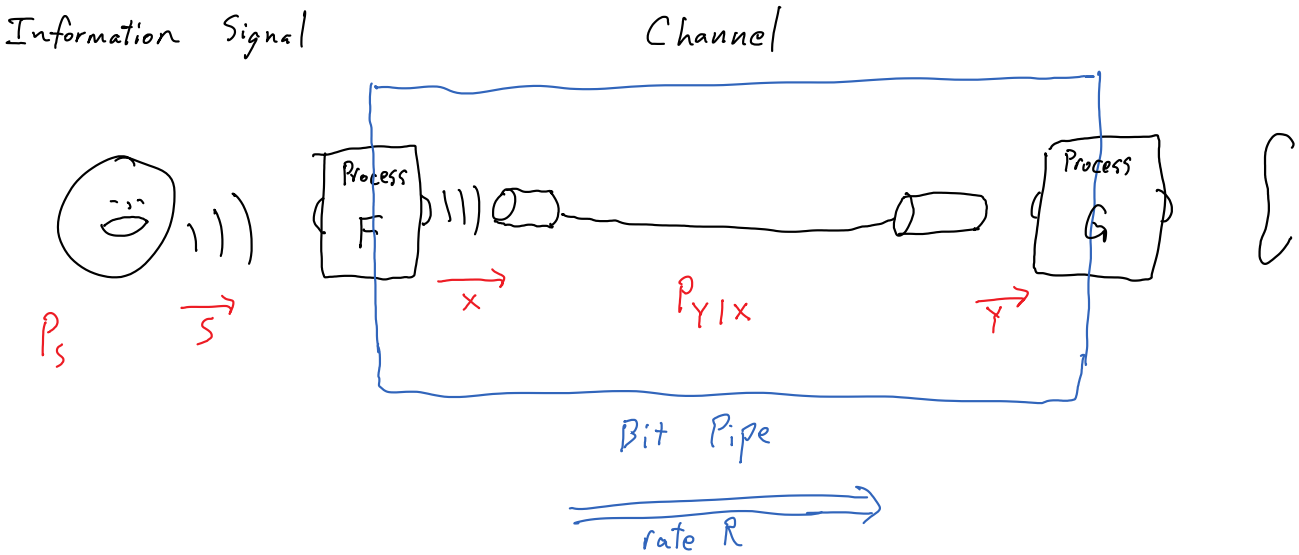
Lecture 1

Wednesday, September 10, 2014
2:44 PM



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

$$\mathcal{X} = \{a, b, c, d\}, \qquad \begin{aligned} p(a) &= \tfrac{1}{2} \\ p(b) &= \tfrac{1}{4} \\ p(c) &= \tfrac{1}{8} \\ p(d) &= \tfrac{1}{8} \end{aligned}$$

Encoding:

$$\begin{aligned} a &\to 00 \\ b &\to 01 \\ c &\to 10 \\ d &\to 11 \end{aligned}$$

requires 2 bits per symbol: $R = 2$ bits

Better encoding:

$$\begin{aligned} a &\to 0 \\ b &\to 10 \\ c &\to 110 \\ d &\to 111 \end{aligned}$$

"variable length"

$$L = E\,\ell(X) = \tfrac{1}{2}\cdot 1 + \tfrac{1}{4}\cdot 2 + \tfrac{1}{4}\cdot 3 = \tfrac{7}{4}$$

Decodable, even when a list is concatenated. (Prefix-free)

Necesary and sufficient condition: (Kraft inequality)

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

Optimal Code:
Given $p(x)$

minimize $E\,\ell(X)$

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

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

Lower bound:

Lagrange multiplier: $\frac{d}{d\,\ell(x)} \sum_x p(x)\,\ell(x) + \lambda\left(1 - \sum_x 2^{-\ell(x)}\right)$

$$= p(x) - \lambda\, 2^{-\ell(x)} \ln 2 = 0$$

$$\ell^*(x) = \log_2 \frac{1}{p(x)} - \underbrace{\log_2 \lambda \ln 2}_{\text{Constant } \forall x}$$

$$\boxed{\ell^*(x) = \log_2 \frac{1}{p(x)}}$$

$$E\,\ell^*(X) = E \log_2 \frac{1}{p(x)}$$

Entropy: $H(X)$

Measure of randomness

$$0 \leq H(X) \leq \log |\mathcal{X}|$$

deterministic uniform

Mutual information: $H(X) + H(Y) \geq H(X,Y)$

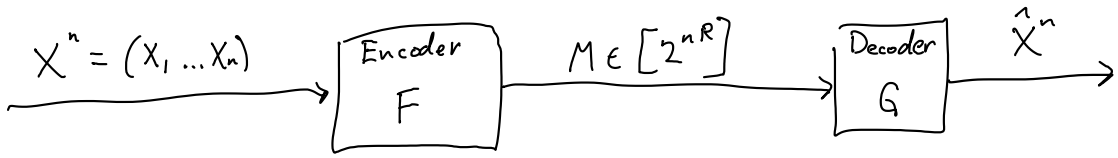
$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

<u>Claim</u>: 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 $\sim P_X$



Source encoding at rate R ("block coding")

Lossless encoding requirement: $X^n = \hat{X}^n$ w.h.p.

