions As corollaries, we get that $C(INNERPRODUCT) \geq \Omega(n)$, $C(DISJOINTNESS) \geq \Omega(n)$.

Log rank conjecture It is actually believed that $C(f) \leq \text{poly}(\log(\mathsf{rank}(f)))$.

Randomized Communication Complexity A very natural extension of the model allows Alice and Bob to use *randomization.* That is, their goal is now to output $f(x,y)$ with probability at least 0.99 (taken over the coins). Define $R(f)$ as smallest randomized communication complexity of f.

Question: do our proof extend to this case? Can we show $R(EQUALITY)$, $R(INNERPRODUCT)$, $R(DISJOINTNESS) \geq \Omega(n)$?

Yao's Min Max principle It seems hard to argue about randomized protocols directly, but the following simple but powerful observation of Yao allows us to talk instead about average-case hardness for *deterministic* protocols.

Theorem 4. *Let* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *and suppose that there is a distribution* D *over* $\{0,1\}^n \times \{0,1\}^n$ *such that for every deterministic protocol* Π *of communication complexity* k *it holds that* $\Pr_{(x,y)\in D}[f(x,y) = \pi(x,y)] \leq 0.9$. *Then* $R(f) \geq k$.

Proof. We can think of a randomized protocol of communication k as a distribution $\mathcal{P}$ over deterministic protocols. So, we know that for every (x,y), the probability for a random $\Pi \in \mathcal{P}$ that $\Pi(x,y) = f(x,y)$ is at least 0.99. In particular this holds when (x,y) is chosen from D, but then by an averaging argument this means that there exists some Π such that $\Pr_{(x,y)\in D}[f(x,y) = \pi(x,y)] \leq 0.99$. □

Inner Product Here is how we can use this idea to show that the INNER PRODUCT function requires $\Omega(n)$ communication.

Theorem 5. $R(INNERPRODUCT) \geq \Omega(n)$

Proof. Let $D = U_n \times U_n$, and suppose towards a contradiction that there is some Π with communication $k = n/100$ that can solve $INNERPRODUCT$ on D with probability 0.99. Then, by averaging there must exist a transcript $\vec{m}$ such that the probability over D that $\Pi(x,y) = \vec{m}$ is at least 2^{-k} and that (setting b to be the last bit of $\vec{m}$) the probability, conditioned on $\Pi(x,y) = \vec{m}$ that $INNER-PRODUCT(x,y) = b$ is at least 0.9.

In other words, we have sets $A, B \subseteq 2^n$ such that $|A| \cdot |B| \geq 2^{2n-k}$ but

$$\sum_{x\in A, y\in B} (-1)^{INNER-PRODUCT(x,y)} \geq 0.8|A||B|$$

Now, if we consider the vectors $1_A, 1_B$ then we get that $1_A^T H 1_B \geq 0.8 \cdot 2^{2n-k}$ where H is the $2^n \times 2^n$ matrix such that $H_{x,y} = (-1)^{INNER-PRODUCT(x,y)}$. But it's easy to see this matrix has orthogonal rows each of norm $2^{n/2}$, which means that it's top eigenvalue λ is at most $2^{n/2}$. That means that

$$1_A^T H 1_B \leq \lambda \|1_A\| \|1_B\| \leq 0.8 \cdot 2^{n/2} \cdot 2^{n/2} \cdot 2^{n/2}$$

or in other words (ignoring the negligible factor 0.8), we get that $2n - k \leq 3n/2$ or $k \geq n/2$. □

The property we used in this proof about the inner product function is that $Disc(INNER-PRODUCT) \leq 2^{-n/2}$ where

$$Disc(f) = \max_{A,B\subseteq\{0,1\}^n} 2^{-2n} \sum_{x\in A, y\in B} (-1)^{f(x,y)}$$

Can you compute $Disc(DISJOINTNESS)$? Why do we care about disjointness so much?