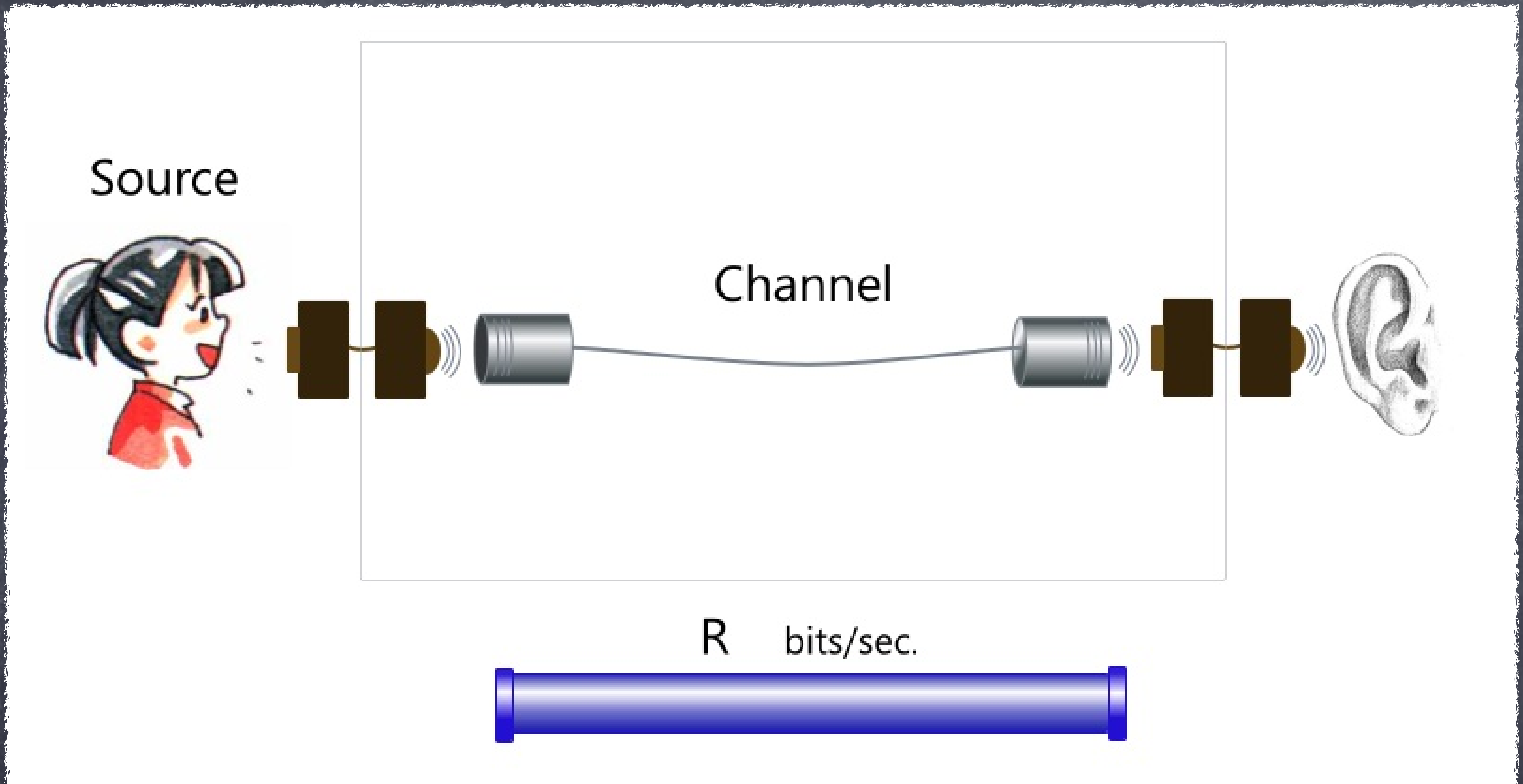
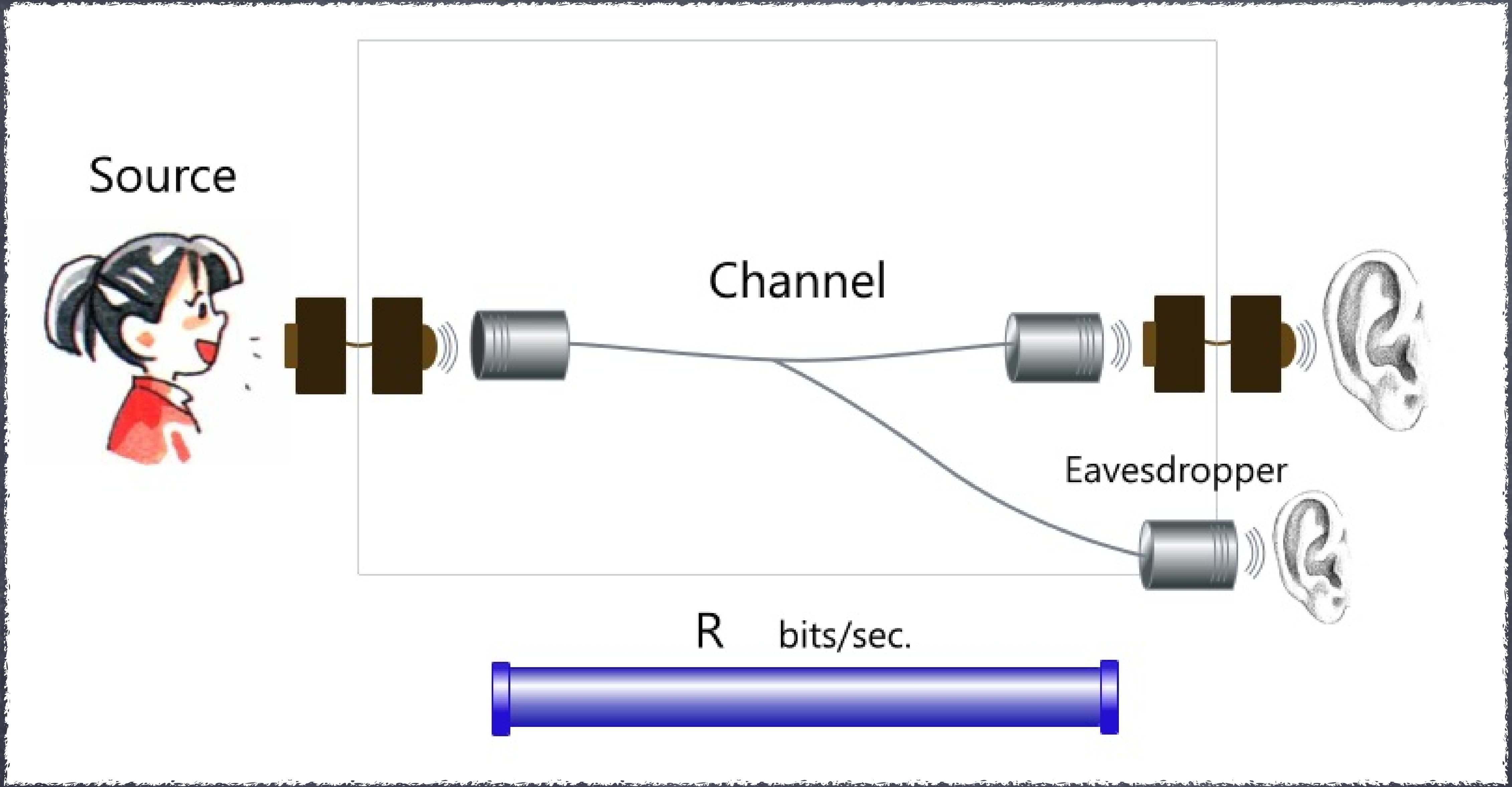
Source Coding (secrecy and embedding)

