ST06 LECTURE 11

Note Title 3/16/2006

Today

Shannon's Coding Theorem (for symmetric channels)

Review of last lecture

Defined Capacity of discrete memoryless channel

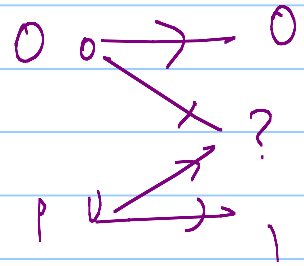
Channel specified by Ω_X, Ω_Y &

$$\left\{ P_{Y|X}(y|x) \right\}_{\substack{y \in \Omega_Y \\ x \in \Omega_X}}$$

Capacity $\quad Cap = \max_{P_X} \left\{ I(X;Y) \right\}$

Example Channels

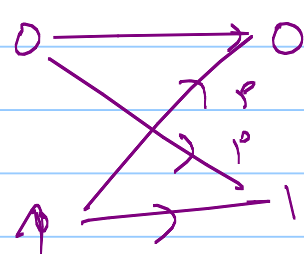
- BEC



? prob. = p.

Cap. = $1-p$

- BSC



Crossover prob. = p

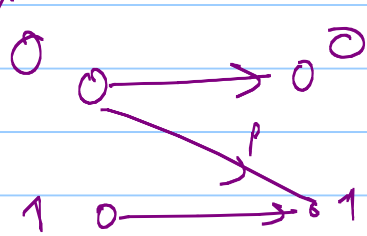
Cap = $1 - H(p)$

- Symmetric Channel (row/column permutations)

$$\text{Cap.} = |\Omega_Y| - H(\text{row}) .$$

Always P_X = uniform

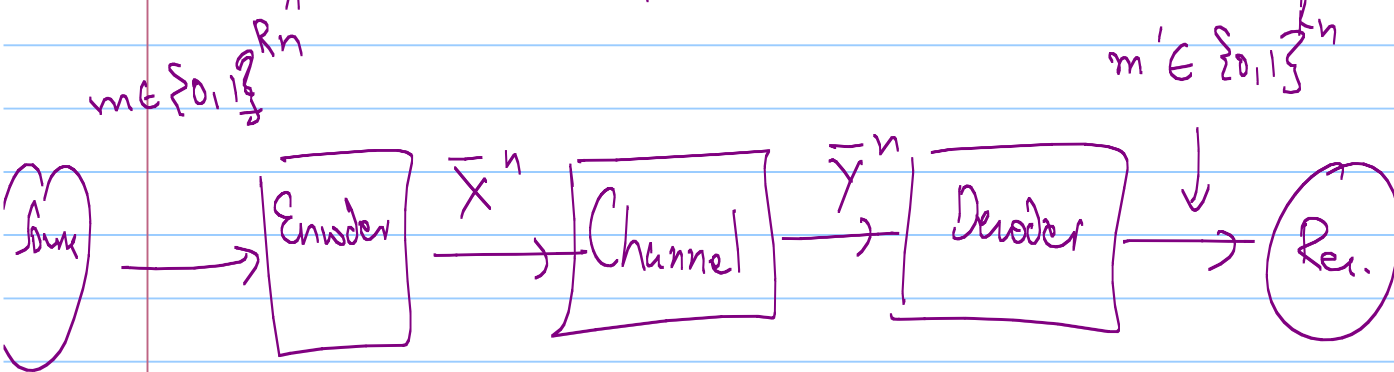
- Example



[Cap ... needs to
be worked out,
by $P_X \neq$ uniform)

Today : realize capacity

"Design" Encoder / Decoder s.t.



Need functions

$$E : \{0,1\}^{Rn} \longrightarrow \Omega_x^n$$

$$D : \Omega_y^n \longrightarrow \{0,1\}^{Rn}$$

So that for

$\bar{X}^n = E(m)$; $\bar{Y}^n = \text{Channel}(\bar{X}^n)$;

$m' = D(\bar{Y}^n)$ we have $m' = m$.

- Can't do this even for erasure channel!
  w.p. p^n $\quad y^n = ?^n \rightarrow$ no info on $\bar{x}^n$;
- So allow prob. of error.

$$P_r[m' \neq m] \text{ is small}$$

- What should E be?
- What should D do?

For fixed E, D may as well be:

$$m' = \arg\max_m \left[\text{Prob}[m \text{ is transmitted}] \cdot \text{Prob}[\bar{y}^n \text{ is rec'd} \mid X^n = E(m) \text{ is trans}] \right]$$

Very hard computationally; but math. well-defined.

