

• Review

- Channel Coding Theorem for DMC.

R is achievable iff $R \leq \max_{p(x)} I(X;Y)$

Achievability $\Rightarrow$

Converse $\Leftarrow$

- Achievability

(i) Find $\arg \max_{p(x)} I(X;Y)$

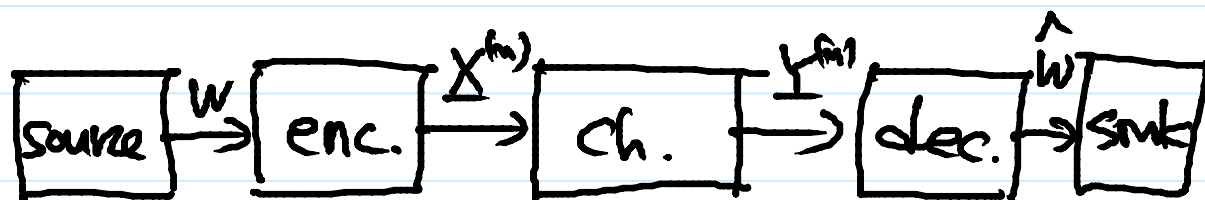
(ii) Find the best code $\mathcal{C}_n$ such that

uniform r.v. W

random encoder

typical-set decoder

$P_e(\mathcal{C}_n)$ is the minimum among all r. codes.



(iii) Find $\tilde{\mathcal{C}}_n$ from $\mathcal{C}_n$ by removing $\approx$ half of the codewords.

$$2^{NR}/2 = 2^{NR-1} = 2^{N(R-\frac{1}{n})}$$

(iv) Choose $\epsilon_k \rightarrow 0$

For each ϵ_k , we can find $\hat{\Sigma}_{nk}$.

$\therefore$ Then, $(\hat{\Sigma}_{nk})_k$ has $R_k \rightarrow R$
 $\epsilon_k \rightarrow 0$.

- Converse

- Fano's Inequality
- Data processing inequality
- DO NOT USE the joint AEP!!
- Independence upper bound
 $I(X; Y) \leq H(Y) - H(Y|X)$
- Chain rule
 $H(Y_1, Y_2, \dots, Y_n | X)$
 $= \sum_{i=1}^n H(Y_i | Y_1, \dots, Y_{i-1}, X)$

