CS 229r Information Theory in Computer Science	Mar 3, 2016
Lecture 12	
Lecturer: Madhu Sudan	*Scribe: Tianren Liu*

1 Today's Topic

- Direct Sum Problem
- Internal Information Cost

2 Direct Sum Problem

For any two party computation problem $f : \{0,1\}^\ell \times \{0,1\}^\ell \to \{0,1\}$, consider its *direct sum problem*

$$f^{\otimes n}(\{0,1\}^\ell \times \{0,1\}^\ell)^n \to \{0,1\}^n$$

Such that

$$f^{\otimes n}(x_1, y_1, \ldots, x_n, y_n) = (f(x_1, y_1), \ldots, f(x_n, y_n))$$

It's obvious that $\mathcal{CC}(f^{\otimes n}) \leq n \cdot \mathcal{CC}(f)$. It might seem that there are no better way to compute $f^{\otimes n}$ than compute each coordinate individually.

While there exists function f such that $\mathcal{CC}(f^{\otimes n}) \ll n \cdot \mathcal{CC}(f)$.

Aside: Computational Complexity

There exists function a f such that $T(f) \geq \ell^2/\log \ell$, while $T(f^{\otimes n}) \ll n \cdot T(f)$ for some n. $T(f)$ is the computional complexity of f, measured by time or circuit size.

For $x \in \{0,1\}^\ell$, let $f(x) = A_\ell x$. $\{A_\ell\}_{\ell=1}^\infty$ is a family of matrix. There exists a family of matrix such that $f(x)$ needs $\Omega(\ell^2/\log \ell)$ size circuits to compute. While its direct product $f^{\otimes n}$ can be speeded up by matrix multiplication.

Theorem 1 ([BBCR10]). *Informal, for all f, μ, $\mathcal{CC}_{\mu^n}(f^{\otimes n}) \gtrapprox \mathcal{CC}_\mu(f) \cdot \sqrt{n}$*

More precisely, we also need to consider the error probability.

$$\mathcal{CC}_{\mu^n,\varepsilon}(f^{\otimes n}) \geq \tilde{\Omega}(\mathcal{CC}_{\mu,\varepsilon}(f) \cdot \sqrt{n}).$$

Notice that the error probability preserves. Compare it with the naïve upper bound

$$\mathcal{CC}_{\mu^n,\varepsilon'}(f^{\otimes n}) \leq n \cdot \mathcal{CC}_{\mu,\varepsilon}(f)$$

where $1 - \varepsilon' = (1-\varepsilon)^n$.

Later work study the asymptotic behavior of the amortized communication, showing that the communication complexity to compute $f^{\otimes n}$ grows linearly.

Theorem 2 ([BR11]). *For all f, μ, ε,*

$$\lim_{n\to\infty} \frac{1}{n} \mathcal{CC}_{\mu^n,\varepsilon}(f^{\otimes n}) = \mathcal{IC}^{\text{int}}_{\mu,\varepsilon}(f)$$

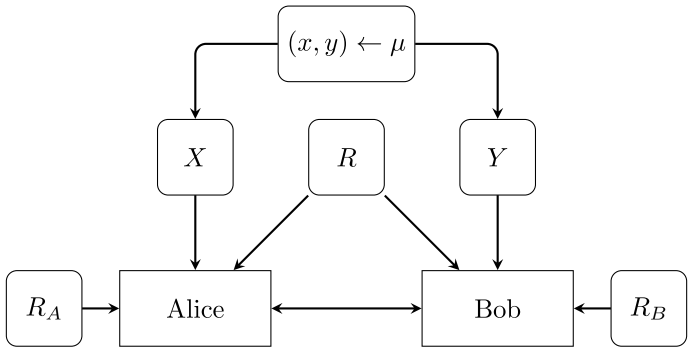
Moreover, in [BBCR10], they prove a stronger result for some functions. Let $f^{+n} : (\{0,1\}^\ell \times \{0,1\}^\ell)^n \to \{0,1\}^n$ denotes the parity of n outputs, or more generally, the sum of n outputs modulo K.

$$f^{+n}(x_1, y_1, \ldots, x_n, y_n) = \sum_{i=1}^n f(x_i, y_i).$$

f^{+n} output much less information then $f^{\otimes n}$, one might expect f^{+n} would be much easier to compute. While there exists function f (and distribution μ) such that $\mathcal{CC}_{\mu^n}(f^{+n}) \gtrapprox \mathcal{CC}_\mu(f) \cdot \sqrt{n}$.

