Optimal Lossless Source Codes for Timely Updates

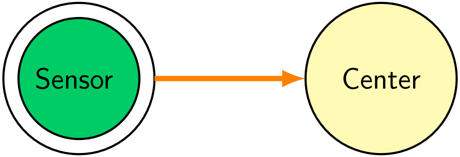
Prathamesh Mayekar

Joint work with

Parimal Parag and Himanshu Tyagi

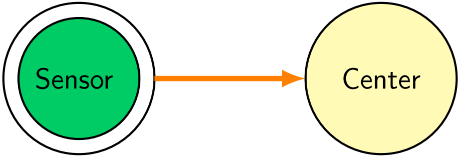
Department of ECE,
Indian Institute of Science

Motivation



Motivation

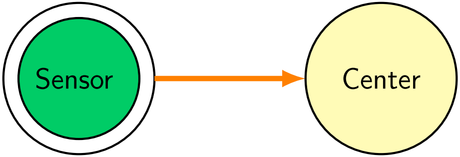
Source - The Hindu



Motivation

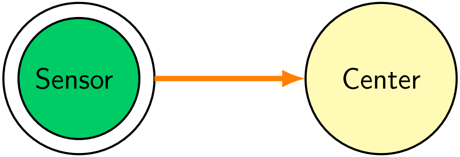
Source - The Hindu

Timely Updates are critical.



Motivation

Source - The Hindu



Timely Updates are critical.

Age of Information[1] - metric to capture timeliness.

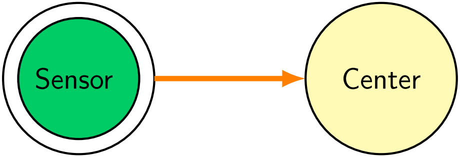
[1] Kaul, S., Yates, R., and Gruteser, M. (2011, December). On piggybacking in vehicular networks. In Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE (pp. 1-5). IEEE.

Age of Information (AOI) - Metric for Timeliness

- AOI: Time lag between the latest information at the RX w.r.t. that at TX.

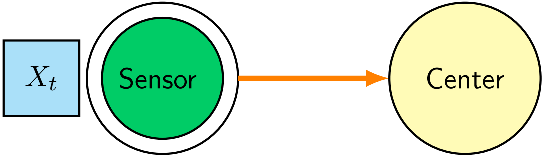
Age of Information (AOI) - Metric for Timeliness

- AOI: Time lag between the latest information at the RX w.r.t. that at TX.



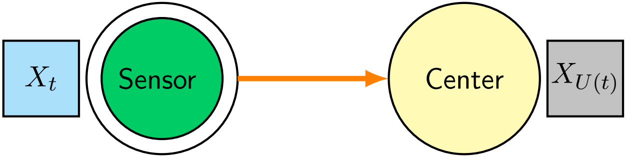
Age of Information (AOI) - Metric for Timeliness

- AOI: Time lag between the latest information at the RX w.r.t. that at TX.



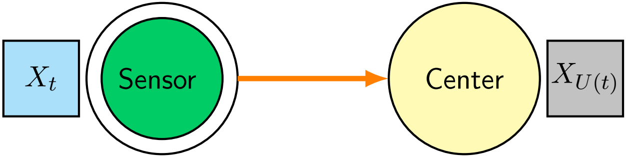
Age of Information (AOI) - Metric for Timeliness

- AOI: Time lag between the latest information at the RX w.r.t. that at TX.



Age of Information (AOI) - Metric for Timeliness

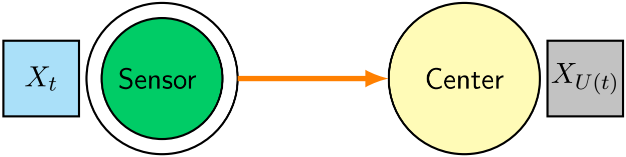
- AOI: Time lag between the latest information at the RX w.r.t. that at TX.



$$A(t) = t - U(t).$$

Age of Information (AOI) - Metric for Timeliness

- AOI: Time lag between the latest information at the RX w.r.t. that at TX.



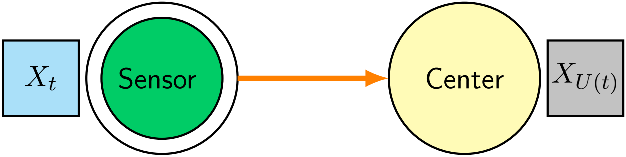
$$A(t) = t - U(t).$$

- We are interested in minimizing the average age

$$\bar{A} \triangleq \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} A(t).$$

Age of Information (AOI) - Metric for Timeliness

- AOI: Time lag between the latest information at the RX w.r.t. that at TX.