$$H(W) = H(W | \hat{W}) + I(W; \hat{W})$$

$$= H(W | Y^{(n)}) + I(W; Y^{(n)}) \quad \text{DPI.}$$

$$= H(W | Y^{(n)}) + I(X^{(n)}; Y^{(n)})$$

$$\leq 1 + nR \cdot P_{e,n} + I(X^{(n)}; Y^{(n)})$$

$$< 1 + nR \cdot P_{e,n} + nC \quad W \rightarrow X^{(n)} \rightarrow Y^{(n)}$$

Fano's inequality

- 1. 10.10.10.10

$$K^m \rightarrow W \rightarrow X^{(m)}$$

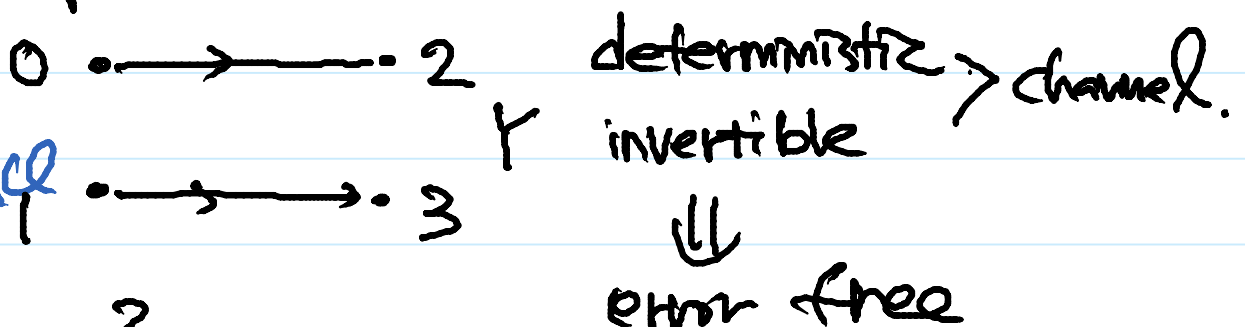
$$nR \leq 1 + nR \cdot P_{e,n} + nC$$

$$\Rightarrow R \leq \frac{1}{n} + R \cdot P_{e,n} + C$$

$$\Rightarrow R \leq C \triangleq \max_{p_X(x)} I(X;Y)$$

- How to solve $\max_{p_X(x)} I(X;Y)$

- Example 1



Experience
 Lesson.

$$C \stackrel{?}{=} \frac{1}{n} \text{ [bits/channel use]}$$

$$\max_{x \in \Omega} f(x)$$

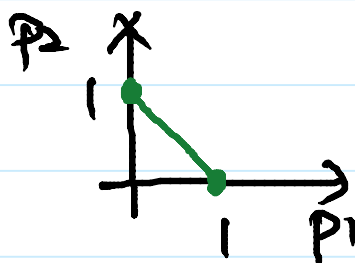
$$\max_{\underline{x} \in \Omega} f(\underline{x})$$

$$\begin{bmatrix} p_X(0) \\ p_X(1) \end{bmatrix} = \underline{x}$$

$$0 \leq x_1 \leq 1$$

$$0 \leq x_2 \leq 1$$

$$x_1 + x_2 = 1$$



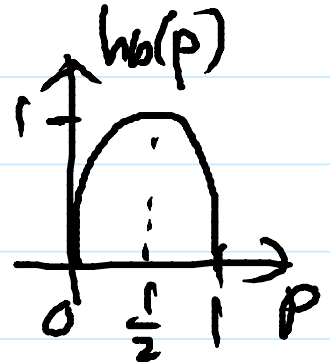
$$I(X;Y) = H(X) - H(X|Y)$$

$$= H(Y) - \cancel{H(X|Y)}$$

$$= H(X)$$

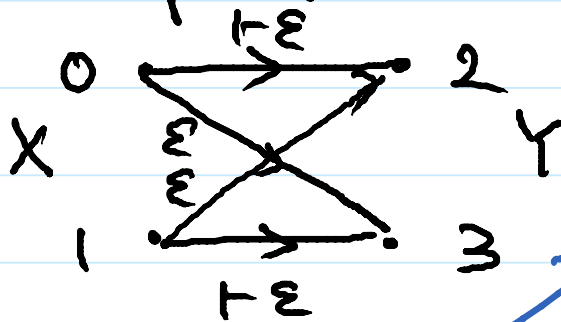
$$= h_b(x_1) \leq 1$$

$$= \text{holds iff } x_1 = \frac{1}{2}$$



$$\therefore C = 1 \text{ [bit/channel use]}$$

Example 2



Binary Symmetric Channel

$$\epsilon = 0 \Rightarrow C = 1$$

$$\epsilon = 1 \Rightarrow C = 1$$

$$0 < \epsilon < 1 \Rightarrow C \leq 1$$

$$\epsilon = \frac{1}{2} \Rightarrow C = 0$$

what we know

intuition

symmetry

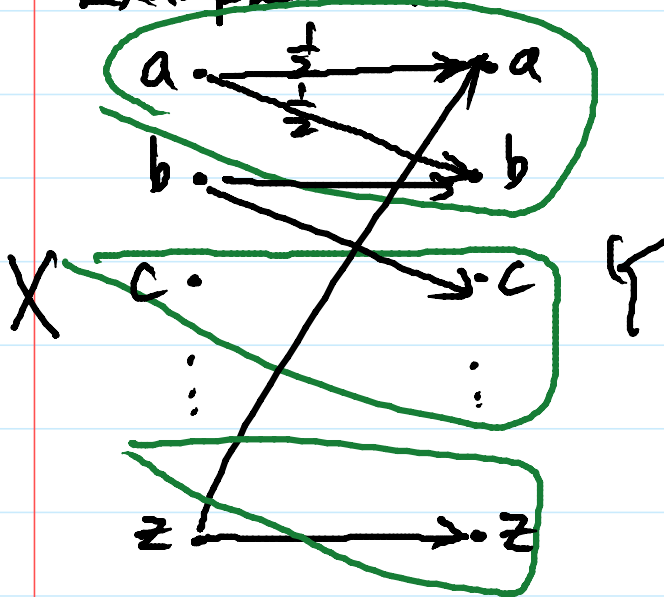
$$\therefore C = 1 - h_b(\epsilon) \text{ [bits/ch. use]}$$