Paul Cuff - Princeton University



Communication

Move a signal from one place to another

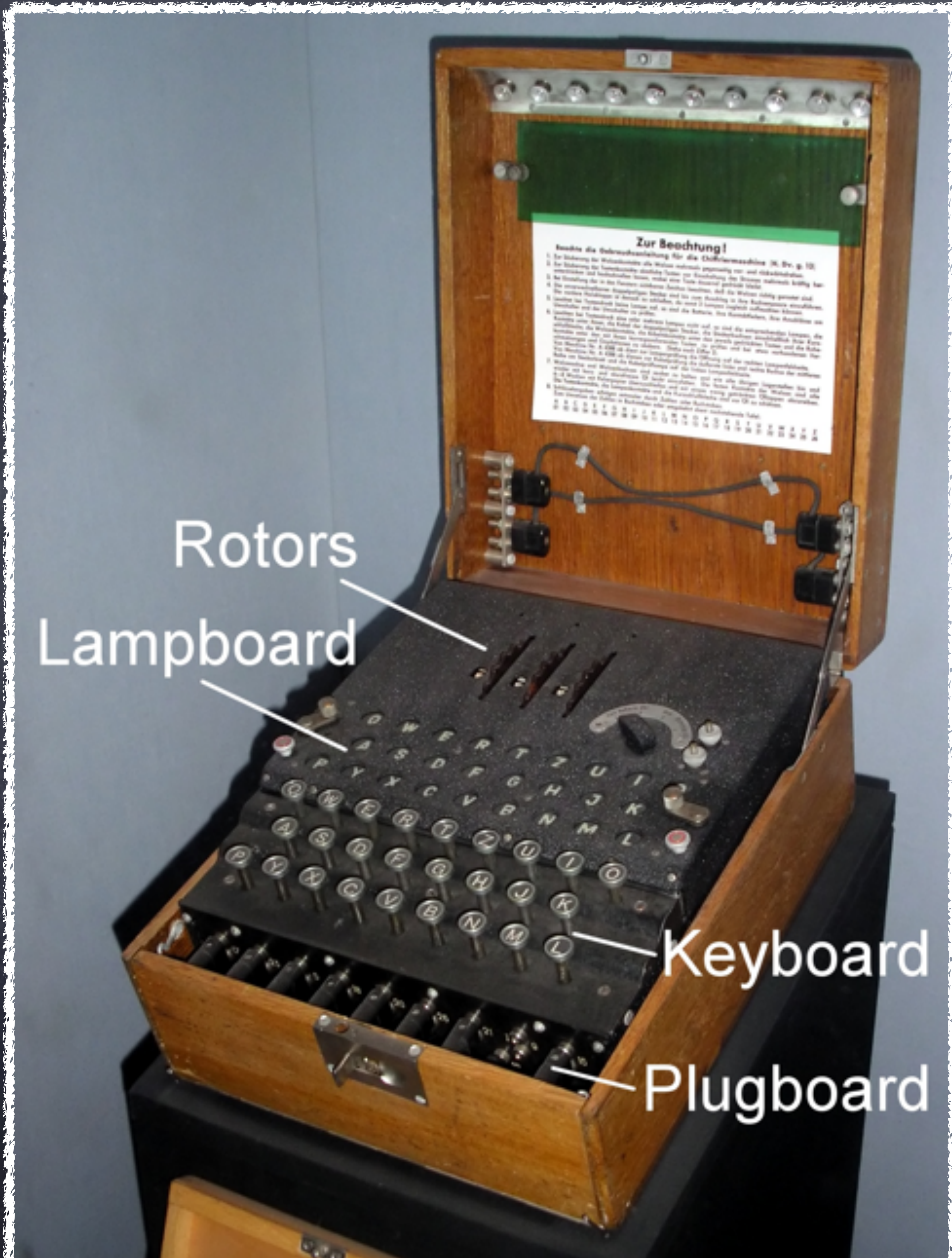


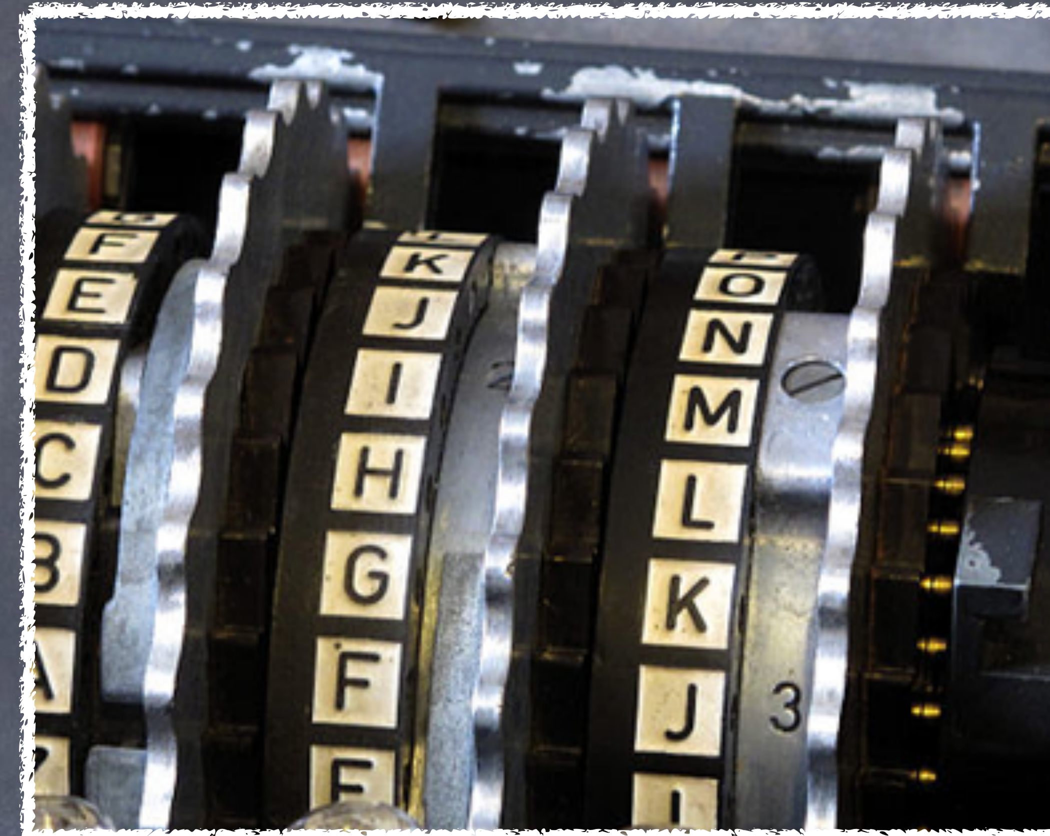
Communication