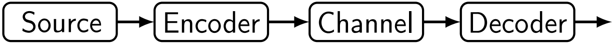
$$A(t) = t - U(t).$$

- We are interested in minimizing the average age

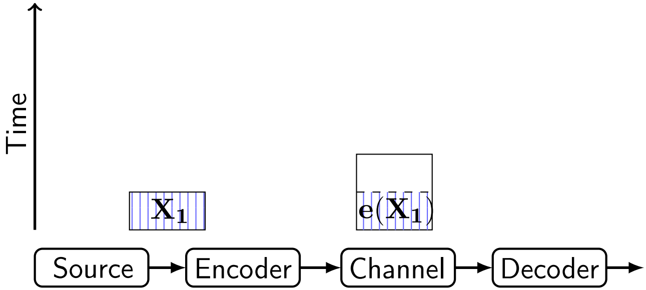
$$\bar{A} \triangleq \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} A(t).$$

- We restrict to Memoryless Update Schemes.

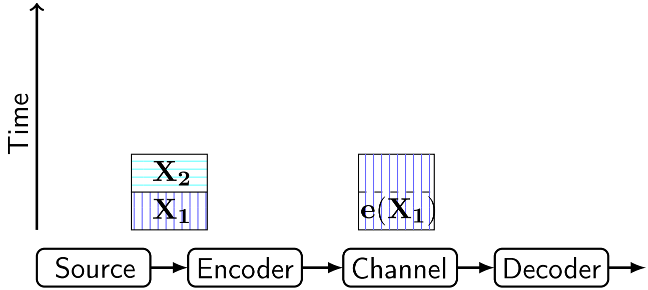
Memoryless Update Schemes



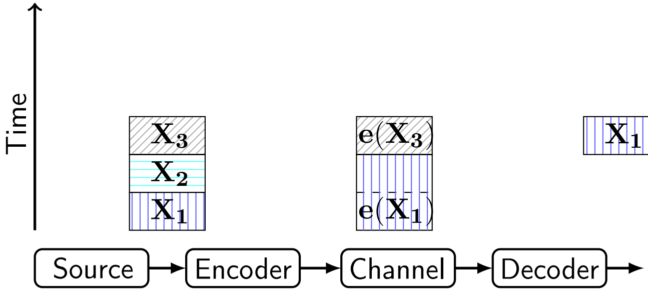
Memoryless Update Schemes



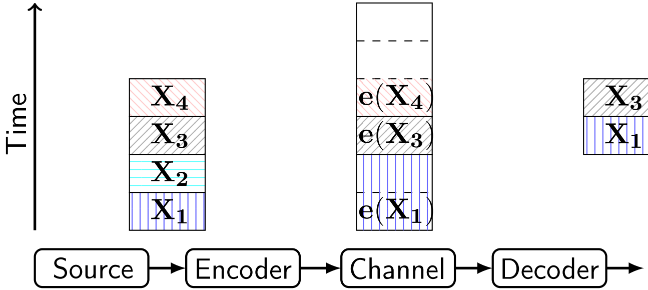
Memoryless Update Schemes



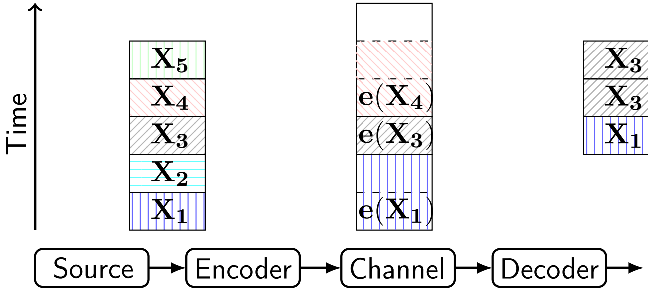
Memoryless Update Schemes



Memoryless Update Schemes



Memoryless Update Schemes



Memoryless Update Schemes

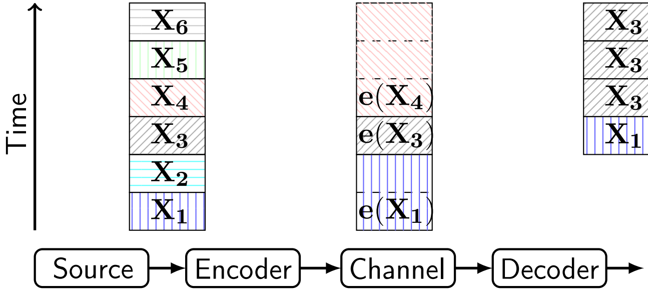
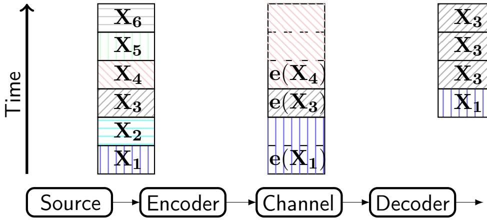