$$\text{sol)} \max_{p_X(x)} I(X;Y)$$

$$\equiv \max_{x \in \mathbb{R}^2} I(X;Y)$$

s.t. $0 \leq x_1, x_2 \leq 1, \quad x_1 + x_2 = 1.$

- Example 3



$A \sim Z : 26$
 $T \sim O : 14$
 $T \sim I : 10$
 $\Delta \neq O : 4$

Noisy Typewriter Ch.
 (can be viewed as a BSC w/ $\epsilon = 1/2$)
generalization

$$P_{Y|X} = \begin{bmatrix} P_{Y_1|X_1} & P_{Y_2|X_2} & \dots \end{bmatrix}$$

BSC

$$\begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}$$

NTC

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \vdots & \vdots \end{bmatrix}$$

• Symmetric Channel

$P_{Y|X}$ matrix has each row obtained by permuting the 1st row $\underline{r}$.

$$\begin{aligned} \max_{P_X(x)} I(X;Y) &= \sum_x H(Y|X=x) P(X=x) \\ &= H(\underline{r}) \\ &= \sum_i r_i \log \frac{1}{r_i} \end{aligned}$$

$$\equiv \max_{P_X(x)} H(Y) - H(Y|X)$$

$$\leq \log_2 |Y| - H(\underline{r})$$

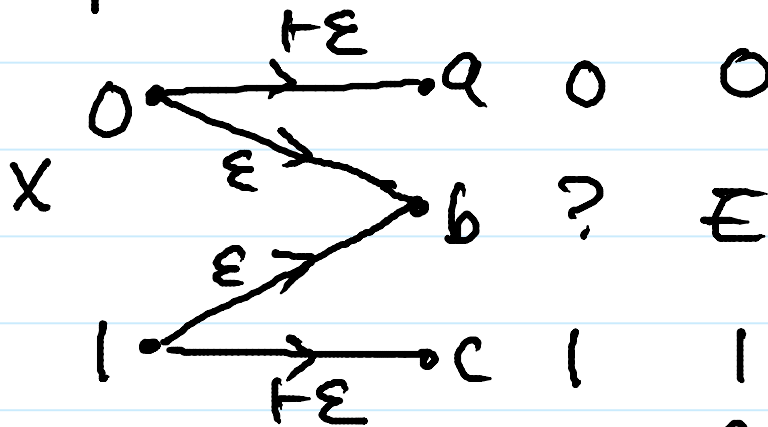
ex/ BSC $\Rightarrow C \leq \log_2(2.3) - h_b(\epsilon)$
 $= 1 - h_b(\epsilon)$

if $P_X(0) = P_X(1) = \frac{1}{2}$.

NTC $\Rightarrow C \leq \log_2 26 - h_b(\frac{1}{2})$
 $= \log_2 26 - 1 = \log_2 13$

if $P_X(a) = P_X(b) = \dots = P_X(z) = \frac{1}{26}$.

Example 4



Binary Erasure Channel (BEC)
input

$$P_{Y|X} = \begin{bmatrix} 1-\epsilon & \epsilon & 0 \\ 0 & \epsilon & 1-\epsilon \end{bmatrix}$$

$$C \leq \log_2 3 - h_b(\epsilon)$$

(frowning face icon)

$$C_{\text{BEC}} = 1 - \epsilon \text{ [bits/ch.use]}$$

