CAS CS 591 B: Communication Complexity

Prof. Mark Bun | Fall 2019

Lecture Notes 2:

Fooling Sets, Rank, and Covers

Reading.

- Rao-Yehudayoff Chapter 1 (Fooling Sets and Rectangle Covers), Chapter 2 (through pg. 58)

1 Fooling Sets

The method of fooling sets refines arguments based on rectangle size. Intuitively, a fooling set is a set of inputs which is hard to cover by disjoint unions of rectangles. Formally, a *fooling set* for a function f is a set $S \subseteq X \times Y$ such that $|R \cap S| \leq 1$ for every rectangle R which is monochromatic with respect to f.

Theorem 1 *If S is a fooling set for f, then $\mathbf{P^{cc}}(f) \geq \log |S|$.*

Proof: Recall that any protocol computing f with c bits of communication partitions $X \times Y$ into at most 2^c monochromatic rectangles. Any partition of $X \times Y$ into monochromatic rectangles must be at least as large as S. Hence $|S| \geq 2^{\mathbf{P^{cc}}(f)}$. ■

Example 2 *Going back to EQ_n, we can take $S = \{(x, x) : x \in \{0,1\}^n\}$ to see that $\mathbf{P^{cc}}(\mathrm{EQ}_n) \geq n$.*

Example 3 *Consider the function GT_n. We claim that the set $S = \{(x, x) : x \in \{0,1\}^n\}$ is a fooling set for GT_n. To see this, suppose T is a monochromatic set which contains both (x, x) and (x', x'), and suppose WLOG that $x < x'$. Then T cannot contain (x, x') so it is not a rectangle.*

Example 4 *One of the central functions in the study of communication complexity is the disjointness function $\mathrm{DISJ}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. This function is defined by $\mathrm{DISJ}_n(x, y) = 1$ iff for every index i, either $x_i = 0$ or $y_i = 0$. Equivalently, this function checks whether the sets for which x and y are indicator vectors are disjoint.*

We can prove a lower bound of n on $\mathbf{P^{cc}}(\mathrm{DISJ}_n)$ using the fooling set $S = \{(x, \bar{x}) : x \in \{0,1\}^n\}$. Note that $\mathrm{DISJ}_n(x, \bar{x}) = 1$ for every pair of inputs in S. To see that this is a fooling set, let T be a monochromatic set containing both $(x, \bar{x})$ and $(x', \bar{x}')$. Then either $(x, \bar{x}')$ or $(x', \bar{x})$ must intersect, so T cannot be a rectangle.

2 Rank

A two-party communication problem $f : X \times Y \to \{0,1\}$ naturally corresponds to a matrix M_f as follows. The rows of M_f are indexed by elements of X and the columns are indexed by elements of Y and $M_f[x, y] = f(x, y)$. Let $\mathrm{rank}(f) = \mathrm{rank}(M_f)$.

To understand the connection between rank and deterministic communication, we need two facts. The first is that combinatorial rectangles have low rank.

Lemma 5 *Let $A \times B \subset X \times Y$ be a non-empty combinatorial rectangle. Then the matrix M where $M[x, y] = 1$ if $(x, y) \in A \times B$ and $M[x, y] = 0$ otherwise has rank 1.*

Proof: Let $v_A \in \{0,1\}^{|X|}$ and $v_B \in \{0,1\}^{|Y|}$ be the indicator vectors for A and B, respectively. Then $M = v_A v_B^T$, and hence M has rank 1. ∎

The second is that rank is subadditive.

Fact 6 *Let M and N be matrices with the same dimensions. Then $\mathrm{rank}(M + N) \leq \mathrm{rank}(M) + \mathrm{rank}(N)$.*

Proof: Let C_M and C_N denote the sets of columns of M and N, respectively, so $\mathrm{rank}(M) = \dim \mathrm{span}(C_M)$ and $\mathrm{rank}(N) = \dim \mathrm{span}(C_N)$. The column span of $M + N$ is contained in the $\mathrm{span}(C_M \cup C_N)$, so $\mathrm{rank}(M + N) \leq \dim \mathrm{span}(C_M \cup C_N) \leq \mathrm{rank}(M) + \mathrm{rank}(N)$. ∎

Theorem 7 *For any function $f : X \times Y \to \{0, 1\}$, we have $\mathbf{P^{cc}}(f) \geq \log \mathrm{rank}(f)$.*