Illustration of Instantaneous Age



$$A(t) = t - U(t)$$

U(t)=Index of latest information at the decoder

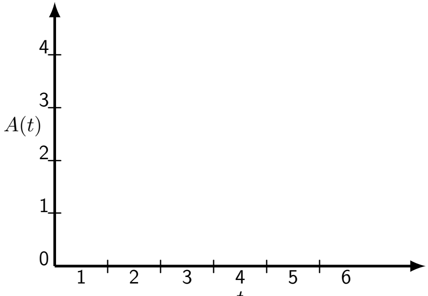
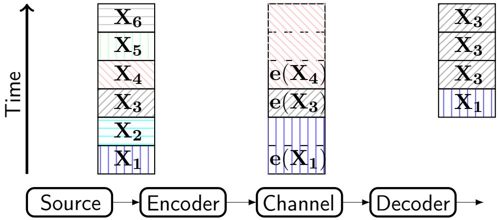


Illustration of Instantaneous Age



$$A(t) = t - U(t)$$

U(t)=Index of latest information at the decoder

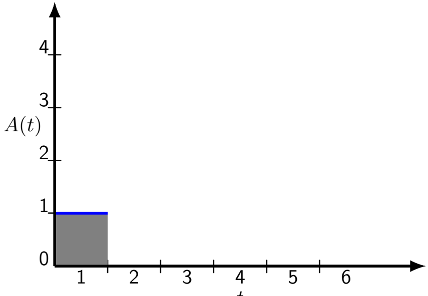
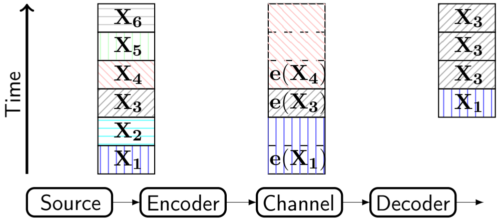


Illustration of Instantaneous Age



$$A(t) = t - U(t)$$

U(t)=Index of latest information at the decoder

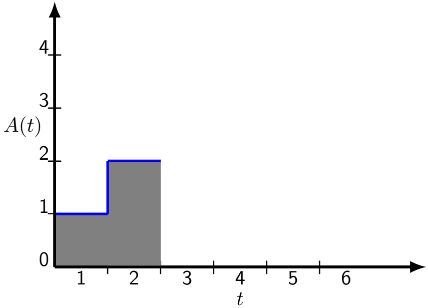
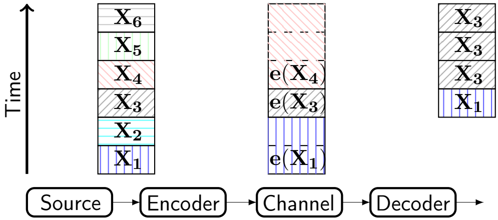


Illustration of Instantaneous Age



$$A(t) = t - U(t)$$

U(t)=Index of latest information at the decoder

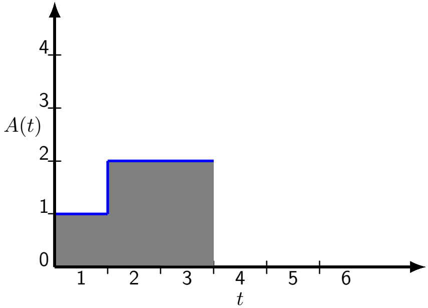
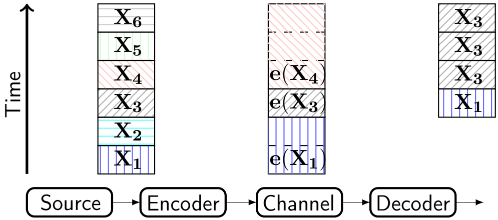


Illustration of Instantaneous Age



$$A(t) = t - U(t)$$

U(t)=Index of latest information at the decoder

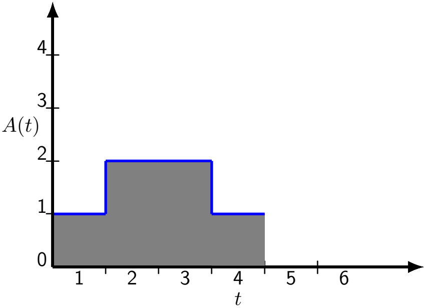
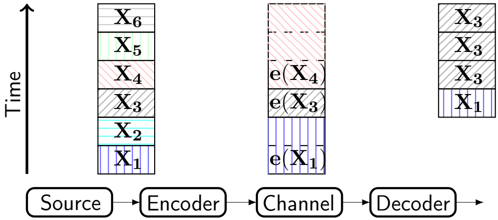
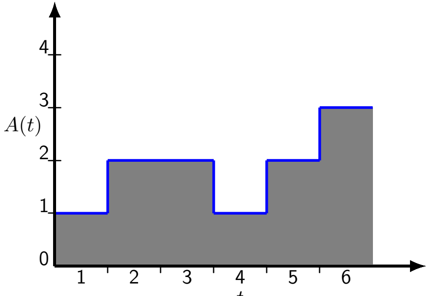


Illustration of Instantaneous Age



$$A(t) = t - U(t)$$

U(t)=Index of latest information at the decoder



Characterization of Average Age

$$\bar{A}(e) \triangleq \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} A(t)$$

$\ell(x) \triangleq$ code-length for a symbol x, $L \triangleq \ell(X)$.