But how do choose E ?

Some issues: $k = R \cdot n = (1-p) \cdot n$

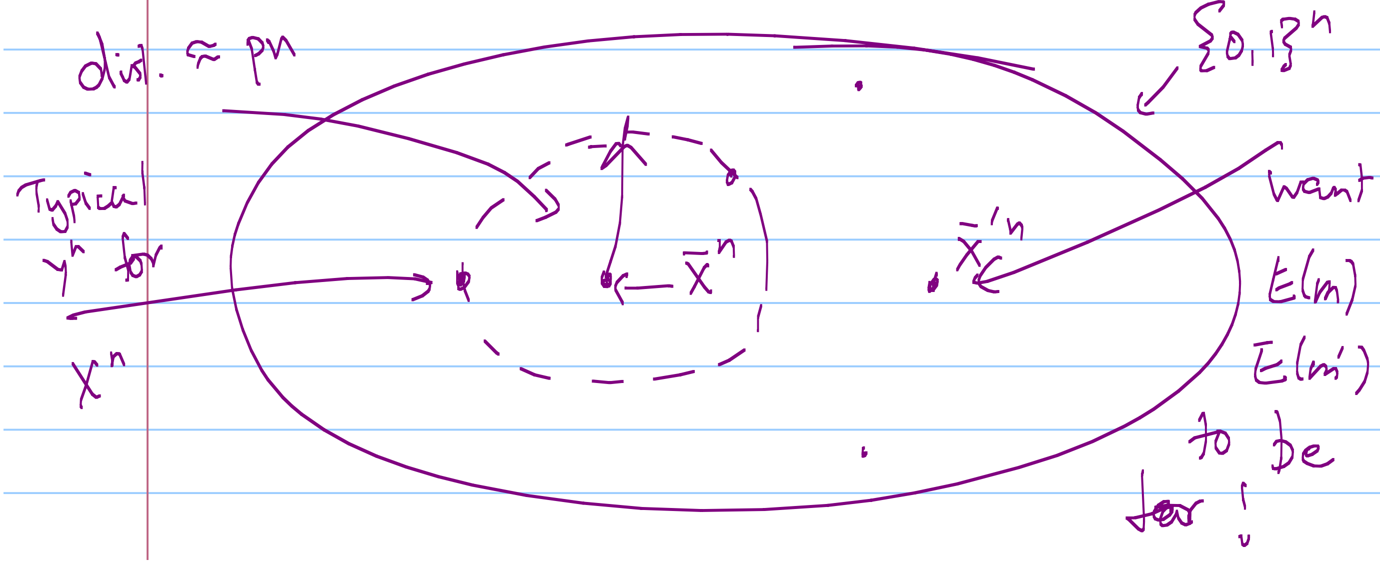
BEC: $m_1 \dots m_k$

Receive: $\{Y_j\}_{j \in S}$ $\quad |S| \approx Rn = (1-p)n$

Need $I(m_1 \dots m_k ; \{Y_j\}_{j \in S}) = k$

for most S !

BSC: $m_1 \dots m_k \to X_1 \dots X_n$



- How can we design such E's?
- Need a new one for each channel?
- BSC: Can we achieve any $R > 0$?

Shannon's remarkable "E"

(Works only for symmetric channels)

for every $m \in \{0,1\}^k$ $[k = R \cdot n]$

$R < C$

- pick $E(m) \in_R \Omega_x^n$.

- Independent for every $m \in \{0,1\}^k$.

$\boxtimes$!!

Main Lemma:

$\forall$ Symmetric Channels $P_{Y|X}$ with capacity C

$\forall$ $R < C$

$$\lim_{n\to\infty} \left\{ \Pr_{\substack{\overline{m}\in\{0,1\}^{Rn} \\ E : \{ \} \to \Omega_x^n \\ \overline{Y}}} \left[m \neq D(\overline{Y}) \right] \right\} \to 0$$

$\Downarrow$

$$\lim_{n\to\infty} \left\{ \min_{E} \left\{ \Pr_{\substack{\overline{m} \\ \overline{Y}}} \left[m \neq D(\overline{Y}) \right] \right\} \right\} \to 0$$

Why does this E work?

in case of BEC $\Rightarrow$ spreads inf. uniformly

in case of BSC $\Rightarrow$ spreads codewords evenly

How to analyse?