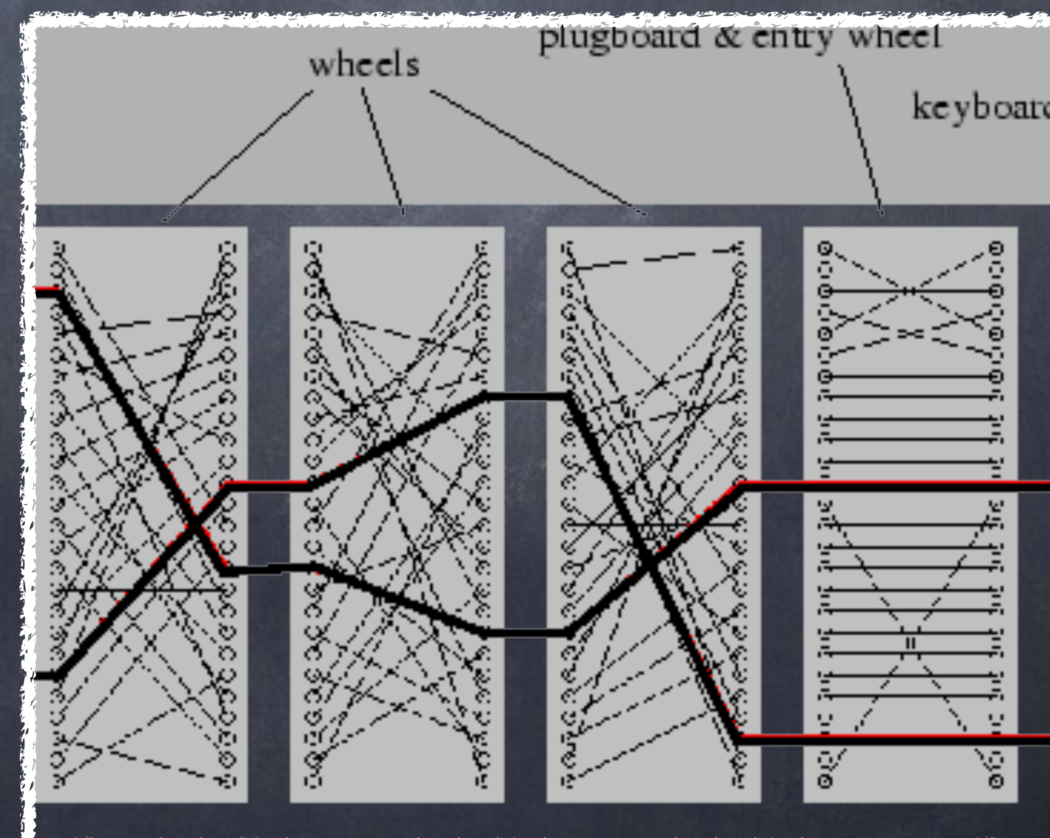
Move a signal from one place to another

Secrecy Systems

- Old Ciphers:
  - Obscure (design kept secret)
  - Complicated
- New Ciphers
  - Computationally challenging
  - Complicated

