Information Complexity vs. Communication Complexity: Hidden Layers Game

Jiahui Liu

Final Project Presentation for Information Theory in TCS
Introduction

- Review of IC vs CC
- Hidden Layers Game
- Upper Bound of IC
- Intuition for lower bound of CC
Communication Complexity and Information Complexity

Communication Complexity

$CC(f; \epsilon)$ is the smallest number of bits that Alice and Bob need to exchange to compute $f$ with error probability $\epsilon$.

Information Complexity

- The protocol $\pi$ on the pair of (random) inputs $(X; Y) \sim \mu$ gives the transcript $\Pi = \Pi(X; Y)$
- The information cost of a protocol $\Pi$ is the amount of information that the protocol reveals to Alice and Bob about their input.
- the amount revealed to Alice – who knows $X$ – about $Y$ is given by the conditional mutual information $I(Y; \Pi|X)$.

Information cost of $\Pi$ is given by:

$$IC_{\mu}(\pi) = I(Y; \Pi|X) + I(X; \Pi|Y)$$
Information Complexity

(continued) The task of finding information complexity of $f$ is the task of minimizing the information complexity of the protocol for $f$:

$$IC_{\mu} = \inf_{\text{protocol } \pi \text{ performing } f} IC_{\mu}(\pi)$$

IC vs CC
Information complexity can be viewed as the interactive analogue of Shannon's entropy. Equality between information and communication complexity is equivalent to compression theorem in the interactive setting: whether it is possible to compress an interactive conversation into its information content like we compress a single message.
A problem with large CC and small IC

Hidden Layers Game problem
Conjectured in [Bra13] with lower bound proved in [GKR16]
- $k$ is a parameter used in the problem
- IC upper bound = $O(\log k)$
- CC Lower Bound = $\Omega(k)$

Exponential separation!
In fact proved separation with external IC (stronger result)
Using embedding of set disjointness inputs.
Hidden Layers Game

Hidden Layers Game is a sampling problem.
Parameters:

- strings over an alphabet $\Sigma$ of size $k$
- another parameter $N = 2^n$. fix $N = 2^n = 2^{2^k}$.

Input:

- Alice and Bob are given a pair of numbers $a, b \sim \{0, ..., N - 1\}$ (i.e. two $n$-bit numbers that take $\Omega(\log n)$ communication to compare)
- Alice is given a uniformly random function $F_A : \Sigma^{2^a} \rightarrow \Sigma$ (the function is only known to Alice)
- Bob is given a uniformly random function $F_B : \Sigma^{2^b+1} \rightarrow \Sigma$ (the function is only known to Bob)
Hidden Layers Game

Alice and Bob need to sample a uniformly random string $s \in \Sigma^{2N}$ subject to the constraints:

- $s_{2a+1} = F_A(s_1...s_{2a})$
- $s_{2b+2} = F_B(s_1...s_{2b+1})$

In other words, they want a $2N$-symbol string over the alphabet, where its first $2a$-symbols substring can be mapped to its $(2a+1)$th symbol by Alice’s function and its first $(2b+1)$-symbols substring can be mapped to its $(2b+2)$th symbol by Bob’s function.
Naive Protocol and IC upper bound

Naive Protocol \( \pi_0 \) for hidden layers game \( H \)

- in odd rounds Alice samples the next symbol of \( s \) and in even rounds Bob does.
- In rounds \( i \neq 2a + 1 \), Alice just sends a uniformly random \( s_i \sim \Sigma \). In round \( i = 2a + 1 \), Alice computes and sends \( s_i = F_A(s_{1..2a}) \).
- Bob does the similar thing for even rounds.