Characterization of Average Age

$$\bar{A}(e) \triangleq \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} A(t)$$

$\ell(x) \triangleq$ code-length for a symbol x, $L \triangleq \ell(X)$.

Theorem

For a prefix-free code e, $\bar{A}(e) = \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} - \frac{1}{2} \quad a.s..$

Characterization of Average Age

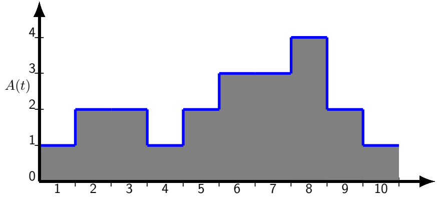
$$\bar{A}(e) \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} A(t)$$

$\ell(x) \triangleq$ code-length for a symbol x, $L \triangleq \ell(X)$.

Theorem

For a prefix-free code e, $\bar{A}(e) = \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}[L]} - \frac{1}{2} \quad a.s.$.

Proof Idea:



Characterization of Average Age

$$\bar{A}(e) \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} A(t)$$

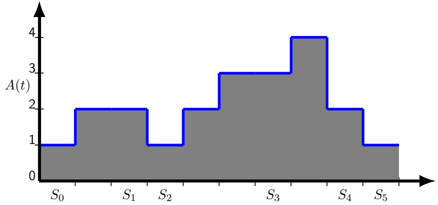
$\ell(x) \triangleq$ code-length for a symbol x, $L \triangleq \ell(X)$.

Theorem

For a prefix-free code e, $\bar{A}(e) = \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}[L]} - \frac{1}{2} \quad a.s.$.

Proof Idea:

- $S_i \triangleq i^{th}$ reception
  $S_{i+1} - S_i$ is iid



Characterization of Average Age

$$\bar{A}(e) \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} A(t)$$

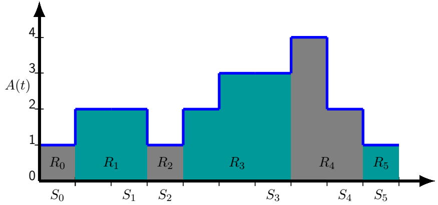
$\ell(x) \triangleq$ code-length for a symbol x, $L \triangleq \ell(X)$.

Theorem

For a prefix-free code e, $\bar{A}(e) = \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} - \frac{1}{2} \quad a.s..$

Proof Idea:

- $S_i \triangleq i^{th}$ reception
  $S_{i+1} - S_i$ is iid
- $(R_{2i+1})_{i\in\mathbb{N}}$ is iid,
  $(R_{2i+2})_{i\in\mathbb{N}}$ is iid



Characterization of Average Age

$$\bar{A}(e) \triangleq \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} A(t)$$

$\ell(x) \triangleq$ code-length for a symbol x, $L \triangleq \ell(X)$.

Theorem

For a prefix-free code e, $\bar{A}(e) = \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}[L]} - \frac{1}{2} \quad a.s.$.

Which source coding scheme is optimal?

Are Shannon Codes Optimal?

Shannon code for P: $\ell(x) = \lceil -\log P(x) \rceil \quad \forall x$.

Are Shannon Codes Optimal?

Shannon code for P: $\ell(x) = \lceil -\log P(x) \rceil \quad \forall x$.

Lemma

Given a pmf P on $\mathcal{X}$, a Shannon code e for P has average age at most $O(\log |\mathcal{X}|)$.

Are Shannon Codes Optimal?

But Shannon codes can be far from optimal!

Are Shannon Codes Optimal?

But Shannon codes can be far from optimal!

Example: Consider $\mathcal{X} = \{0, ..., 2^n\}$ and a pmf P on $\mathcal{X}$ given by

$$P(x) = \begin{cases} 1 - \frac{1}{n}, & x = 0 \\ \frac{1}{n2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Are Shannon Codes Optimal?

But Shannon codes can be far from optimal!

Example: Consider $\mathcal{X} = \{0, ..., 2^n\}$ and a pmf P on $\mathcal{X}$ given by

$$P(x) = \begin{cases} 1 - \frac{1}{n}, & x = 0 \\ \frac{1}{n2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Shannon codes for P have an average age of $\Omega(\log|\mathcal{X}|)$.

Are Shannon Codes Optimal?

But Shannon codes can be far from optimal!

Example: Consider $\mathcal{X} = \{0, ..., 2^n\}$ and a pmf P on $\mathcal{X}$ given by

$$P(x) = \begin{cases} 1 - \frac{1}{n}, & x = 0 \\ \frac{1}{n2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Shannon codes for P have an average age of $\Omega(\log |\mathcal{X}|)$.