3 Internal Information Cost

We use the same notations as previous lectures. The two-party computation scheme is $\Pi = \Pi(X, Y, R, R_A, R_B)$. Use capital letter to denote random variables. X is Alice's private input; Y is Bob's private input; R is common randomness; R_A (R_B) is the private randomness of Alice (Bob).



In previous lectures, we've discussed *external information cost* $\mathcal{IC}^{\text{ext}}(\Pi) = I(XY;\Pi|R)$, what can an external party learn from the transcript.

In this lecture, we consider *internal information cost*, $\mathcal{IC}^{\text{int}}(\Pi) = I(X;\Pi|YR) + I(Y;\Pi|XR)$, what each party can learn about each other's input by reading the transcript.

Definition 1 (internal information cost)**.** $\mathcal{IC}^{\text{int}}(\Pi) = I(X;\Pi|YR) + I(Y;\Pi|XR)$.

A natural definition of internal information cost should be $\mathcal{IC}^{\text{int}}(\Pi) = I(X;\Pi|YRR_B) + I(Y;\Pi|XRR_A)$. Notice that $I(X;\Pi|YR) = I(X;\Pi|YRR_B)$, this justifies our definition.

Claim. $\mathcal{IC}^{\text{int}}(\Pi) \leq \mathcal{IC}^{\text{ext}}(\Pi) \leq \mathcal{CC}(\Pi)$

Proof. Let Π is a k-bit transcript, then

$$I(X;\Pi|YR) = \sum_{i=1}^{k} I(\Pi_i; X|Y, R, \Pi_1 \ldots \Pi_{i-1})$$
$$I(Y;\Pi|XR) = \sum_{i=1}^{k} I(\Pi_i; Y|X, R, \Pi_1 \ldots \Pi_{i-1})$$
$$I(X,Y;\Pi|R) = \sum_{i=1}^{k} I(\Pi_i; X, Y|R, \Pi_1 \ldots \Pi_{i-1})$$

Let $Z_1, \ldots, Z_k \in \{a, b\}$ be random variables, Z_i is the party who send the i-th bit. By the constraint of two-party computation, Z_i is determined by $R, \Pi_1 \ldots \Pi_{i-1}$. Conditional on a assignment of $R = r, \Pi_1 \ldots \Pi_{i-1} = \pi_1 \ldots \pi_{i-1}$, w.o.l.g. assume $Z_i = b$ (Bob would send the i-th bit), then Bob learn nothing from the next bit as it's generated by him. Based on this intuition, it's easy to prove that $I(\Pi_i; X|Y, R, \Pi_1 \ldots \Pi_{i-1}, Z_i = b) = 0$.

$$\begin{aligned}
&I(\Pi_i; X, Y|R, \Pi_1 \ldots \Pi_{i-1}, Z_i = b) \\
=\ &I(\Pi_i; X|R, \Pi_1 \ldots \Pi_{i-1}, Z_i = b) + I(\Pi_i; Y|X, R, \Pi_1 \ldots \Pi_{i-1}, Z_i = b) \\
\geq\ &\underbrace{I(\Pi_i; X|Y, R, \Pi_1 \ldots \Pi_{i-1}, Z_i = b)}_{=0} + I(\Pi_i; Y|X, R, \Pi_1 \ldots \Pi_{i-1}, Z_i = b)
\end{aligned}$$

Similar inequality holds when conditional on $Z_i = a$. Then

$$\begin{aligned}
& I(\Pi_i; X, Y|R, \Pi_1 \dots \Pi_{i-1}) \\
= & \sum_{z\in\{a,b\}} \Pr[Z_i = z] I(\Pi_i; X, Y|R, \Pi_1 \dots \Pi_{i-1}, Z_i = b) \\
\geq & \sum_{z\in\{a,b\}} \Pr[Z_i = z] \Big(I(\Pi_i; X|Y, R, \Pi_1 \dots \Pi_{i-1}, Z_i = z) + I(\Pi_i; Y|X, R, \Pi_1 \dots \Pi_{i-1}, Z_i = z) \Big) \\
= & I(\Pi_i; Y|X, R, \Pi_1 \dots \Pi_{i-1}) + I(\Pi_i; X, Y|R, \Pi_1 \dots \Pi_{i-1})
\end{aligned}$$