Proof: Let Π be a protocol computing f with cost c. For each leaf ℓ of Π on which the protocol outputs 1, let M_ℓ be the matrix with $R_\ell[x, y] = 1$ if (x, y) reaches ℓ and $R_\ell[x, y] = 0$ otherwise. Since M_ℓ is the indicator of a rectangle, $\mathrm{rank}(M_\ell) = 1$. Moreover, $M_f = \sum_\ell M_\ell$, so $\mathrm{rank}(M_f) \leq \sum_\ell \mathrm{rank}(M_\ell) \leq 2^c$. Hence $c \geq \log \mathrm{rank}(f)$. ∎

Example 8 *The equality and greater-than functions correspond to matrices of full rank 2^n, and hence have deterministic communication complexity $\geq n$.*

2.1 The Log Rank Conjecture

Deterministic communication can also be *upper* bounded by rank itself. To see this, suppose M_f has rank r. Then we can decompose $M_f = A \cdot B$ where A is an $|X| \times r$ *Boolean* matrix and Y is a $r \times |Y|$ *real* matrix. Observe that $f(x, y) = e_x^T M_f e_y = e_x^T A B e_y$, where e_x and e_y are the indicators for x and y respectively. So Alice can send Bob the r-dimensional Boolean vector $e_x^T A$ which suffices for him to compute $f(x, y)$.

There's an exponential gap between the best known bounds on $\mathbf{P^{cc}}$ in terms of rank: $\log \mathrm{rank}(f) \leq \mathbf{P^{cc}}(f) \leq \tilde{O}(\sqrt{\mathrm{rank}(f)})$. (The improved upper bound was obtained by Lovett in 2013.) Which bound is closer to the truth? Thirty years ago, Lovász and Saks conjectured that the lower bound is tight up to polynomial factors, namely that that there is a constant $c \geq 1$ such that $\log \mathrm{rank}(f) \leq \mathbf{P^{cc}}(f) \leq O(\log^c \mathrm{rank}(f))$. This is probably the most famous conjecture in communication complexity.

While we are far from resolving the log rank conjecture itself, we do know a lot about related formulations in other models. For instance, the log of the *sign*-rank of a matrix characterizes its $\mathbf{UPP^{cc}}$ complexity up to an additive constant. On the other hand, a recent breakthrough of Chattopadhyay, Mande, and Sherif showed an exponential gap between the log of the *approximate*-rank of a matrix and its $\mathbf{BPP^{cc}}$ complexity, refuting the "approximate log rank conjecture." More on these later in the course.

3 Covers and Nondeterminism

Recall that we can lower bound deterministic communication by showing that we need many monochromatic rectangles to partition the input space. A stronger statement (which could perhaps be easier to prove in practice for specific functions) is to show that we need many rectangles just to *cover* the 0-inputs or the 1-inputs to a function. More precisely let us define the 0-cover and 1-cover numbers as follows.

Definition 9 *For $b \in \{0,1\}$, let $C^b(f)$ be the minimum number of rectangles needed to cover the b-inputs to f. Namely, it is the smallest set of rectangles whose union equals $f^{-1}(b)$.*

Clearly $\log C^0(f)$ and $\log C^1(f)$ are both lower bounds on $\mathbf{P^{cc}}(f)$.

Example 10 *Let's look at the covering numbers for the Equality function. Recall that we used the fooling set method to show that $C^1(\mathrm{EQ}_n) \geq 2^n$. On the other hand, $C^0(\mathrm{EQ}_n) \leq 2n$. To see this, consider the set of rectangles of the form $R_{i,b} = \{(x,y) : x_i = b, y_i \neq b\}$.*

The above example points out an interesting property of the Equality function: It has low co-nondeterministic complexity. Namely, on an pair of inputs (x,y) for which $\mathrm{EQ}_n(x,y) = 0$, and a *proof* of length only $\log n$ capturing the index i on which $x_i \neq y_i$, Alice and Bob can verify that $x \neq y$ using only 2 bits of communication. On the other hand, the nondeterministic complexity of EQ is still high, as it is hard for Alice and Bob to verify that $x = y$ even with an $O(n)$-length proof.

This motivates us to study nondeterministic and co-nondeterministic protocols.