Instead, use Shannon codes for pmf $P'(x)$, where

$$P'(x) = \begin{cases} \frac{1}{2^{\sqrt{n}}}, & x = 0 \\ \frac{1-2^{-\sqrt{n}}}{2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Are Shannon Codes Optimal?

But Shannon codes can be far from optimal!

Example: Consider $\mathcal{X} = \{0, ..., 2^n\}$ and a pmf P on $\mathcal{X}$ given by

$$P(x) = \begin{cases} 1 - \frac{1}{n}, & x = 0 \\ \frac{1}{n2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Shannon codes for P have an average age of $\Omega(\log |\mathcal{X}|)$.

Instead, use Shannon codes for pmf $P'(x)$, where

$$P'(x) = \begin{cases} \frac{1}{2^{\sqrt{n}}}, & x = 0 \\ \frac{1-2^{-\sqrt{n}}}{2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Shannon codes for P' have an average age of $O(\sqrt{\log |\mathcal{X}|})$.

Are Shannon Codes Optimal?

But Shannon codes can be far from optimal!

Example: Consider $\mathcal{X} = \{0, ..., 2^n\}$ and a pmf P on $\mathcal{X}$ given by

$$P(x) = \begin{cases} 1 - \frac{1}{n}, & x = 0 \\ \frac{1}{n2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Shannon codes for P have an average age of $\Omega(\log |\mathcal{X}|)$.

Instead, use Shannon codes for pmf $P'(x)$, where

$$P'(x) = \begin{cases} \frac{1}{2^{\sqrt{n}}}, & x = 0 \\ \frac{1-2^{-\sqrt{n}}}{2^n}, & x \in \{1, \ldots, 2^n\}. \end{cases}$$

Shannon codes for P' have an average age of $O(\sqrt{\log |\mathcal{X}|})$.

Shannon codes are order-wise suboptimal!

Our Approach

Reduction to a simpler problem

Need to solve IP;

$$\min \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]}$$

$$\text{s.t.} \quad \ell \in \mathbb{Z}_+^{|\mathcal{X}|},$$

$$\sum_{x \in \mathcal{X}} 2^{-\ell(x)} \leq 1$$

Reduction to a simpler problem

Need to solve IP;

$$\min \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]}$$

s.t. $\ell \in \mathbb{Z}_+^{|\mathcal{X}|}$,

$$\sum_{x\in\mathcal{X}} 2^{-\ell(x)} \leq 1$$

Instead solve RP;

$$\min \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]}$$

s.t. $\ell \in \mathbb{R}_+^{|\mathcal{X}|}$,

$$\sum_{x\in\mathcal{X}} 2^{-\ell(x)} \leq 1$$

Reduction to a simpler problem

Need to solve IP;

$$\min \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]}$$
$$\text{s.t.} \quad \ell \in \mathbb{Z}_+^{|\mathcal{X}|}, \quad \sum_{x \in \mathcal{X}} 2^{-\ell(x)} \leq 1$$

Instead solve RP;

$$\min \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]}$$
$$\text{s.t.} \quad \ell \in \mathbb{R}_+^{|\mathcal{X}|}, \quad \sum_{x \in \mathcal{X}} 2^{-\ell(x)} \leq 1$$

and use $\ell(x) = \lceil \ell^*(x) \rceil \quad \forall x \in \mathcal{X}$

Reduction to a simpler problem

Need to solve IP;

$$\min \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]}$$

s.t. $\ell \in \mathbb{Z}_+^{|\mathcal{X}|}$,

$$\sum_{x \in \mathcal{X}} 2^{-\ell(x)} \leq 1$$

Instead solve RP;

$$\min \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]}$$

s.t. $\ell \in \mathbb{R}_+^{|\mathcal{X}|}$,

$$\sum_{x \in \mathcal{X}} 2^{-\ell(x)} \leq 1$$

and use $\ell(x) = \lceil \ell^*(x) \rceil \quad \forall x \in \mathcal{X}$

Proposition

Cost using this approach will be atmost 2.5 bits away from the optimal cost.

Structural Result for RP

Main Theorem

Optimal solution for RP is unique, and is given by

$$\ell^*(x) = -\log P_P^*(X) \quad \forall x \in \mathcal{X},$$

where P_P^* is a tilting of source distribution P.