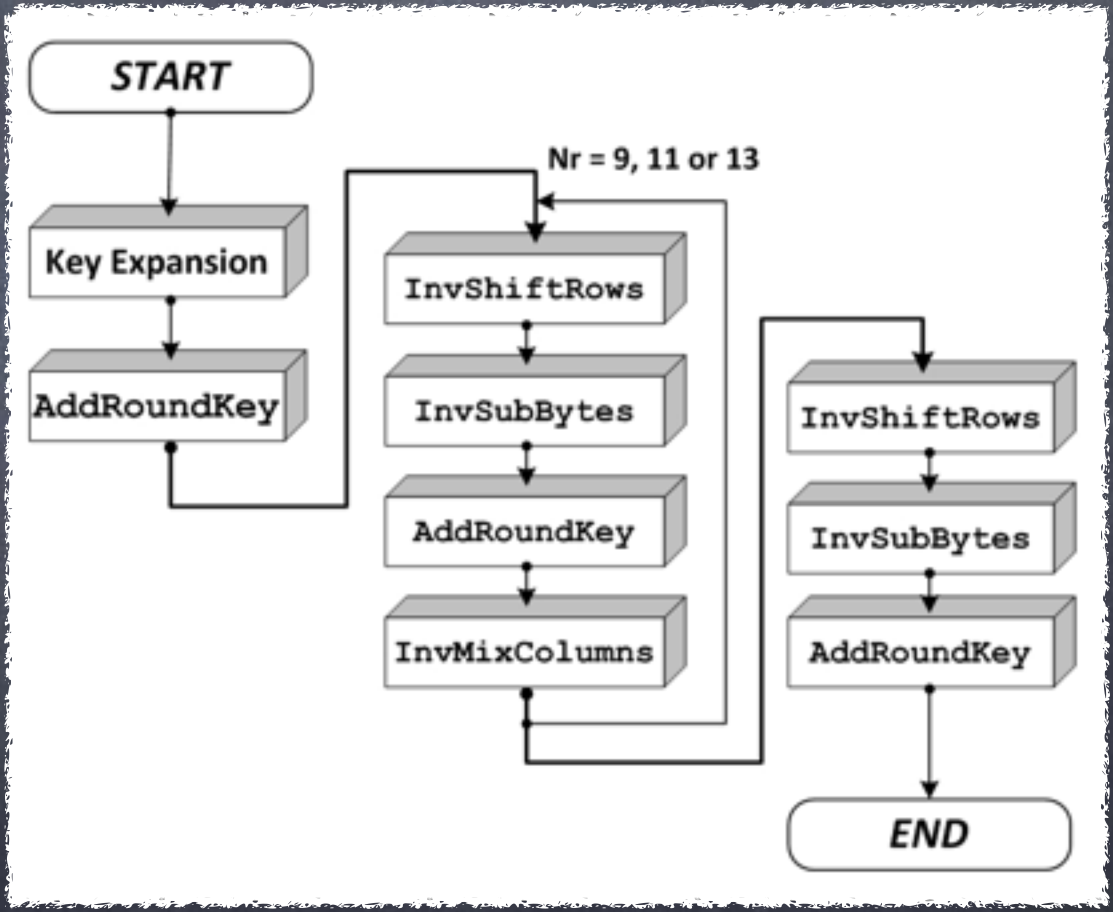

Zur Beachtung!
Rotors
Lampboard
Keyboard
Plugboard




wheels
plugboard & entry wheel

AES

- Complicated
- Not known how to undo



Modern Cryptography

- Began in 1970's (from IT community)
- Built on fundamental mathematical problems that are believed difficult to compute
- Diffie-Hellman
- RSA

Information Theoretic Security

- Shannon's '49 paper
  - Fewer assumptions (omnipotent adversary)
  - More resources required for mathematical guarantees

Perfect Secrecy

- Independence between the transmitted message and the information
- Perfect Secrecy achieved with one-time pad (Vernam cipher).
- Shannon: This method is also necessary. (i.e. $R_k > H(X)$)

One-time Pad

Information: 011010100010110

Key: 110100110100101

$\oplus$

Message: 101110010110011

Overview

Channel Capacity

Source Coding

Secure Channel Coding

Secure Source Coding

Compression

Encoder: $f: \quad \mathcal{X}^n \to [2^{nR}]$

Decoder: $g: \quad [2^{nR}] \to \mathcal{Y}^n$

IID Model

- Let the information signal be i.i.d. with a known distribution P_X
- Represent as: $X_1, X_2, \ldots, X_n = X^n$
- Codec functions over the block

Lossless Compression

Lossless Compression

- R is achievable if for any $\varepsilon>0$, there exists an n, f, and g such that $P(X^n \neq Y^n) < \varepsilon$
- Minimum achievable rate is $H(X)$

Entropy

$$H(X) = \mathbf{E} \log \frac{1}{P_X(X)}$$

$H(X) \geq 0$ (deterministic)

$H(X) \leq \log|X|$ (uniform)

$H(X,Y,Z) = H(X) + H(Y|X) + H(Z|X,Y)$

Binary Entropy

- X is Bern(p)
  - $P(X=1) = p$, $P(X=0) = 1-p$
- $H(X) = p \log(1/p) + (1-p) \log(1/(1-p))$
- Call this the binary entropy function
  - $h(p)$

Mutual Information

$$
\begin{aligned}
I(X;Y) &= H(X) + H(Y) - H(X,Y) \\
&= H(X) - H(X|Y) \\
&= H(Y) - H(Y|X)
\end{aligned}
$$

$I(X;Y) \geq 0$ (independent)

$I(X;Y) \leq H(X)$ (function)

