CMPSCI 711: More Advanced Algorithms

Lowerbounds 2: Information Statistics

Andrew McGregor

Last Compiled: March 27, 2018

Information Statistics Approach

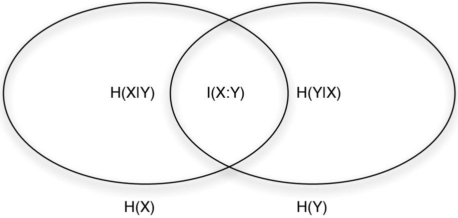
- Information statistics approach is based on analyzing the "information revealed" about the input from the messages.
- Useful for proving bounds on complicated functions in terms of simpler problems, e.g., proving a bound on

$$\mathrm{DISJ}_t(M) = \bigvee_{j \in [n]} \mathrm{AND}_t(M_{1,j}, \ldots, M_{t,j})$$

  by first establishing a bound on AND_t.
- We'll first give some definitions and then run through an example.

Information Theory Definitions

- Let X and Y be random variables.
- *Entropy:* $H(X) := \sum_i -\mathbb{P}[X = i] \lg \mathbb{P}[X = i]$
- *Conditional Entropy:* $H(X|Y) := \mathbb{E}_{y \sim Y}[H(X|Y = y)] \leq H(X)$
- *Mutual Information:* $I(X : Y) = H(X) - H(X|Y)$



- Useful Facts:
  - If X takes at most 2^ℓ values, then $H(X) \leq \ell$.
  - Chain rule: $H(XY) = H(X) + H(Y|X)$.
  - Subadditivity: $H(XY) \leq H(X) + H(Y)$; equality if independent.

Mutual Information

Lemma

If X and Y are independent, then $I(XY:Z) \geq I(X:Z) + I(Y:Z)$.

Proof.

$$\begin{aligned} I(XY:Z) &= H(XY) - H(XY|Z) \\ &= H(X) + H(Y) - H(XY|Z) \\ &\geq H(X) + H(Y) - H(X|Z) - H(Y|Z) \\ &= I(X:Z) + I(Y:Z) \end{aligned}$$

□

Information Cost

- Suppose you have a protocol Π for a two-party communication problem P in which Alice and Bob have random inputs X and Y.
- Let M be the (random) message sent by Alice and define:

$$\mathrm{cost}(\Pi) = \max |M|$$

and

$$\mathrm{icost}(\Pi) = I(M : X)$$

- Note $\mathrm{icost}(\Pi) = I(M : X) \leq H(M) \leq \mathrm{cost}(\Pi)$.

Example: Indexing

- We'll prove a lower bound on the information cost of $\textsc{Index}$ where $X \in_R \{0,1\}^n$ in terms a simpler problem "$\textsc{Echo}$"
- $\textsc{Echo}$: Alice has a single bit $B \in_R \{0,1\}$ and Bob wants to output B with probability at least $1-\delta$.
- A protocol $\Pi_{\textsc{Index}}$ for $\textsc{Index}$ yields a protocol $\Pi_{\textsc{Echo},i}$ for $\textsc{Echo}$:
  1. Given B, Alice picks $X_j \in_R \{0,1\}$ for $j \neq i$ and generates:

     $$X = (X_1, X_2, \ldots, X_{i-1}, B, X_{i+1}, \ldots, X_n)$$

  2. She sends the message M she'd have sent in $\Pi_{\textsc{Index}}$ if she'd had X.
  3. Bob receives message and outputs the value he'd have returned in $\Pi_{\textsc{Index}}$ had his input been i.

Relating Information Cost of $\textsc{Index}$ and $\textsc{Echo}$

- Since $X_1, X_2, \ldots, X_n$ are independent:

$$\begin{aligned} \mathsf{cost}(\Pi_{\textsc{Index}}) &\geq \mathsf{icost}(\Pi_{\textsc{Index}}) \\ &= I(X_1 X_2 \ldots X_n : M) \\ &\geq I(X_1 : M) + I(X_2 : M) + \ldots + I(X_n : M) \\ &= \mathsf{icost}(\Pi_{\textsc{Echo},1}) + \mathsf{icost}(\Pi_{\textsc{Echo},2}) + \ldots + \mathsf{icost}(\Pi_{\textsc{Echo},n}) \end{aligned}$$

- *Lemma:* Any protocol solving $\textsc{Echo}$ with probability $\geq 1 - \delta$, needs

$$\mathsf{icost}(\Pi_{\textsc{Echo},i}) \geq 1 - H_2(\delta)$$

where $H_2(p) = -p \lg p - (1-p) \lg(1-p)$.

- Hence, $\mathsf{cost}(\Pi_{\textsc{Index}}) \geq (1 - H_2(\delta))n$.

Proof of Lemma

1. Fano's inequality: Let A and B be random variables. If you can guess B correctly with probability at least $1 - \delta$ given A, then

$$H(B|A) \leq H_2(\delta) \ .$$

2. Let $A = M$ be message and B be the bit needing echoed.
3. Hence,

$$\text{icost}(\Pi_{\text{ECHO}}) = H(B) - H(B|M) \geq 1 - H_2(\delta)$$

Outline for $\textsc{disj}_t$ Lower Bound

- Express $\textsc{disj}_t$ in terms of $\textsc{and}_t$ where $\textsc{and}_t(x_1, \ldots, x_t) = \prod_i x_i$:

$$\textsc{disj}_t(M) = \bigvee_{j \in [n]} \textsc{and}_t(M_{1,j}, \ldots, M_{t,j})$$

- Consider a random input M to $\textsc{disj}_t$ where $M_{D_j j} \in_R \{0,1\}$ for $D_j \in_R [t]$. All other entries are 0.
- Let $T = (T_1, \ldots, T_{t-1})$ be the messages sent in a t-party protocol and define the information cost of a protocol as:

$$\text{icost}(\Pi|D) = I(T : M|D) \quad \text{where} \quad D = (D_1, \ldots, D_t) \ .$$

- A protocol for $\textsc{disj}_t$ yields n different protocols $\Pi_{\textsc{and}_t,i}$ for $\textsc{and}_t$:

$$\text{icost}(\Pi_{\textsc{disj}_t}|D) \geq \sum_{i \in [n]} \text{icost}(\Pi_{\textsc{and}_t,i}|D) \ .$$

- Result follows by showing $\text{icost}(\Pi_{\textsc{and}_t,i}|D) = \Omega(1/t)$.