Structural Result for RP

Main Theorem

Optimal solution for RP is unique, and is given by

$$\ell^*(x) = -\log P_P^*(X) \quad \forall x \in \mathcal{X},$$

where P_P^* is a tilting of source distribution P.
Furthermore,

$$P_P^*(x) = \frac{g(y^*, x)}{\sum_{x \in \mathcal{X}} g(y^*, x)},$$

where

$$y^* = \max_{\substack{y \in \mathcal{Y}, \\ g(y,\cdot) \geq 0}} \sum_{x \in \mathcal{X}} g(y, x) \log \frac{\sum_{x \in \mathcal{X}} g(y, x)}{g(y, x)}$$

Structural Result for RP

Main Theorem

Optimal solution for RP is unique, and is given by

$$\ell^*(x) = -\log P_P^*(X) \quad \forall x \in \mathcal{X},$$

where P_P^* is a tilting of source distribution P.
Furthermore,

$$P_P^*(x) = \frac{g(y^*, x)}{\sum_{x\in\mathcal{X}} g(y^*, x)},$$

where

$$y^* = \max_{\substack{y\in\mathcal{Y},\\ g(y,\cdot)\geq 0}} \sum_{x\in\mathcal{X}} g(y, x) \log \frac{\sum_{x\in\mathcal{X}} g(y, x)}{g(y, x)}$$

How does the tilted distribution look like?

Illustration of Optimal Tilting

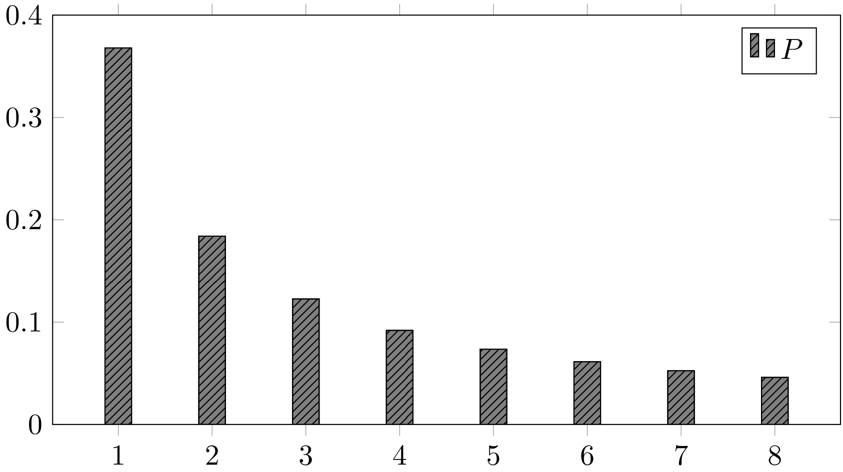
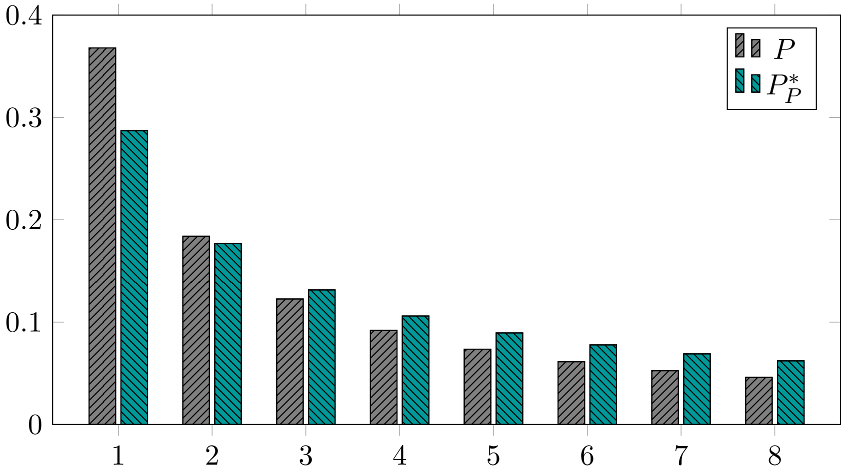


Illustration of Optimal Tilting



Proof sketch of Main theorem

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) \triangleq \min_{\ell \in \Lambda} \mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]}$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \min_{\ell \in \Lambda} \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Optimal lengths for RP

$$y^* := \arg\max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Optimal lengths for RP

$$\ell^* = \arg\min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y^*, x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Optimal lengths for RP

$$\ell^*(x) = \log \frac{\sum_{x \in \mathcal{X}} g(y^*, x)}{g(y^*, x)}$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} g(y,x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y\in\mathcal{Y}} \min_{\ell\in\Lambda} \sum_{x\in\mathcal{X}} g(y,x)\ell(x)$$

- Optimal lengths for RP

$$\ell^*(x) = \log \frac{\sum_{x\in\mathcal{X}} g(y^*,x)}{g(y^*,x)}$$

- Reduction of max-min to max problem

$$\Delta^*(P) = \max_{y\in\mathcal{Y}} \min_{\ell\in\Lambda} \sum_{x\in\mathcal{X}} g(y,x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Optimal lengths for RP

$$\ell^*(x) = \log \frac{\sum_{x \in \mathcal{X}} g(y^*, x)}{g(y^*, x)}$$

- Reduction of max-min to max problem

$$\Delta^*(P) = \max_{\substack{y \in \mathcal{Y}, \\ g(y,\cdot) \geq 0}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Optimal lengths for RP

$$\ell^*(x) = \log \frac{\sum_{x \in \mathcal{X}} g(y^*,x)}{g(y^*,x)}$$

- Reduction of max-min to max problem

$$\Delta^*(P) = \max_{\substack{y \in \mathcal{Y}, \\ g(y,\cdot) \geq 0}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