From receivers' perspective

sees $\bar{Y}^n \rightarrow$

tries $X^n(1) \longrightarrow Y^n$?

$X^n(2) \longrightarrow Y^n$

$X^n(m) \longrightarrow Y^n$

$\vdots$

$X^n(2^k) \longrightarrow Y^n$?

How to pick message?

What can we expect Y to look like given m?

1. Y^n is equally likely to be a

"typical" received word given x

(E.g. in BSC ... $Y_i = X_i$ w.p. $1-p$

$= \overline{X_i}$ w.p. p)

$\Rightarrow$ Y_i will be like X_i in all but pn places.

etc.)

2. Can it also be a typical received word for $m' \neq m$?

But note:

1. Y^n is distributed uniformly over Ω_Y^n

(Using Symmetry)

② $E(m')$ independent of Y^n

depends on $E(m)$; Channel; but not $\underline{\underline{E(m')}}$!

$$\Pr\left[Y^n \text{ typical for } E(\bar{m}) \right] \le \frac{|\text{Size of typical set for } E(\bar{m})|}{(|\Omega_Y|)^n}$$

<u>Defn</u>: for any $\bar{x} \in \Omega_x^n$

typical Set $A^{(n)}_{\epsilon,\bar{x}} = \left\{ \bar{y} \;\middle|\; \Pr[\bar{y} | \bar{x}] \approx 2^{-(H(\bar{r}) \pm \epsilon) n} \right\}$

<u>Claim</u>: $\left| A^{(n)}_{\epsilon,\bar{x}} \right| \le 2^{(H(\bar{r}) + \epsilon) \cdot n}$

C = Capacity $= \log|\Omega_y| - H(\bar{r})$

$R < C \Rightarrow$ "

For every $\bar{y}$ $\left(\in A^{(n)}_{\epsilon, E(m)}\right)$.

<u>Claim</u>: $\Pr\left[\exists\, m' \neq m \;\middle|\; y \in A^{(n)}_{\epsilon, E(m')}\right]$

$$\leq \frac{2^{R \cdot n} \sum_{m'} \left|A^{(n)}_{\epsilon, E(m')}\right|}{|\Omega_y|^n}$$

$\longrightarrow 0$ as $R < C$

& $\epsilon \leq \frac{C - R}{3}$

⋮

Ready to Prove Shannon's Theorem

Events

E1
- Pick $X^n = E(m)$ at random ; receive $\overline{Y}^n$ ...
- $E1: \quad \overline{Y}^n \notin A^{(n)}_{\epsilon, E(m)}$

$$\Pr[E1] \longrightarrow 0 \quad \text{as} \quad \begin{array}{l} n \to \infty \\ \epsilon \to 0 \end{array}$$

E2
- Fix $E(m)$, $\overline{y}^n$
- Now pick $E(m') \quad \forall m' \neq m$
- $\underline{E2}$: $\exists m' \neq m$ s.t. $\overline{y} \in A^{(n)}_{\epsilon, E(m')}$

$$\Pr[E2] \to 0$$

if neither E1 nor E2 occur

then $m = D(\bar{y}^n)$

Proves Shannon's Coding Theorem

- But ONLY for symmetric channel.
- Covers BSC ✓

Noisy typewriter (Big Deal!)

But not BEC

Why not BEC ?

E still works but analysis needs to change

- $|\Omega_Y|^n$ is irrelevant ....

- $\bar{Y}^n$ is not random

Broader Cases :

" $E(m)$ uniform in Ω_X^n "

is not right !

(Channel capacity may not be maximized by unif. dist on X^n; then can't be that $E(m)$ varying from that can achieve capacity) ;

Next Lecture

Will consider

"E" picked as follows

Let P on Ω_X be distribution

maximizing $I(X;Y)$.

Then for every $m \in \{0,1\}^k$, $i \in \{1 \ldots n\}$

$E(m)_i$ are i.i.d with dist P.

- To analyze this decoding will consider

joint dist. on $(\bar{X}^n, \bar{Y}^n)$

where $\bar{X}^n$ dist according to P^n

$\bar{Y}^n$ dis according to $(P_{Y|X})^n$

---x---

Decoding procedure: $\bar{Y}$

if $(E(m), \bar{Y})$ typical for dist.;

& $(E(m'), \bar{Y})$ not typical for dist;

then decoding of $\bar{Y} = m$;

else ERROR.

---x---

Will refine E1, E2 } for this setting

$\boxtimes$