EECS 126 — 2.25 LECTURE 11

ACHIEVABILITY

We'll use the AEP to prove this part of the Source Coding Theorem.
Fix $\varepsilon > 0$, and for each $n \geq 1$ define the "typical set":

$$A_\varepsilon^{(n)} = \left\{ (x_1 \dots x_n) : P(x_1 \cdots x_n) \geq 2^{-n(H(X)+\varepsilon)} \right\}$$

= subset of possible observed sequences

Properties: $P((x_1 \cdots x_n) \in A_\varepsilon^{(n)}) \longrightarrow 1$ as $n \to \infty$, by AEP

$|A_\varepsilon^{(n)}| \leq 2^{n(H(X)+\varepsilon)}$, because $1 \geq \sum_{(x_1 \cdots x_n) \in A_\varepsilon^{(n)}} P(x_1 \cdots x_n) \geq \sum 2^{-n(H(X)+\varepsilon)}$
$= |A_\varepsilon^{(n)}| 2^{-n(H(X)+\varepsilon)}$

With n objects, how many bits to store each? $\log n$
Source Coding protocol: if we observe $(x_1 \cdots x_n) \in A_{\varepsilon/2}^{(n)}$, we describe it
using $\log |A_{\varepsilon/2}^{(n)}|$ bits $\leq n(H(X) + \varepsilon/2)$
if we observe $(x_1 \cdots x_n) \notin A_{\varepsilon/2}^{(n)}$, brute force $n \log |\mathcal{X}|$

$$E[\#\text{bits}] = n(H(X) + \varepsilon/2) \underbrace{P((x_1 \cdots x_n) \in A_{\varepsilon/2}^{(n)})}_{\hookrightarrow \leq 1} + n \log |\mathcal{X}| \underbrace{P((x_1 \cdots x_n) \notin A_{\varepsilon/2}^{(n)})}_{\leq \varepsilon/2 \text{ for large } n}$$

$$\leq n(H(X) + \varepsilon)$$

This proves achievability. The converse is true, but the proof is omitted.

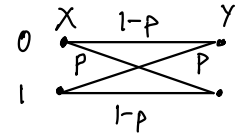
INFORMATION TRANSMISSION

How do we reliably send data over an unreliable channel? This is channel coding.
Fix some "rate" $R > 0$.

message $\longrightarrow$ encoder $\longrightarrow$ noisy channel $\longrightarrow$ decoder
$M \sim \text{Uniform}(\{1 \dots 2^{nR}\}) \longrightarrow X^n(M) \longrightarrow Y^n \longrightarrow \hat{M}(Y^n)$
↳ nR bits for $H(M)$ ↳ $(X_1(M), \dots, X_n(M))$ ↳ $(Y_1 \dots Y_n)$
↳ Y_i is some noisy X_i

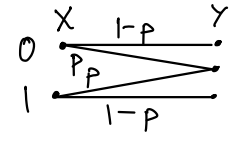
Parameters: $R = \frac{H(M)}{n} = \frac{\#\text{ info bits}}{\#\text{ channel uses}}$
$P_e^{(n)} = \Pr\{\hat{M} \neq M\}$ = error probability

Types of noisy channels: BSC(p) binary symmetric



flip wp. p independently

BEC(p) binary erasure



erase wp. p independently

Usually, channels are represented by PMFs. Example, BSC: $P_{Y|X}(y|x) = \begin{cases} p & y \neq x \\ 1-p & y = x \end{cases}$

Mutual Information: for a channel $P_{Y|X}$ and input distribution P_X:

$P_{XY}(x,y) = P_X(x)\,P_{Y|X}(y|x)$ joint dist. of in & out

$P_Y(y) = \sum_x P_{XY}(x,y)$ marginal dist. of outputs

then, mutual info $I(X;Y) = \sum P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}$

Capacity: for a channel $P_{Y|X}$, $C = \max_{P_X} I(X;Y)$ max mutual info between input & output over all input distributions

doesn't depend on n!

THEOREM Shannon's Channel Coding Theorem: fix a channel $P_{Y|X}$, $\varepsilon > 0$, $R < C$. For all n sufficiently large, there exists a rate-R communication scheme that achieves $P_e < \varepsilon$. If $R > C$, then $P_e \to 1$ for any sequence of schemes.

We'll prove this for $BEC(p)$ but not in general. $C = 1-p$

PROOF 1) $R > C$ is not possible: consider n channel uses. suppose we can somehow know which bits are erased. then, transmitter can just distribute M over the unerased bits. we know we have $\leq n(1-p+\varepsilon)$ unerased positions, so we can't reliably send more than $n(1-p)$ bits $\longrightarrow R \leq 1-p$.

2) all $R < C$ allow reliability: no need to analyze an explicit scheme, just show one exists! don't need to construct it either. fix $\varepsilon > 0$ & $R < 1-p-\varepsilon$. generate random matrix $C = \left[\begin{smallmatrix} & \\ n & \end{smallmatrix}\right]^{2^{nR}}_{1}$ where $C_{ij} \sim_{IID} Bern(1/2)$. give to both encoder and decoder. send row M of C, and the receiver matches Y in C (mod erasures). this will error if >1 row matches.

we will continue next time.