Communication complexity = \( \Theta(N \log k) = \Theta(2^k \log k) \) (2N rounds, we can view the encoding of each char \( s_i \) as \( \log k \) since the alphabet has size \( k \)).
Naive Protocol and IC upper bound

$s$ is sampled uniformly from the subset $S$ of strings which satisfy the two constraints. The size of $S$, $|S| = k^{2N-2}$; uniformly random except $(2a + 1)$th and $(2b + 2)$th symbols. Thus the KL-divergence between $s$ and the uniform distribution on $\Sigma^{2N}$ is $2 \log k$. 
Naive Protocol and IC upper bound

The transcript of $\pi_0$ is distributed exactly as the output $s$ of $H$ given $a, F_A, b, F_B$.
Denote $\mu$ as the distribution of the inputs $a, F_A, b, F_B$ to $H$.

$$IC_\mu(H) = IC_\mu(\pi_0) = I_\mu(s; a, F_A, b, F_B)$$
$$= \mathbb{E} \mathbb{D}(s|a; F_A, b, F_B || s) = 2 \log k$$
CC Lower Bound Intuition 1: Disjointness

- A randomly selected string $t \in \Sigma^{2N}$ has a probability of exactly $1/k$ of being consistent with Alice’s input: $t$ just needs its $(2a + 1)$th symbol happen to satisfy $t_{2a+1} = F_A(t_{1..2a})$
  Same for Bob.

- So $t$ has probability $1/k^2$ to be consistent with both Alice and Bob
Protocol 1:
- Using public randomness and no communication; sample $k^2$ strings $s_1, \ldots, s_{k^2}$ drawn uniformly at random from $\Sigma^{2N}$.
- Let $A$ be the subset (of approximately $k$) strings satisfying Alice’s constraint.
- Let $B$ be the subset satisfying Bob’s constraint.
- Alice and Bob communicate to determine whether $A \cap B = \emptyset$; if not, they output the first element of $A \cap B$; otherwise they repeat the entire process.
CC Lower Bound Intuition 1: Disjointness

Correctness

- the first string in the intersection between A and B must satisfy distribution of s
- The probability that $A \cap B \neq \emptyset$ is approximately $1 - 1/e$, and the process will terminate after an expected constant number of iterations.

Lower bound

- CC upper bound for disjointness of two sets with size $k$ is $O(k)$.
- if we can reduce to Disjointness...
- CC lower bound for disjointness of two sets with size $k$ is $\Omega(k)$. 
CC Lower Bound Intuition 2: Greater Than (GT)

Protocol 2:

- find a $c$ such that $a \leq c \leq b$ or $a \geq c \geq b$, wlog assume that $a \leq c \leq b$
- Use public randomness and no communication, generate strings $t_1, t_2.. \in \Sigma^{2c+1}$ uniformly randomly
- Alice sends Bob the index of the first string $s_0$ satisfying her constraint
- Using public randomness and no communication, generate strings $r_1, r_2.. \in \Sigma^{2N}$ which extend $s_0$ with uniformly random symbols
- Bob sends Alice the index of the first string $s$ satisfying his constraint. This $s$ is the output.
The first step of comparing $a, b$ requires CC of the Greater Than function.
As discussed above, the probability of a string to be acceptable is $1/k$, and therefore communicating the index of the first acceptable string requires $O(\log k)$ bits.

**Lower bound for GT**

$a$ and $b$ are $n-bit$ numbers.

$n = \log N$. Therefore the lb for GT here is

$\Omega(\log n) = \Omega(\log \log N) = \Omega(k)$.
Sampling vs Decision?

- Can this sampling problem be generalized to a decision problem?
- If no such decision problem exists, what property makes protocols for decision problems easier to compress?
- Sampling problems require a lot of public randomness. Decision problems protocols: the answer is determined by messages sent by Alice and Bob.
Sampling vs Decision?

- Exponential separation of IC and CC for Boolean functions has been shown in [GKR15].
- With the introduction of a lower bound method: Relative discrepancy method.
- But is Relative discrepancy method to separate all boolean functions?
- Actually No! More in my report.
Questions?
Mark Braverman.
A hard-to-compress interactive task?

Anat Ganor, Gillat Kol, and Ran Raz.
Exponential separation of information and communication for boolean functions.

Anat Ganor, Gillat Kol, and Ran Raz.
Exponential separation of communication and external information.