Lecture 1
Wednesday, September 10, 2014
2:44 PM

Information Signal → Channel

\[ P_s \rightarrow S \]

Channel Coding: Convert channel to noise-free digital channel (create digital resources from physical ones)

Source Coding: Encode information source to digital signal (consume digital resources)

Example: Suppose we know that information source produces data

\[ X = \{a, b, c, d\} \]

\[ P(a) = \frac{1}{2} \]
\[ P(b) = \frac{1}{4} \]
\[ P(c) = \frac{1}{8} \]
\[ P(d) = \frac{1}{8} \]

Encoding:

\[ a \rightarrow 00 \]
\[ b \rightarrow 01 \]
\[ c \rightarrow 10 \]
\[ d \rightarrow 11 \]

Requires 2 bits per symbol: \( R = 2 \) bits

Better encoding:

\[ a \rightarrow 0 \]
\[ b \rightarrow 10 \]
\[ c \rightarrow 110 \]
\[ d \rightarrow 111 \]

"Variable length"

\[ L = E(X) = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 = \frac{7}{4} \]

Decodable, even when a list is concatenated. (Prefix-free)
Necessary and sufficient condition: (Kraft inequality)

\[ \sum_{x \in X} 2^{-L(x)} \leq 1 \]

Optimal Code:

Given \( p(x) \)

minimize \( E \ell(x) \)

subject to \( \sum_{x \in X} 2^{-E \ell(x)} \leq 1 \)

\( \ell(x) \in \{1, 2, \ldots\} = N \)

Lower bound:

Lagrange multiplier:

\[ \frac{d}{d \ell(x)} \left( \sum_{x} p(x) \ell(x) + \lambda \left( 1 - \sum_{x} 2^{-E \ell(x)} \right) \right) = p(x) - \lambda 2^{-L(x) \ln 2} = 0 \]

\( \ell^*(x) = \log_2 \frac{1}{p(x)} - \frac{\lambda}{\ln 2} \)

Constant \( L^* \)

\[ \mathbb{E} L^*(x) = \mathbb{E} \log_2 \frac{1}{p(x)} \]

Entropy:

\( H(X) \)

Measure of randomness

\[ 0 \leq H(X) \leq \log |X| \]

deterministic \quad uniform

Mutual information:

\( H(X) + H(Y) \geq H(X, Y) \)

\[ I(x; y) = H(x) + H(y) - H(x, y) \]

Claim: For variable-length source coding:

\[ H(X) \leq L^* \leq H(X) + 1 \text{ bit} \]

Standard Information Theory Approach (Memoryless, IID)
Let \( X_1, X_2 \ldots \) be an i.i.d. process \( \sim P_X \)

\[
X^n = (X_1 \ldots X_n) \xrightarrow{\text{Encoder}} F \xrightarrow{M \in [2^n]^{R}} \xrightarrow{\text{Decoder}} \hat{X}^n
\]

Source encoding at rate \( R \) ("block coding")
Lossless encoding requirement: \( X^n = \hat{X}^n \) w.h.p.
That is, rate \( R \) is achievable if \( \forall \varepsilon > 0 \exists n, F, G \) such that \( P(X^n \neq \hat{X}^n) \leq \varepsilon \)

Claim: \( \inf \{ R : \text{achievable} \} = H(X) \)

---

Protect against transmission noise:
"Error correction"
"Channel Coding"

\[
M \in [2^n]^{R} \xrightarrow{\text{Encoder}} F \xrightarrow{X^n} \xrightarrow{\text{Channel}} Y^n \xrightarrow{\text{Decoder}} \hat{M}
\]

Reliable communication requirement: \( M = \hat{M} \) w.h.p.

Example of error detection: Parity check: \( \overline{0110110} \) over 7 bits

Example of error correction:
Repetition: 111
Decide by majority

Hamming (7, 4) - code
4 information bits 3 parity
Correct 1 error

\[
\begin{align*}
\text{Correct 1 error:} & \quad \begin{bmatrix} b_1, b_2, b_3, b_4 \end{bmatrix} \quad \begin{bmatrix} p_1, p_2, p_3 \end{bmatrix} \\
\text{Parity check:} & \quad \begin{bmatrix} \overline{0110110} \end{bmatrix}
\end{align*}
\]
In technology, Reed-Solomon codes are ubiquitous.

\[ \text{Prob. of error} \approx 2^{11e^2} \]

Binary Symmetric Channel (BSC)

Claim: \( \sup \{ R: \text{achievable} \} \leq C = \max_{p(x)} I(X;Y) \)

Again, entropy gives fundamental limit:

Information Theory ignores complexity, latency, etc.

---

Another aspect of info. theory: Rate-distortion Theory.

Puzzle: How many bits needed to comm. an \( n \)-bit uniformly random source signal such that it can be recovered with at most one bit-error? \( (n^-?) \)

Security:

Quantum:

Why?
- Entropy
- We move information physically
- Quantum computing
- Secure protocols: Quantum Key Exchange.

Revisit Channel Coding:

Classical Capacity

1. Revisit Channel Coding:
Entanglement assisted classical capacity:
- Solved.
- Elegant extension to classical information theory

Quantum Capacity:
- Preserve the quantum state.
- Open problem

Output must be measured. This optimization changes the mathematics of the problem.