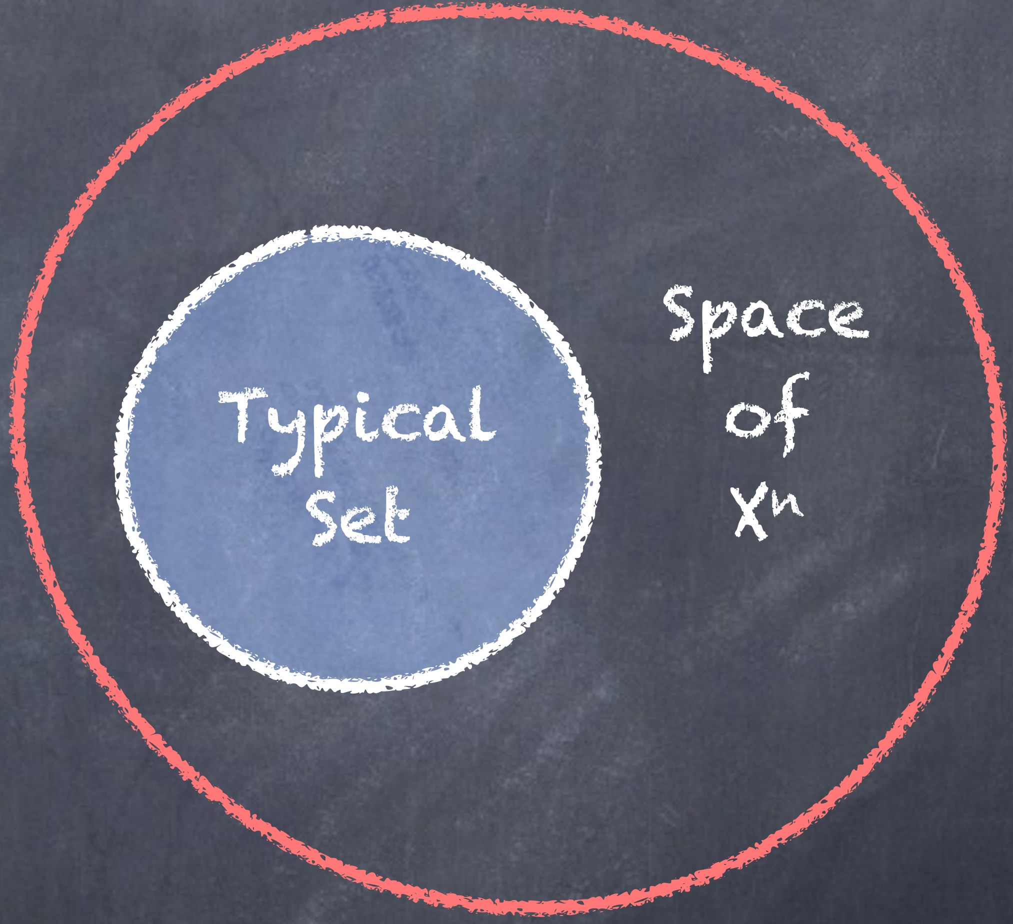
$I(X;Y,Z) = I(X;Y) + I(X;Z|Y)$

Data-processing Inequality

- Markov chain: $X-Y-Z$
- $I(X;Z) \leq I(X;Y)$

Lossless Compression

- Enumerate the typical set:



- Error if X^n not typical (negligible probability)
- $R \approx H(X)$

Typical Set and A.E.P.

- $A = \{x^n : P(x^n) \approx 2^{-nH(X)}\}$
    - exponent within ε of $H(X)$
- $P(A) \to 1$ (law of large numbers)
    - $1/n \, \Sigma_i \log(P(X_i)) \to E \log P(X_i) = -H(X)$
- $|A| \approx 2^{nH(X)}$

Lossy Compression

Rate-Distortion Theory

Average Distortion

- We need a relevant metric for lossy transmission
- Bounded distortion function: $d(x,y)$
- Average distortion:
  - $d(x^n,y^n) = 1/n \sum_i d(x_i,y_i)$

Puzzle

- Given an n-bit random sequence
- 1-bit distortion:
  - How many bits of description are needed to have no more than one bit of distortion? (n - ?)
- 1-bit transmission:
  - How much distortion can be achieved with only a one-bit description? (n/2 - ?)

Definition of achievability

- Rate R is achievable for distortion D if there exist n, f, and g operating at rate R such that

$$E\, d(X^n, Y^n) \leq D$$

Rate-Distortion Theorem [Shannon]

$$R(D) = \min_{P_{Y|X}\,:\,\mathbf{E}d<D} I(X;Y)$$

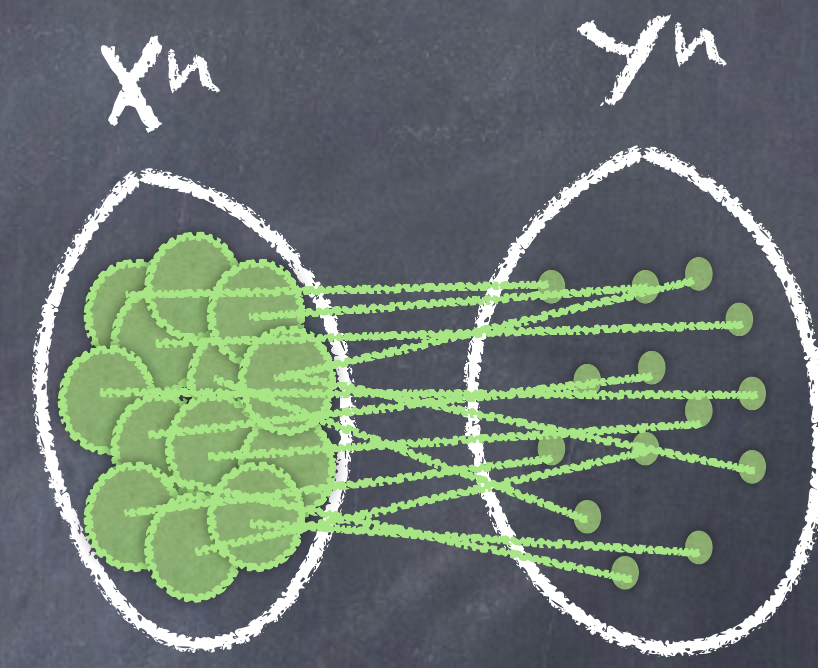
Formula is still an optimization problem

- Choose $P_{Y|X}$ (given P_X and $d(x,y)$):
  - $R > I(X;Y)$
  - $D > E\, d(X,Y)$

Examples

- Binary signal, Hamming distortion:
  - $R(D) = [h(p) - h(D)]_+$
  - Choice of $P_{Y|X}$ is s.t. $P_{X|Y}$ is BSC(D)
- Gaussian signal, squared-error:
  - $R(D) = [1/2 \log(\sigma/D)]_+$
  - Choice of $P_{Y|X}$ is s.t. $P_{X|Y}$ is AGN(D)