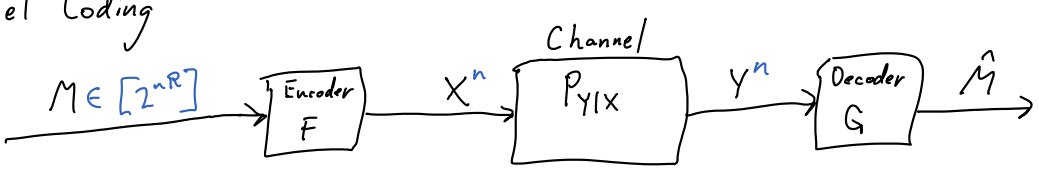
That is, rate R is achievable if $\forall \varepsilon > 0\ \exists n, F, G$ such that $P(X^n \neq \hat{X}^n) < \varepsilon$

<u>Claim:</u> $\inf\{R : \text{achievable}\} = H(X)$

Protect against transmission noise:

"Error correction"

"Channel Coding"



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

Example of error detection: Parity check: $\underbrace{0110110}_{7 \text{ bits}}\,0$

Example of error correction:

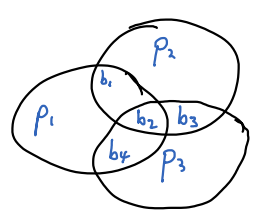
Repetition: 111 Decode by majority

Hamming (7,4) - code

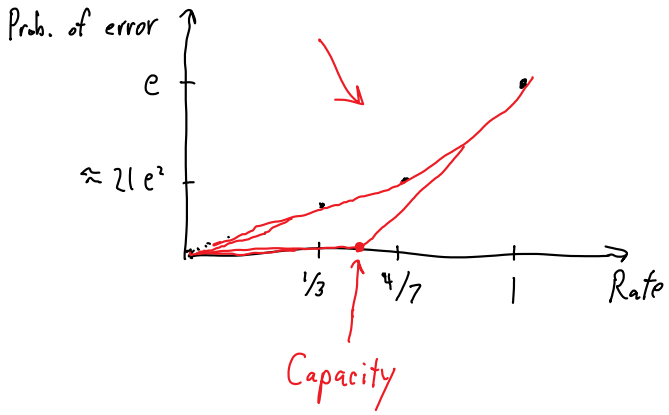
Correct 1 error.

4 information bits 3 parity

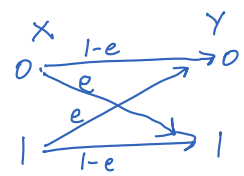
$\boxed{b_1 | b_2 | b_3 | b_4}$ $\boxed{p_1 | p_2 | p_3}$



In technology, Reed-Soloman codes are ubiquitous



Binary Symmetric Channel
BSC



<u>Claim</u>: $\sup\{R : \text{achievable}\} \triangleq C = \max_{p(x)} I(X;Y)$

Again, entropy gives fundamental limit.

Information Theory ignores complexity, latency, etc.

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

Not covered in depth

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-?)

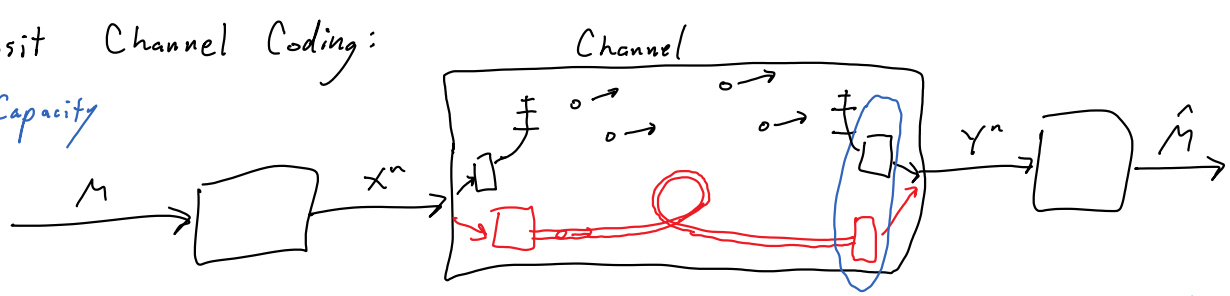
<u>Security</u>:

<u>Quantum</u>: Why?

- Entropy
- We move information physically
- Quantum computing
- Secure protocals: Quantum Key Exchange.

① Revisit Channel Coding:

Classical Capacity



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

2. Entanglement assisted classical capacity.

- Solved.
- Elegant extension to classical information theory

3. Quantum Capacity

- Preserve the quantum state.
- Open problem