Proof sketch of Main theorem

- Linear form for the average age cost in terms of code-lengths

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Saddle point claim

$$\Delta^*(P) = \max_{y \in \mathcal{Y}} \min_{\ell \in \Lambda} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Optimal lengths for RP

$$\ell^*(x) = \log \frac{\sum_{x \in \mathcal{X}} g(y^*,x)}{g(y^*,x)}$$

- Reduction of max-min to max problem

$$\Delta^*(P) = \max_{\substack{y \in \mathcal{Y}, \\ g(y,\cdot) \geq 0}} \sum_{x \in \mathcal{X}} g(y,x) \log \frac{\sum_{x \in \mathcal{X}} g(y,x)}{g(y,x)}$$

Linearizing the Average Age Cost

$$\mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

Linearizing the Average Age Cost

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Note that,

$$\mathbb{E}[L] + \frac{\mathbb{E}[L^2]}{2\mathbb{E}[L]} = \max_{z \geq 0} \left(1 - \frac{z^2}{2}\right) \mathbb{E}[L] + z\sqrt{\mathbb{E}[L^2]}.$$

How to linearize $\sqrt{\mathbb{E}[L^2]}$?

Variational Formula for p-norm of a Random Variable

Theorem

For a random variable X with distribution P and $p \geq 1$ such that $\|X\|_p < \infty$, we have

$$\|X\|_p = \max_{Q \ll P} \mathbb{E}\left[\left(\frac{dQ}{dP}\right)^{\frac{1}{p'}} |X|\right],$$

where $p' = p/(p-1)$ is the Hölder conjugate of p.

Variational Formula for p-norm of a Random Variable

Theorem

For a random variable X with distribution P and $p \geq 1$ such that $\|X\|_p < \infty$, we have

$$\|X\|_p = \max_{Q \ll P} \mathbb{E}\left[\left(\frac{dQ}{dP}\right)^{\frac{1}{p'}} |X|\right],$$

where $p' = p/(p-1)$ is the Hölder conjugate of p.

Proof Idea:

For $Q \ll P$, $\alpha = p'$,

$$D_\alpha(P_P, Q) = \frac{1}{\alpha - 1} \log \mathbb{E}_P\left[\left(\frac{dQ}{dP}\right)^{\alpha} \left(\frac{dP_p}{dP}\right)^{1-\alpha}\right] \geq 0,$$

where $\frac{dP_p}{dP} = \frac{1}{\|X\|_p^p} \cdot |X|^p$.

Variational Formula for p-norm of a Random Variable

Theorem

For a random variable X with distribution P and $p \geq 1$ such that $\|X\|_p < \infty$, we have

$$\|X\|_p = \max_{Q \ll P} \mathbb{E}\left[\left(\frac{dQ}{dP}\right)^{\frac{1}{p'}} |X|\right],$$

where $p' = p/(p-1)$ is the Hölder conjugate of p.

Application:

$$\sqrt{\mathbb{E}\left[L^2\right]} = \max_{Q \ll P} \sum_{x \in \mathcal{X}} \sqrt{Q(x)P(x)}\ell(x)$$

Linearizing the Average Age Cost

$$\mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y, x)\ell(x)$$

- Note that,

$$\mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} = \max_{z \geq 0} \left(1 - \frac{z^2}{2}\right)\mathbb{E}\left[L\right] + z\sqrt{\mathbb{E}\left[L^2\right]}.$$

Linearizing the Average Age Cost

$$\mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} = \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} g(y,x)\ell(x)$$

- Note that,

$$\mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} = \max_{z \geq 0} \left(1 - \frac{z^2}{2}\right) \mathbb{E}\left[L\right] + z\sqrt{\mathbb{E}\left[L^2\right]}.$$

- Using variational formula for $\sqrt{\mathbb{E}\left[L^2\right]}$, average age equals

$$= \max_{z \geq 0} \left(1 - \frac{z^2}{2}\right) \mathbb{E}\left[L\right] + z \max_{Q \ll P} \sum_{x \in \mathcal{X}} \sqrt{Q(x)P(x)}\ell(x),$$

$$= \max_{z \geq 0} \max_{Q \ll P} \sum_{x \in \mathcal{X}} g(z,Q,x)\ell(x).$$

Almost optimal recipe for minimizing average age

- Solve the maximization problem

$$\Delta^*(P) = \max_{\substack{z \geq 0, Q \ll P, \\ g(y,\cdot) \geq 0}} \sum_{x \in \mathcal{X}} g(z, Q, x) \log \frac{\sum_{x \in \mathcal{X}} g(z, Q, x)}{g(z, Q, x)}$$