Achievability Proof

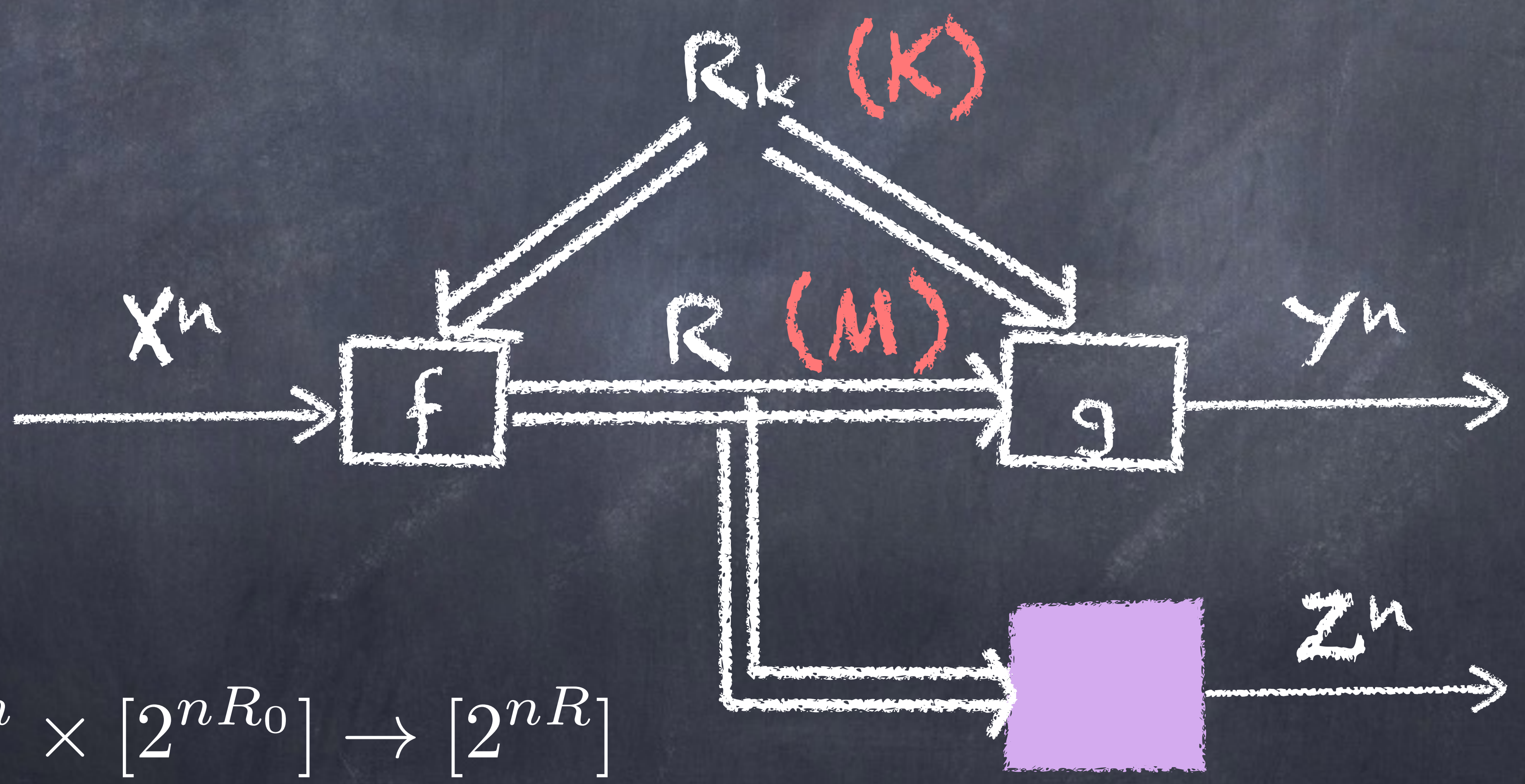


- General idea—Covering

Source Coding for Secrecy

Setting

- Shannon cipher



$$f \; : \; \mathcal{X}^n \times [2^{nR_0}] \to [2^{nR}]$$

$$g \; : \; [2^{nR}] \times [2^{nR_0}] \to \mathcal{Y}^n$$

Performance Metric

- Theory will include:
  - Lossy reconstruction
  - Imperfect secrecy
- Secrecy also measured by distortion
  - Lowest distortion achievable by the eavesdropper

Two distortion functions

- Distortion at the intended receiver:
  - $d_1(x^n,y^n) = 1/n \sum_i d_1(x_i,y_i)$
- Distortion at the eavesdropper
  - $d_2(x^n,z^n) = 1/n \sum_i d_2(x_i,z_i)$

Pessimistic

- Assume eavesdropper makes best use of the information:

$$D_2 = \min_{Z^n = z^n(M)} \mathbf{E}\, d_2(X^n, Z^n)$$

Definition of achievability

- (R, R_k, D_1, D_2) is achievable if there exists an n, f, and g operating at rates R and R_k such that

