Finite-state Rate Distortion

Dharmendra S. Modha[†] and Daniela P. de Farias[‡]

[†]IBM Almaden Research Center

[‡]MIT

ISIT 2004, Chicago, IL, July 2, 2004

Finite-state Compressibility

Given an individual sequence $x \in B^\infty$, Lempel and Ziv (1978) defined

$$\rho(x) \equiv \rho(x, B) = \lim_{s\to\infty} \limsup_{n\to\infty} \min_{\mu\in \mathbb{M}^i(s, B)} \frac{L(x_1^n, \mu)}{n},$$

Finite-state Compressibility

Given an individual sequence $x \in B^{\infty}$, Lempel and Ziv (1978) defined

$$\rho(x) \equiv \rho(x, B) = \lim_{s \to \infty} \limsup_{n \to \infty} \min_{\mu \in \mathbb{M}^i(s, B)} \frac{L(x_1^n, \mu)}{n},$$

where

- $\mathbb{M}^i(s, B)$ is set of *information lossless* machines with at most s states that map sequences in B^{∞} to $\{0, 1\}^{\infty}$

Finite-state Compressibility

Given an individual sequence $x \in B^\infty$, Lempel and Ziv (1978) defined

$$\rho(x) \equiv \rho(x, B) = \lim_{s\to\infty} \limsup_{n\to\infty} \min_{\mu \in \mathbb{M}^i(s,B)} \frac{L(x_1^n, \mu)}{n},$$

where

- $\mathbb{M}^i(s, B)$ is set of *information lossless* machines with at most s states that map sequences in B^∞ to $\{0, 1\}^\infty$
- $L(x_1^n, \mu)$ is the number of bits output by the machine μ upon seeing the first n bits of the input sequence

Notation

- source alphabet: B

Notation

- source alphabet: B
- reproduction alphabet: $\hat{B}$

Notation

- source alphabet: B
- reproduction alphabet: $\hat{B}$
- distortion measure $d : B \times \hat{B} \to [0, \infty)$

$$\max_{b \in B} \min_{\hat{b} \in \hat{B}} d(b, \hat{b}) = 0;$$

Notation

- source alphabet: B
- reproduction alphabet: $\hat{B}$
- distortion measure $d : B \times \hat{B} \to [0, \infty)$

$$\max_{b \in B} \min_{\hat{b} \in \hat{B}} d(b, \hat{b}) = 0;$$

- Desired average per-letter distortion D;

Prior Work

- Ziv (1980) and Yang and Kieffer (1996)

Prior Work

- Ziv (1980) and Yang and Kieffer (1996)
- Essence of the definition:

$$R(D|x) = \inf\{\rho(y, \hat{B}) : y \in \hat{B}^\infty, d_\infty(x, y) \le D\}$$

where $d_\infty(x, y) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i)$

Prior Work

- Ziv (1980) and Yang and Kieffer (1996)

- Essence of the definition:

$$R(D|x) = \inf\{\rho(y, \hat{B}) : y \in \hat{B}^\infty, d_\infty(x, y) \le D\}$$

where $d_\infty(x, y) = \limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i)$

- $R(D|x)$ is an asymptotically attainable lower bound

Prior Work

- Ziv (1980) and Yang and Kieffer (1996)
- Essence of the definition:

$$R(D|x) = \inf\{\rho(y, \hat{B}) : y \in \hat{B}^\infty, d_\infty(x, y) \le D\}$$

  where $d_\infty(x, y) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i)$

- $R(D|x)$ is an asymptotically attainable lower bound
- If x is drawn from a stationary, ergodic process, then $R(D|x)$ becomes the rate-distortion

A class of D-Lossy Machines

A deterministic machine $\hat{E} : B^\infty \to \hat{B}^\infty$ is D-lossy if

$$\sup_{x \in B^\infty} d_\infty(x, \hat{E}x) \le D.$$

A class of D-Lossy Machines

A deterministic machine $\hat{E} : B^\infty \to \hat{B}^\infty$ is <u>D-lossy</u> if

$$\sup_{x \in B^\infty} d_\infty(x, \hat{E}x) \le D.$$

Let $\hat{\mathbb{M}}(s, D, B, \hat{B})$, $s \in \mathbb{N}$, denote the class of D-lossy machines with at most s states.

A class of D-Lossy Machines

A deterministic machine $\hat{E} : B^\infty \to \hat{B}^\infty$ is D-lossy if

$$\sup_{x \in B^\infty} d_\infty(x, \hat{E}x) \leq D.$$

Let $\hat{\mathbb{M}}(s, D, B, \hat{B})$, $s \in \mathbb{N}$, denote the class of D-lossy machines with at most s states.

Lemma 1: Let $\hat{E}$ be in $\cup_{D \geq 0} \cup_{s=0}^{\infty} \hat{\mathbb{M}}(s, D, B, \hat{B})$, then

$$\rho(x) = \rho(\hat{E}x) + \rho(x|\hat{E}x),$$

where $\rho(x|\hat{E}x)$ is as in Merhav (2000).

Finite-state Rate Distortion

For a fixed source sequence $x \in B^\infty$, define

$$\pi(x, D) = \lim_{s \to \infty} \min_{\hat{E} \in \hat{\mathbb{M}}(s, D, B, \hat{B})} \rho(\hat{E}x, \hat{B}).$$

Finite-state Rate Distortion

For a fixed source sequence $x \in B^\infty$, define

$$\pi(x, D) = \lim_{s \to \infty} \min_{\hat{E} \in \hat{\mathbb{M}}(s,D,B,\hat{B})} \rho(\hat{E}x, \hat{B}).$$

Lemma 2: $\pi(x, D) = R(D|x)$.

Main Theorem

If $B = \hat{B} = \{0, 1\}$ and d is Hamming distortion, then

$$\pi(x, D) \geq \rho(x, B) - h(D).$$

Main Theorem

If $B = \hat{B} = \{0, 1\}$ and d is Hamming distortion, then

$$\pi(x, D) \geq \rho(x, B) - h(D).$$

Proof:

$$\begin{aligned}
\pi(x, D) &= \rho(x) - \lim_{s\to\infty} \max_{\hat{E}\in\hat{\mathbb{M}}(s,D,B,\hat{B})} \rho(x|\hat{E}x) \\
&\geq \rho(x) - \lim_{s\to\infty} \max_{\hat{E}\in\hat{\mathbb{M}}(s,D,B,\hat{B})} \rho(x \oplus \hat{E}x|\hat{E}x) \\
&\geq \rho(x) - \lim_{s\to\infty} \max_{\hat{E}\in\hat{\mathbb{M}}(s,D,B,\hat{B})} \rho(x \oplus \hat{E}x) \\
&\geq \rho(x) - h(D)
\end{aligned}$$