COMP760, SUMMARY OF LECTURE 19.

HAMED HATAMI

1. Braverman's $2^{O(I)}$ compression

Consider a protocol π, and fix the public randomness to $R = r$. Recall that for an input (x, y) the probability that a leaf t will be the transcript of the protocol for a given pair of inputs x, y is given by

$$\Pr_{R_A,R_B}[\Pi_{xy} = t] = \left(\prod_{\substack{i\in[\ell-1] \\ \text{Alice owns } v_i}} p^x_{v_i}(v_{i+1}) \right) \times \left(\prod_{\substack{i\in[\ell-1] \\ \text{Bob owns } v_i}} p^y_{v_i}(v_{i+1}) \right) = p^x_A(t)p^y_B(t),$$

where $(v_1, v_2, \ldots, v_\ell(= t))$ denote the unique path from the root to the leaf t. Note that in general $p^x_A(t)$ is known only to Alice and $p^y_B(t)$ is known only to Bob. However Bob has an estimate of $p^x_A(t)$ and Alice has an estimate of $p^y_B(t)$. Namely

$$q^x_B(t) = \prod_{\substack{i\in[\ell-1] \\ \text{Bob owns } v_i}} q^x_{v_i}(v_{i+1})$$

is Alice's estimate of $p^y_B(t)$ and

$$q^y_A(t) = \prod_{\substack{i\in[\ell-1] \\ \text{Alice owns } v_i}} q^y_{v_i}(v_{i+1})$$

is Bob's estimate of $p^x_A(t)$.

Hence Alice's estimate of $p^x_A(t)p^y_B(t)$ is $p^x_A(t)q^x_B(t)$, and Bob's estimate is $q^y_A(t)p^y_B(t)$. The expected divergence betweens the actual probability distribution on the leaves and the estimates of Alice and Bob adds up to the information cost.

Proposition 1. *We have*

$$\mathrm{IC}_\mu(\pi_r) = \mathbb{E}_{xy\in\mu}\left[D(p^x_A p^y_B \parallel p^x_A q^x_B) + D(p^x_A p^y_B \parallel q^y_A p^y_B)\right].$$

Proof. Using the formula

$$I(A; B|C) = \mathbb{E}_{ac\sim AC}\left[D\left(B|_{A=a,C=c} \parallel B|_{C=c}\right)\right]$$

we have

$$I(X; \Pi|Y) = \mathbb{E}_{xy\sim\mu} D(\Pi|_{XY=xy} \parallel \Pi|_{Y=y}) = \mathbb{E}_{xy\sim\mu} D(p^x_A p^y_B \parallel q^y_A p^y_B),$$

and similarly

$$I(Y; \Pi|X) = \mathbb{E}_{xy\sim\mu} D(\Pi|_{XY=xy} \parallel \Pi|_{X=x}) = \mathbb{E}_{xy\sim\mu} D(p^x_A p^y_B \parallel p^x_A q^x_B).$$

□

1.1. **The protocol.** To simplify the notation we will assume $\epsilon \leq 1/4$, and $I := \mathrm{IC}_\mu(\pi) \geq 1$ (otherwise in the following, we have to replace ϵ with $\epsilon/4$ and I with $I+1$).

Let $\mathcal{U}$ be the set of all leaves. Firstly, Alice and Bob use public randomness to generate a sequence of points (t_i, α_i, β_i) each chosen independently and uniformly from $\mathcal{U} \times [0,1] \times [0,1]$. They only consider $T = 2|\mathcal{U}| \ln 1/\epsilon$ such points. Their goal is to find the first index j such that $\alpha_j \leq p_A^x(t_j)$ and $\beta_j \leq p_B^y(t_j)$. Note that then $\Pr[t_j = t] = p_A^x(t)p_B^y(t)$ for every $t \in \mathcal{U}$. Hence indeed t_j has the correct distribution and this would give a simulation of the original protocol. The challenge is that Alice can only verify the condition for α_i's and Bob can only verify the condition for β_i's. Note that the probability that such a $j \leq T$ does not exist is equal to

