Randomized Algorithms 2015A
Lecture 13
Course Recap via Communication Complexity Lower Bounds*

Robert Krauthgamer

1 Communication Complexity

Model: Two parties, called Alice and Bob, receive inputs x, y respectively. They can exchange messages, in rounds, until one of them (or both) reports an output $f(x, y)$.

Main measure is communication complexity, i.e., total communication between the parties (in bits).

Variants of randomization: none (deterministic), shared/public, or private.

Number of rounds: zero (simultaneous, i.e., without direct communication), one (one-way communication), or more/unbounded.

Connection to sketching: simultaneous protocols can be viewed as a sketch, and vice versa.

Examples: follow from our skeching examples.

We will focus on these models.

Indexing problem:

Alice's input is $x \in \{0,1\}^n$ (equivalently a subset $T \subset [n]$), Bob's input is an index $i \in [n]$.

Their goal is to output x_i.

Theorem [Kremer, Nisan, and Ron, 1999]: The randomized one-way communication complexity of indexing is $\Omega(n)$, even with shared randomness.

It's therefore a "canonical" problem for reductions (in this model).

Proof by [Jayram, Kumar and Sivakumar, 2008]:

Assume there is a protocol with (constant) error probability $\delta > 0$ and communication complexity t. Fix an error correcting code with Hamming distance 4δ, namely, a subset $A \subset \{0,1\}^n$ of size

*These notes summarize the material covered in class, usually skipping proofs, details, examples and so forth, and possibly adding some remarks, or pointers. The exercises are for self-practice and need not be handed in. In the interest of brevity, most references and credits were omitted.

$|A| \geq 2^{\alpha n}$ for $\alpha = \alpha(\delta)$, where the for all $x \neq y \in A$, the Hamming distance is $\|x - y\|_1 \geq 4\delta n$. Consider an input chosen uniformly at random from this code A.

By taking the "best" coins in the assumed randomized protocol (we're actually using Yao's minimax principle), we get that there is also a *deterministic* protocol, whose error probability on this input distribution is $\leq \delta$.

Now suppose Alice sends the same message m for several inputs $x, x', \ldots$. On at most one of these inputs, the protocol errs on $\leq 2\delta n$ coordinates i; indeed, let $z = z(m) \in \{0,1\}^n$ be the protocol's outputs when Alice send message m to Bob, who follows the protocol with different i's as his input; then at most one of these inputs can be $< 2\delta n$-close to z (otherwise, we have $x, x' \in A$ such that by triangle inequality $\|x - x'\|_1 \leq 4\delta n$).

Overall, for at most 2^t of Alice's inputs $x \in A$, the protocol errs on $< 2\delta n$ coordinates $i \in [n]$, thus looking at the "rest", we have

$$\tfrac{2^{\alpha n} - 2^t}{2^{\alpha n}} \cdot 2\delta \leq \Pr_{\text{input}}[\text{det. protocol errs}] \leq \delta.$$

Simplifying, we get $t \geq \alpha n - 1$.

QED.

Exer: Use Yao's minimax principle to prove an $\Omega(n)$ lower bound for the following problem. The input is an array of n bits (accessed only by reading a single bit each time), and the goal is to find a position where the array contains 1.

2 Gap Hamming Distance (GHD)

Problem definition of GHD: Alice and Bob's inputs are $x, y \in \{0,1\}^n$, respectively, and their goal is to determine whether the hamming distance between x, y is $\leq \frac{n}{2} - \sqrt{n}$ or $\geq \frac{n}{2} + \sqrt{n}$.

Theorem [Woodruff, 2004]: The randomized one-way communication complexity of GHD is $\Omega(n)$, even with shared randomness.

Proof from [Jayram, Kumar and Sivakumar, 2008]: We reduce from the indexing problem, so consider inputs $u \in \{-1, +1\}^n$ and $e_i \in \{0,1\}^n$ for indexing. We shall show how to solve this instance assuming there is a protocol for GHD that uses $t = t(N)$ bits. Without loss of generality, we assume n is odd.

Alice and Bob can pick, using the shared randomness, a common $r \in \{+1, -1\}^n$, and compute, without using any communication, $x := \text{sgn}(\langle u, r\rangle)$ and $y := \text{sgn}(\langle e_i, r\rangle) = r_i$, respectively. The key idea is that for some absolute constant $c > 0$,

$$\Pr_r[x \neq y] = \Pr_r[\text{sgn}(\langle u, r\rangle) \neq r_i] \begin{cases} \geq \frac{1}{2} + \frac{c}{\sqrt{n}} & \text{if } u_i = -1; \\ \leq \frac{1}{2} - \frac{c}{\sqrt{n}} & \text{if } u_i = +1. \end{cases} \tag{1}$$

Assume for now the bound (1) holds. Then, Alice and Bob can repeat this process $N = 16n/c^2$ times, and produce $\bar{x}, \bar{y} \in \{0,1\}^N$ whose Hamming distance is WHP either $\geq (\frac{1}{2} + \frac{c}{\sqrt{n}})N - 3\sqrt{N} =$

$\frac{1}{2}N + \sqrt{N}$ or $\leq \frac{1}{2}N - \sqrt{N}$. If they apply protocol we assumed for GHD, which succeeds WHP, they can distinguish between the two cases, i.e., determine u_i, using communication of $t(N)$ bits. Applying our lower bound for indexing, $t(N) \geq \Omega(n) = \Omega(c^2 N)$.

To prove the bound (1), write $\langle u, r\rangle = u_i r_i + w$ where $w := \sum_{j\neq i} u_j r_j$; note w is random but independent of u_i. Observe that if $w \neq 0$ then necessarily $|w| \geq 2$, and then the desired probability is exactly 1/2. But with probability at least $2c/\sqrt{n}$, we have $w = 0$, in which case $\text{sgn}(\langle u, r\rangle) = u_i r_i$, and then the desired event becomes $u_i r_i \neq r_i$, and its probability is 1 when $u_i = -1$, and is 0 when $u_i = +1$. The theorem follows by the total probability formula.

QED.

Corollary: The one-way communication complexity of determining whether the Hamming distance between $x, y \in \{0,1\}^n$ is $\leq R$ or $\geq (1+\varepsilon)R$ is at least $\Omega(1/\varepsilon^2)$ bits (for suitable $R = \Theta(n)$ and assuming $n \geq 1/\varepsilon^2$).

Exer: Prove it formally.

Corollary: Approximating the ℓ_1-norm in the data stream model requires $\Omega(1/\varepsilon^2)$ bits.

Proof: Suppose there is a streaming algorithm with space requirement s. The we could design the following one-way protocol for GHD on inputs x, y. Alice executes the streaming algorithm on x, send her entire memory, which is only s bits, to Bob, who continues executing the streaming algorithm on $-y$, and then $(1+\varepsilon)$-approximates (in the above promise model) $\|x - y\|_1$. Thus $s \geq \Omega(1/\varepsilon^2)$.

Theorem [Chakrabarti and Regev, 2011]: The communication complexity (with unbounded number of rounds) of GHD is $\Omega(n)$, even with shared randomness.

Remark: Such communication complexity methods were recently used also to give tight lower bounds for cut sparsifiers [Andoni, Krauthgamer and Woodruff, 2014].