Take the sum of both sides of the inequality for $i = 1, \dots, n$ finish the proof. ☐

4 Direct Sum Problem (Continued)

Informally, the following lemma shows that the (internal) information cost of direct sum $f^{\otimes n}$ is n times that of f.

Lemma 3. *If you have protocol for $f^{\otimes n}$ with information cost $\mathcal{I}$ and communication $\mathcal{C}$. Then you can get protocol for f with communication $\mathcal{C}$ and information cost $\leq \mathcal{I}/n$.*

The following lemma shows that if there is a long protocol has low information cost, it can be compressed.

Lemma 4. *If you have protocol for f with communication $\tilde{\mathcal{C}}$ and information cost $\tilde{\mathcal{I}}$. Then there exists a protocol for f with communication $O(\sqrt{\tilde{\mathcal{I}}\tilde{\mathcal{C}}} \log \tilde{\mathcal{C}})$.*

Suppose $\mathcal{CC}(f^{\otimes n}) = k$. Then Lemma 3 shows that there exists protocol Π' computing f such that $\mathcal{CC}(\Pi) \leq k$ and $\mathcal{IC}(\Pi) \leq \frac{k}{n}$. Then apply Lemma 4, $\mathcal{CC}(f) \leq \frac{k}{\sqrt{n}} \cdot \sqrt{\log k}$.

Proof of Lemma 3. Alice and Bob are given input x, y sampled from μ. They know a protocol Π that compute $f^{\otimes n}$. They want to use protocol the same protocol to solve the problem $f(x, y)$.

1. Pick a random location $j \in \{1, \dots, n\}$.
2. Construct input pair $(x_1, \dots, x_n), (y_1, \dots, y_n)$ such that $(x_j, y_j) = (x, y)$.
   For $i < j$, x_i is sampled from μ_X using public randomness, and y_i is sampled from $\mu_{Y|X=x_i}$ using Bob's private coins.
   For $i > j$, y_i is sampled from μ_Y using public randomness, and x_i is sampled from $\mu_{X|Y=y_i}$ using Alice's private coins.
3. Run the protocol $f^{\otimes n}$ and use the j-th bit of the output.

Denote above protocol by Π'. The communication complexity of Π' is the same as Π. The first term of the internal information cost of Π' is

$$\mathop{\mathbb{E}}_{j}\Big[I(X_j; \Pi|Y_j, R, j, X_1, \dots X_{j-1}, Y_{j+1} \dots Y_n) \Big]$$

We claim that it's no more than (in fact, equals to)

$$\frac{1}{n} \underbrace{I(X_1, \dots, X_n; \Pi|Y_1 \dots Y_n, R)}_{\text{first term of } \mathcal{IC}^{\text{int}}(\Pi)}.$$

$$\begin{aligned}
& \mathop{\mathbb{E}}_{j}\Big[I(X_j; \Pi|Y_j, R, j, X_1, \dots X_{j-1}, Y_{j+1} \dots Y_n) \Big] \\
= & \frac{1}{n} \sum_{j=1}^{n} I(X_j; \Pi|X_1 \dots X_{j-1}, Y_1 \dots Y_n, R) \\
= & \frac{1}{n} I(X_1, \dots, X_n; \Pi|Y_1 \dots Y_n, R)
\end{aligned}$$

Similar equality holds for the second term of the internal information cost of Π', Π. Thus

$$\mathcal{IC}^{\text{int}}(\Pi') = \frac{1}{n}\mathcal{IC}^{\text{int}}(\Pi) \qquad \square$$

References

[BBCR10] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. In Leonard J. Schulman, editor, *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 67–76. ACM, 2010.

[BR11] Mark Braverman and Anup Rao. Information equals amortized communication. *CoRR*, abs/1106.3595, 2011.