$$(1 - 1/|\mathcal{U}|)^T < e^{-T/|\mathcal{U}|} < \epsilon/4.$$

Alice constructs her set of candidates for j:

$$\mathcal{C}_A := \{i \leq T \ : \ \alpha_i \leq p_A^x(t_i),\ \beta_i \leq K q_B^x(t_i)\},$$

and Bob constructs his candidates

$$\mathcal{C}_B := \{i \leq T \ : \ \alpha_i \leq K q_B^y(t_i),\ \beta_i \leq p_A^y(t_i)\},$$

where $K = 2^{8I/\epsilon}$. Namely the both leave a multiplicative factor of K as a possible margin of error for the probabilities that they don't know exactly, and only have a estimate for.

Claim 2. *For a random $xy \sim \mu$, the probability that j does not belong to both $\mathcal{C}_A$ and $\mathcal{C}_B$ is at most $\epsilon/4$.*

Proof. Fix x and y. Since t_j has distribution $p_A^x p_B^y$, we have

$$\begin{aligned}\Pr[j \notin \mathcal{C}_A] &= \Pr[K q_B^x(t_j) < \beta_j] \leq \Pr[K q_B^x(t_j) < p_B^y(t_j)] \leq \frac{1}{\log K}\mathbb{E}\log\frac{p_B^y(t_j)}{q_B^x(t_j)} \\ &= \frac{1}{\log K}\mathbb{E}\log\frac{p_A^x(t_j)p_B^y(t_j)}{p_A^x(t_j)q_B^x(t_j)} = \frac{1}{\log K} D(p_A^x p_B^y \parallel p_A^x q_B^x).\end{aligned}$$

Similarly

$$\Pr[j \notin \mathcal{C}_B] = \frac{1}{\log K} D(p_A^x p_B^y \parallel q_A^y p_B^y).$$

Hence taking the expectation with respect to $xy \sim \mu$, Proposition 1 implies

$$\Pr_{r, xy\sim\mu}\Pr[j \notin \mathcal{C}_A \cap \mathcal{C}_B] \leq \frac{1}{\log K}\mathrm{IC}_\mu(\pi) = \frac{I}{8I/\epsilon} \leq \frac{\epsilon}{4}.$$

□

Note that in the likely case that j belongs to both $\mathcal{C}_A$ and $\mathcal{C}_B$, it will in fact be the first index in $\mathcal{C}_A \cap \mathcal{C}_B$. Hence their task now reduces to finding the first element in the intersection of their sets. To this end, Alice sends many hash values for each element in her set of candidates $\mathcal{C}_A$, and Bob finds the first element $b \in \mathcal{C}_B$ that matches the hash values of one of the elements in $C_{\mathcal{A}}$, and and declares that this is the index j that they were seeking. Note that he cannot just simply send the value of j to Alice as that would requite $\log |T|$ bits of communication, which can be very large. Howeve Bob can simply send k to Alice if the hash values of b matched the hash values of the k-th element of $\mathcal{C}_A$.

- They generate a sequence of points $(t_i, \alpha_i, \beta_i)_{i \leq T}$ each chosen independently and uniformly from $\mathcal{U} \times [0,1] \times [0,1]$.

- Alice and Bob create the sets $\mathcal{C}_A$ and $\mathcal{C}_B$ as above. If any of these sets is larger than $2^{9I/\epsilon}$, they declare fail.
- For each $a \in \mathcal{C}_A$, Alice sends the value of $d = 30I/\epsilon$ random public hash function $h_1(a), \ldots, h_d(a)$ to Bob.
- Bob finds the smallest index $b \in \mathcal{C}_B$ whose hash values $h_1(b), \ldots, h_d(b)$ match the hash values of one of the elements in $\mathcal{C}_A$. He declares $j = b$ and sends to Alice the place of j in $\mathcal{C}_A$.