Almost optimal recipe for minimizing average age

- Solve the maximization problem

$$\Delta^*(P) = \max_{\substack{z \geq 0, Q \ll P, \\ g(y,\cdot) \geq 0}} \sum_{x \in \mathcal{X}} g(z, Q, x) \log \frac{\sum_{x \in \mathcal{X}} g(z, Q, x)}{g(z, Q, x)}$$

- Use

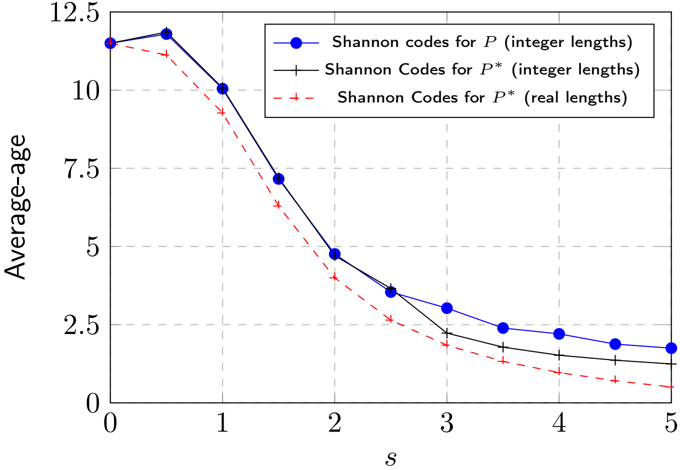
$$\ell(x) = \left\lceil \log \frac{\sum_{x \in \mathcal{X}} g(z^*, Q^*, x)}{g(z^*, Q^*, x)} \right\rceil \quad \forall x$$

Simulation Results

`Zipf`(s, N) is given by $P(i) = \frac{i^{-s}}{\sum_{j=1}^{N} j^{-s}}, \quad 1 \leq i \leq N$.

Simulation Results

$\texttt{Zipf}(s, N)$ is given by $P(i) = \frac{i^{-s}}{\sum_{j=1}^{N} j^{-s}}, \quad 1 \leq i \leq N$.



Comparison of proposed codes and Shannon codes for $\texttt{Zipf}(s, 256)$ w.r.t. s.

Conclusion

[2]Humblet, P. A. (1978). Source coding for communication concentrators.

[3]Larmore, L. L. (1989). Minimum delay codes. SIAM Journal on Computing, 18(1), 82-94.

Conclusion

- Structural property for the optimal solution for RP.

[2] Humblet, P. A. (1978). Source coding for communication concentrators.
[3] Larmore, L. L. (1989). Minimum delay codes. SIAM Journal on Computing, 18(1), 82-94.

Conclusion

- Structural property for the optimal solution for RP.
- A recipe for minimizing the average age cost.

[2] Humblet, P. A. (1978). Source coding for communication concentrators.

[3] Larmore, L. L. (1989). Minimum delay codes. SIAM Journal on Computing, 18(1), 82-94.

Conclusion

- Structural property for the optimal solution for RP.
- A recipe for minimizing the average age cost.
- Extensions:
  - Randomized Memoryless Update Schemes.

[2] Humblet, P. A. (1978). Source coding for communication concentrators.
[3] Larmore, L. L. (1989). Minimum delay codes. SIAM Journal on Computing, 18(1), 82-94.

Conclusion

- Structural property for the optimal solution for RP.
- A recipe for minimizing the average age cost.
- Extensions:
  - Randomized Memoryless Update Schemes.
  - Our Recipe for the Minimum Delay Problem[2] [3]

[2] Humblet, P. A. (1978). Source coding for communication concentrators.

[3] Larmore, L. L. (1989). Minimum delay codes. SIAM Journal on Computing, 18(1), 82-94.

Conclusion

- Structural property for the optimal solution for RP.
- A recipe for minimizing the average age cost.
- Extensions:
  - Randomized Memoryless Update Schemes.
  - Our Recipe for the Minimum Delay Problem[2] [3]

Average Delay

$$\bar{D}(e) = \begin{cases} \mathbb{E}\left[L\right] + \frac{\lambda \mathbb{E}\left[L^2\right]}{2(1-\lambda \mathbb{E}[L])}, & \lambda \mathbb{E}\left[L\right] < 1, \\ \infty, & \lambda \mathbb{E}\left[L\right] \geq 1. \end{cases}$$

Average Age

$$\bar{A}(e) = \mathbb{E}\left[L\right] + \frac{\mathbb{E}\left[L^2\right]}{2\mathbb{E}\left[L\right]} - \frac{1}{2}$$

[2] Humblet, P. A. (1978). Source coding for communication concentrators.

[3] Larmore, L. L. (1989). Minimum delay codes. SIAM Journal on Computing, 18(1), 82-94.