$$\mathbb{E}\, d_1(X^n, Y^n) \le D_1$$

$$\min_{(\text{adversary})} \mathbb{E}\, d_2(X^n, Z^n) \ge D_2$$

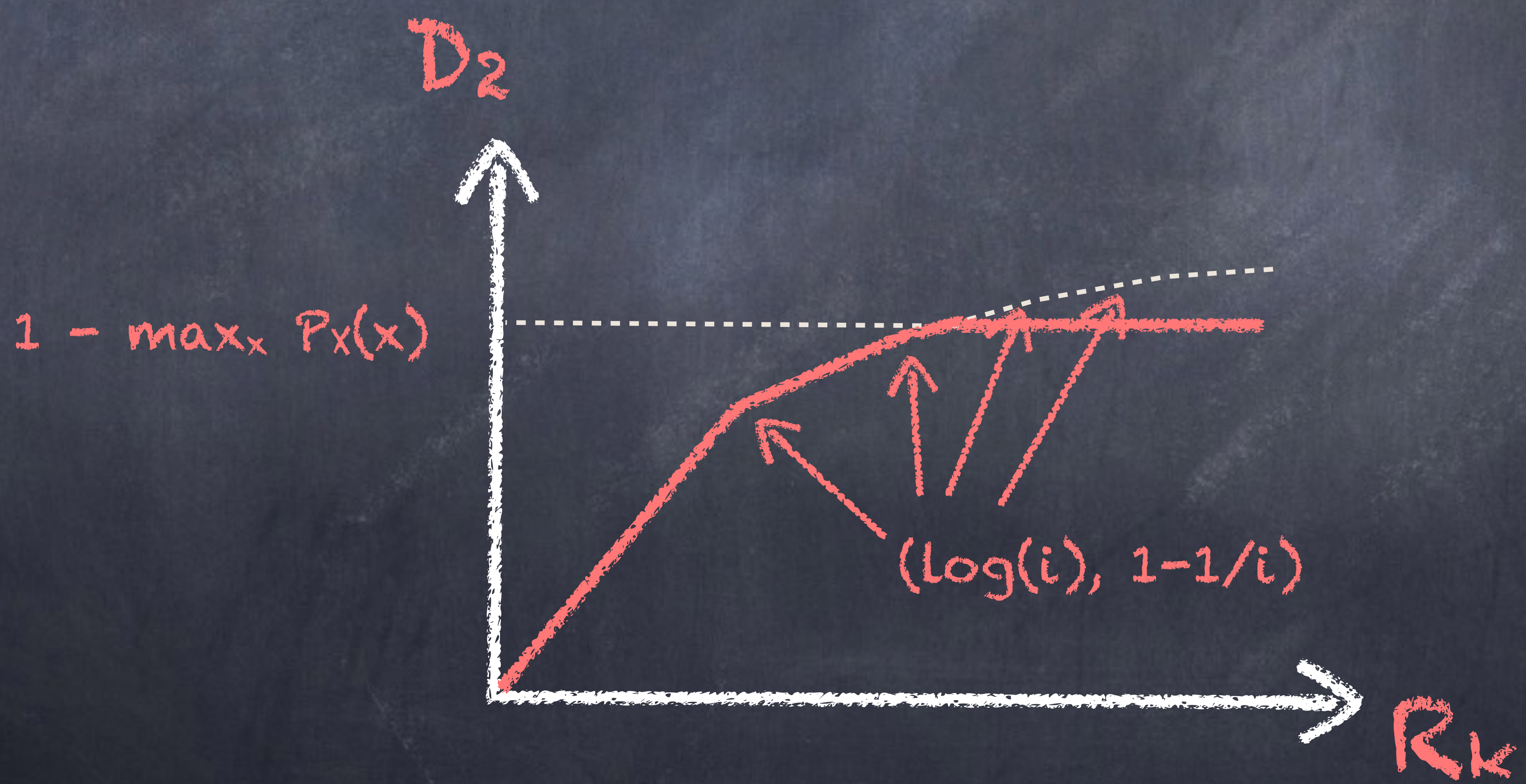
Rate-Distortion Theory for Secrecy Systems

- (R, R_k, D_1, D_2) achievable if and only if
  - There exist U, V, Y s.t.
  - $X-(U,V)-Y$
  - $R \geq I(X;U,V)$
  - $R_k \geq I(X,Y;V|U)$
  - $E\, d_1(X,Y) \leq D_1$
  - $\min_{z(.)} E\, d_2(X,z(U)) \geq D_2$

Recall that we are given P_X

Hamming Distortion

- Any source distribution, with $D_1=0$:



Equivocation Region

Alternative Problem

- No distortion at the eavesdropper
- Measure secrecy by equivocation rate
  - $\Delta_X = 1/n\ H(X^n|M)$
- Or, for lossy compression
  - $\Delta_{X,Y} = 1/n\ H(X^n,Y^n|M)$

Recover equivocation from RD Theory

- Equivocation rate is one special case of rate-distortion theory
- Causal disclosure is necessary

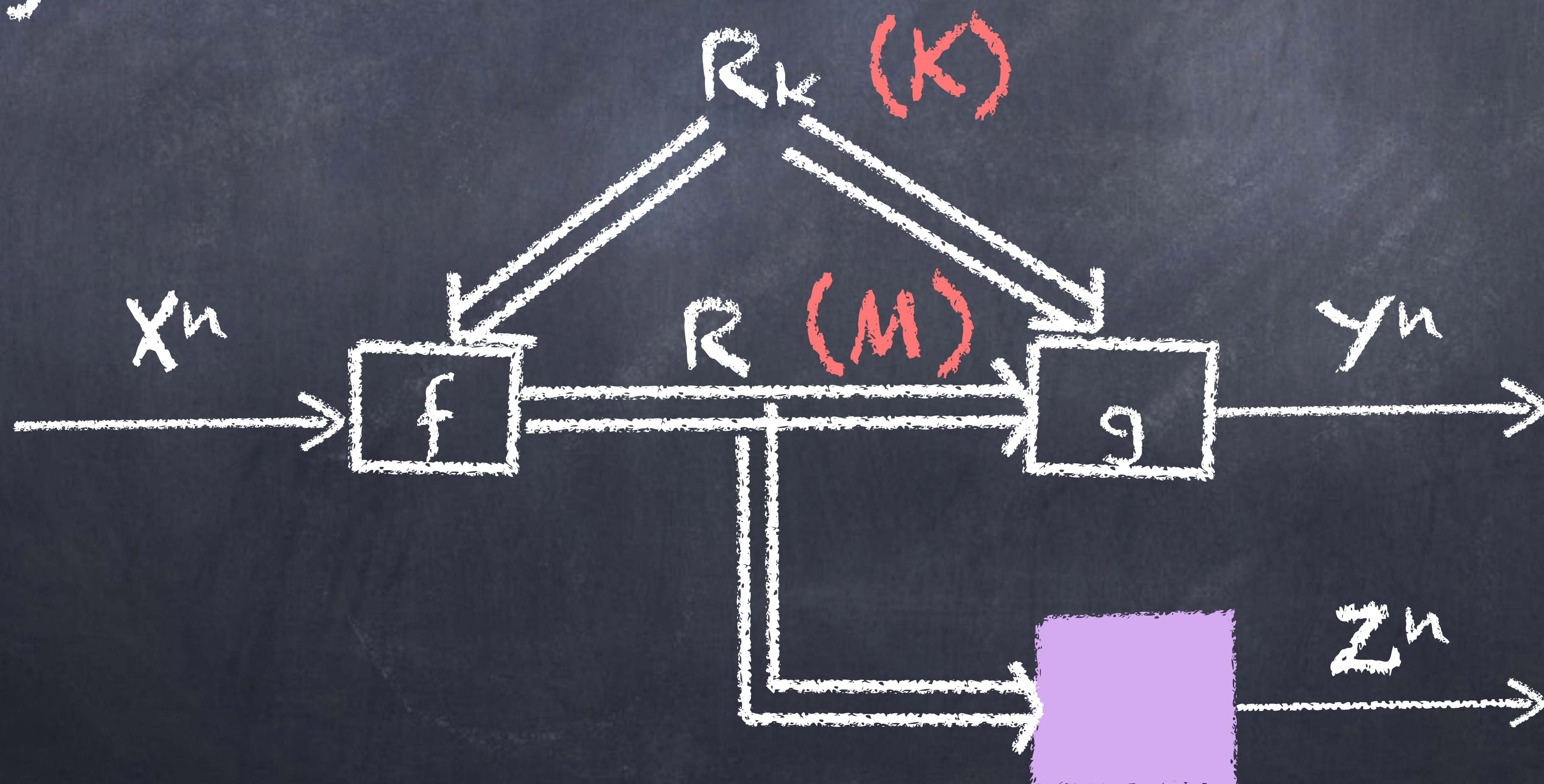
Log-loss function

- Distortion function
  - $d(x,z) = \log 1/z(x)$
  - z is a distribution
- Notice: $\min_{z(m)} E\, d(X,z(M)) = H(X|M)$