1.2. **Analysis.** Let $\mathcal{E}_1$ be the "bad" event that the index $j \leq T$ does not exist, let $\mathcal{E}_2$ be the bad event that $j \notin \mathcal{C}_A \cap \mathcal{C}_B$, let $\mathcal{E}_3$ be the bad event that $\max(|\mathcal{C}_A|, |\mathcal{C}_B|) > 2^{9I/\epsilon}$, and let $\mathcal{E}_4$ be the bad event that Bob's declared $j = b$ does not belong to $\mathcal{C}_A$ (i.e. the hash function detected incorrectly some $b \notin \mathcal{C}_A$). Note that

$$\Pr[\text{success}] \geq 1 - \Pr[\mathcal{E}_1 \vee \mathcal{E}_2 \vee \mathcal{E}_3 \vee \mathcal{E}_4].$$

We have already shown

$$\Pr[\mathcal{E}_1] \leq \frac{\epsilon}{4},$$

and

$$\Pr[\mathcal{E}_2] \leq \frac{\epsilon}{4}.$$

Also by Markov's inequality

$$\Pr[\mathcal{E}_3] \leq \frac{\mathbb{E}[|\mathcal{C}_A|]}{2^{9I/\epsilon}} + \frac{\mathbb{E}[|\mathcal{C}_B|]}{2^{9I/\epsilon}} = \frac{K}{2^{9I/\epsilon}} + \frac{K}{2^{9I/\epsilon}} \leq \frac{\epsilon}{4}.$$

and finally since the protocol guarantees that $|\mathcal{C}_A|, |\mathcal{C}_B| \leq 2^{9I/\epsilon}$ if Alice and Bob get to the stage of the protocol where the hash values are compared, we have

$$\begin{aligned} \Pr[\mathcal{E}_4] &\leq \Pr[\exists a \in \mathcal{C}_A, b \in \mathcal{C}_B, a \neq b, h_i(a) = h_i(b)\ \forall i = 1, \ldots, d] \\ &\leq 2^{9I/\epsilon} \times 2^{9I/\epsilon} \times 2^{-30I/\epsilon} \leq \epsilon/4. \end{aligned}$$

We conclude

$$\Pr[\text{success}] \geq 1 - \epsilon.$$

Finally the number of communicated bits is

$$2^{9I/\epsilon} d + \log |\mathcal{C}_A| \leq 2^{10I/\epsilon}.$$

2. The AND function

In the next few sections we will study the 0-error information complexity of the AND function (of two bits).

2.1. The external information cost of AND.

Theorem 3. *We have*

$$\mathrm{IC}^{\mathrm{ext}}(\mathrm{AND}) := \max_{\mu} \mathrm{IC}^{\mathrm{ext}}_{\mu}([\mathrm{AND}, \mu, 0]) = \log 3.$$

Proof. First we prove the upper-bound. Consider the following trivial protocol:

- Alice sends X Bob.
- If $X = 1$, then Bob sends Y to Alice.

Obviously in the end, both Alice and Bob know the value of $X \wedge Y$. Note that the transcript Π is either 0, 00 or 01. Hence

$$\mathrm{IC}^{\mathrm{ext}}_{\mu}(\pi) = I(\Pi; XY) \leq H(\Pi) \leq \log 3,$$

as Π is distributed over three elements.

To prove the lower-bound consider the measure μ with $\mu(01) = \mu(10) = \mu(11) = \frac{1}{3}$, and $\mu(00) = 0$. Let π be a protocol that computes the AND function with 0-error under μ, and let Π denote its transcript. Then because of the rectangle property

$$\Pr_{R_A,R_B}[\Pi_{xy} = t] = p^x_A(t)p^y_B(t),$$

the value of t uniquely determines the value of XY. Indeed if

$$\Pr[\Pi_{01} = t], \Pr[\Pi_{10} = t] > 0$$

then $\Pr[\Pi_{11} = t] > 0$, and this would contradict the assumption that π is a 0-error protocol. Moreover obviously 11 and 01 (or 10) cannot lead to the same leaf as the value of the AND function is different for them.

Since Π determines the value of xy, we have

$$H(\Pi; XY) = H(XY) = \log 3.$$

□

References

School of Computer Science, McGill University, Montréal, Canada
E-mail address: hatami@cs.mcgill.ca