Definition 11 *A nondeterministic protocol is a collection of deterministic protocols $\{\Pi_w : w \in \{0,1\}^s\}$. Such a protocol computes a function f if*

- *For every $(x,y) \in f^{-1}(1)$ there* exists *a $w \in \{0,1\}^s$ such that $\Pi_w(x,y) = 1$*
- *For every $(x,y) \in f^{-1}(0)$,* for all *$w \in \{0,1\}^s$ we have $\Pi_w(x,y) = 0$.*

The cost of such a protocol is $s + \max_{w\in\{0,1\}^s} \mathrm{cost}(\Pi_w)$, and the nondeterministic communication complexity of a function f, denoted $\mathbf{NP^{cc}}(f)$, is the least cost of such a protocol computing f. Similarly, a co-nondeterministic protocol computes f if 1) for every $(x,y) \in f^{-1}(1)$, we have $\Pi_w(x,y) = 1$ for all $w \in W$ and 2) for every $(x,y) \in f^{-1}(0)$ there exists $w \in W$ such that $\Pi_w(x,y) = 0$. The co-nondeterministic communication complexity of f is denoted by $\mathbf{coNP^{cc}}(f)$.

Nondeterministic communication complexity is tightly characterized by covering numbers.

Theorem 12 *For every function $f : X \times Y \to \{0,1\}$, we have $\mathbf{NP^{cc}}(f) = \log C^1(f) + \Theta(1)$ and $\mathbf{coNP^{cc}}(f) = \log C^0(f) + \Theta(1)$.*

Proof: We prove this for $\mathbf{NP^{cc}}$. First, suppose f admits a nondeterministic protocol with witness size s and protocol length c. Then the 1-inputs to f are covered by a union of $2^s \cdot 2^c$ 1-monochromatic rectangles. Hence $C^1(f) \leq 2^{\mathbf{NP^{cc}}(f)}$.

Now we show how to design a nondeterministic protocol for f which has a small covering number. Let $\{R_w\}$ be a cover of the 1-inputs of f. Then $f(x,y) = 1$ iff there exists a w such that $(x,y) \in R_w$. So given w as a proof, Alice and Bob can verify whether $(x,y) \in R_w$ using 2 bits of communication, and hence $\mathbf{NP^{cc}}(f) \leq \lceil \log C^1(f) \rceil + 2$. ∎

Theorem 13 *For every function f, we have $\mathbf{P^{cc}}(f) \leq O(\log C^0(f) \cdot \log C^1(f)) \leq O(\mathbf{NP^{cc}}(f) \cdot \mathbf{coNP^{cc}}(f))$.*

Proof: We sketch an "algorithmic" proof of this result which is reminiscent of the protocol for Clique vs. Independent Set. Given a small 0-cover and a small 1-cover, we will design a protocol which uses binary search to try to find a 0-rectangle containing (x, y), thus proving that $f(x, y) = 0$. (If the protocol fails, then we can conclude that $f(x, y) = 1$.) In each of stages $i = 1, \ldots, \log C^0(f)$, let T_i be the set of "live" 0-rectangles which could still contain (x, y). In every round, Alice and Bob will transmit $O(\log C^1(f))$ bits to decrease the size of T_i by a factor of 2, giving a total of $O(\log C^0(f) \cdot \log C^1(f))$ bits of communication.

The way to do this pruning is as follows. If there is a 1-rectangle containing row x whose row set intersects at most half of the rows sets of the rectangles in T_i, then Alice can send its identity to Bob with $O(\log C^1(f))$ bits and the parties can remove the remaining half of T_i from consideration. Otherwise, if there is a 1-rectangle containing column y that intersects at most half of T_i in columns, Bob can send its identity to Alice. Note that a 0-rectangle containing x will always survive this pruning process. If neither Alice nor Bob can find such a 1-rectangle, then the parties can conclude that $f(x, y) = 0$. To see this, suppose instead that (x, y) were contained in a 1-rectangle R. Then by the pigeonhole principle, R could intersect at most half of T_i in rows and at most half of T_i in columns, so Alice or Bob would be able to prune.

■

Theorem 13 tells us that if f has both a good nondeterministic protocol and a good co-nondeterministic protocol, then it has a good deterministic protocol. One way to state this succinctly is as follows. By abusing notation, define the complexity classes $\mathbf{P^{cc}}, \mathbf{NP^{cc}}, \mathbf{coNP^{cc}}$ as the sequences of functions $\{f_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}\}$ which have $(\log n)^{O(1)}$-cost deterministic, nondeterministic, and co-nondeterministic protocols, respectively. Then Theorem 13 implies $\mathbf{P^{cc}} = \mathbf{NP^{cc}} \cap \mathbf{coNP^{cc}}$. Also note that Example 10 shows that $\mathbf{P^{cc}} \subsetneq \mathbf{coNP^{cc}}$, and by complementing, $\mathbf{P^{cc}} \subsetneq \mathbf{NP^{cc}}$.