  $\min_{z(x,m)} E \log 1/z(x,m)$

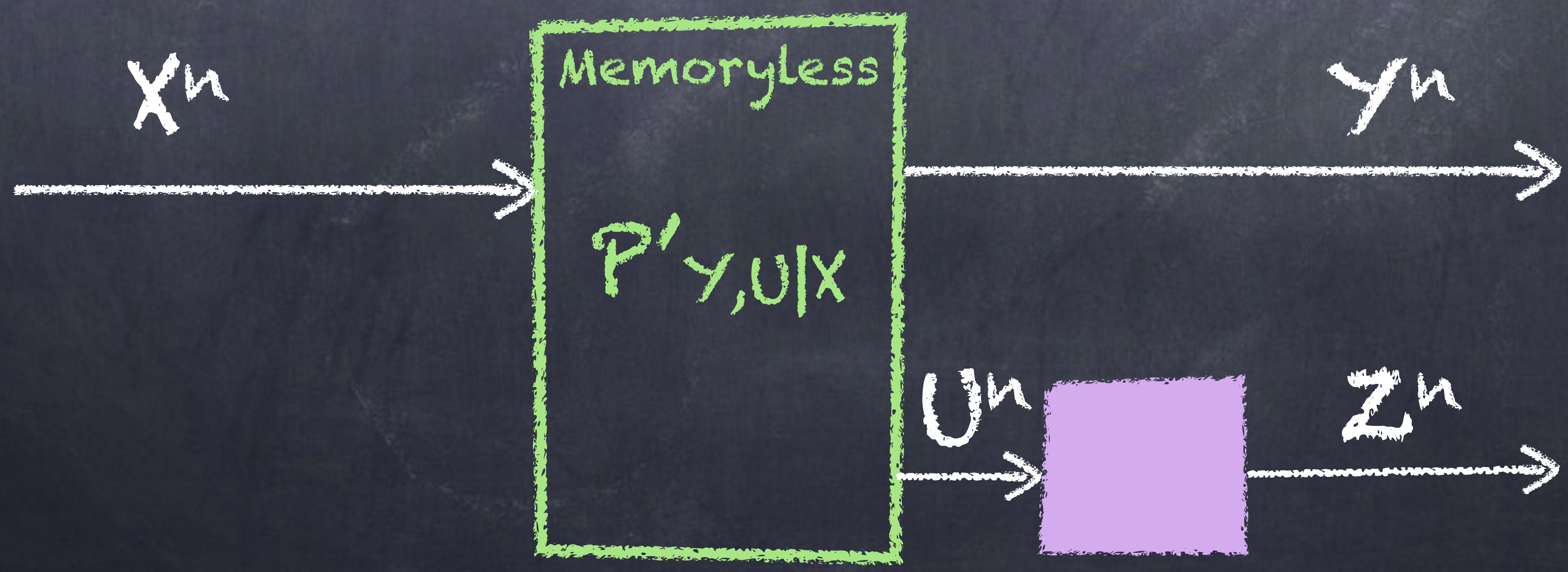
Synthetic noise

- RDSS with full causal disclosure extracts an intuitive communication system



Synthetic noise

- RDSS with full causal disclosure extracts an intuitive communication system



Embedding Information

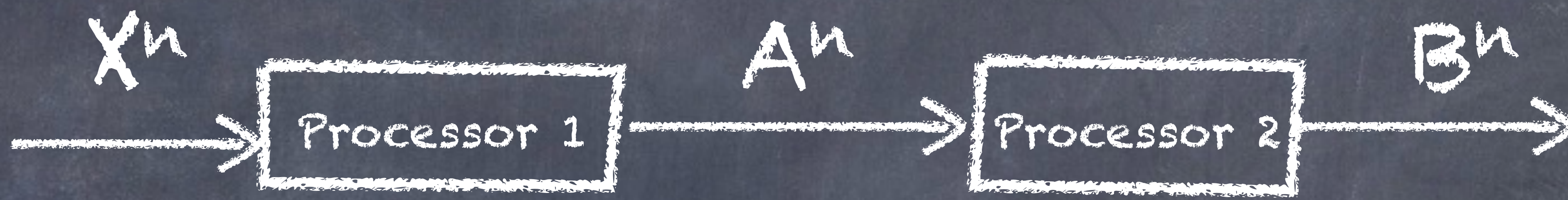
Online Penny Matching

- Two players attempt to guess a binary sequence (both must be correct to earn a point).
- The sequence is revealed to Player 1.
- [Gossner et. al.]

Optimal Score

- $h(V^*) + (1 - V^*) \log_2 3 = 1$
- $V^* \approx 0.82$

General Solutions



		d_2			
		$-\infty$	0	$k>0$	∞
d_1	$-\infty$	$H(A) \geq I(X;A,B)$	$X-(A,U)-B$ $H(A) \geq I(X;A,U)+I(A;U)$	$H(A) \geq I(X;A,B)+I(A;B)$	$X \perp B$
	0	$X \perp U, \quad H(A\|X,U)=0$ $H(A) \geq I(X;A,B\|U)$	$X-A-B$	$X \perp B$	$X \perp B$
	$k>0$	$X \perp A$ $H(A) \geq I(X;A,B)$	$X \perp (A,B)$	$X \perp (A,B)$	$X \perp (A,B)$
	∞	$X \perp (A,B)$	$X \perp (A,B)$	$X \perp (A,B)$	$X \perp (A,B)$