Last time:

Def

For a channel N, a rate R is achievable if $\exists$ sequence of $(\lfloor 2^{nR} \rfloor, n)$ codes C_n s.t. $P_e(C_n) \to 0$ as $n \to \infty$.

Def The capacity of N, $C(N)$, is the supremum over achievable rates.

Thm: $C(N) = \max_{p(x)} I(X:Y)$

Direct coding thm:

For any $p(x)$, $\exists (M, n)$ codes C_n s.t.

- $\frac{1}{n} \log M \geq I(X:Y) - 3\eta - \frac{1}{n} \log\left(\frac{1}{\eta}\right) - \frac{1}{n}$ for any $\eta > 0$
- $P_e(C_n) \leq 4\eta$.

We need $\eta \to 0$ as $n \to \infty$. $\varepsilon < \eta$, $\delta < \eta$ in HEP.

We can choose $\delta = n^{-\frac{1}{8}}$, $\varepsilon = n^{-\frac{1}{4}}$, $n_0 \geq (\mathrm{Var}\, Y)^2$ in the AEP.

Converse (today):

For any achievable R, $R \leq \max_{p(x)} I(X:Y)$

Contrapositive:

If $R > \max_{p(x)} I(X:Y)$, any sequence of $(\lfloor 2^{nR} \rfloor, n)$ codes has $P_e(C_n) \not\to 0$.

Strong converse: $1 - P_e(C_n) \sim \exp(-n \times \text{const})$.

Pf of converse:

Let R be achievable

Let C_n be the sequence of $(\lfloor 2^{nR} \rfloor, n)$ codes with $P_e(C_n) \to 0$

Let W_n be a rv describing a random element in $\{1, 2, \ldots, \lfloor 2^{nR} \rfloor\}$ (the index set for messages from C_n).

Let $p_n = P_e(C_n)$

The following holds from definition:

$$nR = H(W_n) = \underbrace{H(W_n|Y^n)}_{\textcircled{1}} + \underbrace{I(W_n : Y^n)}_{\textcircled{2}}$$

$$\textcircled{1} \leq 1 + p_n nR$$

$$\textcircled{2} \leq I(\mathcal{E}_n(W_n) : Y^n)$$

$$\textcircled{3} \leq n \max_{p(x)} I(X:Y)$$

Together $nR \leq 1 + p_n nR + n \max_{p(x)} I(X:Y)$

Divide by n, take $n \to \infty$, $p_n \to 0$, we have

$$R \leq \max_{p(x)} I(X:Y).$$

① Thm [Fahos ineq]

Consider r.v.s A, B, C, $C = f(B)$, f function

Let q = prob $(A \neq C)$

Ω = sample space of A

Then $h(q) + q \log(|\Omega| - 1) \geq H(A|B)$.

Pf: Define new rv E s.t $E = \begin{cases} 0 & \text{if } A = C \\ 1 & \text{otherwise} \end{cases}$

$$H(EA|B) = H(A|B) + H(E|AB) = H(E|B) + H(A|EB)$$

$H(EAB) - H(B)$ $\quad$ $H(AB) - H(B)$ $\quad$ $H(EAB) - H(AB)$ $\quad$ exchange E & A

$$\therefore \quad H(A|B) + 0 = H(E|B) + H(A|EB)$$

$$\leq H(E) + \sum_b p(b) \left[q\, H(A|E=1, B=b) + (1-q) H(A|E=0, B=b) \right]$$

$$\therefore \quad H(A|B) \leq h(q) + q\, H(A|E=1, B)$$

$$\leq h(q) + q \log\left[|\Omega| - 1\right]$$

- Set $A = W_n$, $B = Y^n$, $|\Omega| = 2^{nR}$, $f = D_n$, $q = p_n$

Then $H(W_n|Y^n) \leq h(p_n) + p_n \cdot nR$.

$$\leq 1 + p_n \cdot nR.$$

② Thm [Data processing ineq (DPI)]

If 3 r.v's A, B, C form a Markov chain $A \to B \to C$ (processing)

ie $I(A:C|B) = 0$

then $I(A:B) \geq I(A:C)$

Pf: in notes of lec 08, reading ex.

NB: Exchanging A & C:

$I(B:C) \geq I(A:C)$ ⊛

- Set $A = W_n$, $B = \xi_n(W_n)$, $C = Y^n$

then $I(A:C|B) = 0$

So $I(\xi_n(W_n) : Y^n) \geq I(W_n : Y^n)$

NB: $I(A:C|B) = 0 \iff H(A|B) = H(A|BC)$

$\iff H(C|B) = H(C|AB)$ from (I 10)

Given $B = \xi_n(W_n)$

$C = Y^n = D_n \circ N^{\otimes n} \circ \xi_n(W_n)$

independent of $A = W_n$

③ Lemma: Let $Y^n = N^{\otimes n}(X^n)$

Then $I(X^n : Y^n) \leq \sum_{i=1}^{n} I(X_i : Y_i)$

NB: neither X^n nor Y^n need to be iid

cf. X^n from codewords, $C_i = \underbrace{x_{i1} x_{i2} \cdots x_{in}}_{\text{concluded}}$

Pf: $I(X^n : Y^n) = H(Y^n) - H(Y^n | X^n)$

$= H(Y^n) - \sum_{i=1}^{n} H(Y_i | Y_1 Y_2 \cdots Y_{i-1} X^n)$ (chain rule with cond.)

$= H(Y^n) - \sum_{i=1}^{n} H(Y_i | X_i)$ with X_i, Y_i does not depend on the rest.

$\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i | X_i)$ subadditivity

$= \sum_{i=1}^{n} I(X_i : Y_i)$.

NB: All of ①–③ unaffected if Bob has free communication (classical) to Alice, used in arbitrary way!

∴ Feedback does not increase capacity. $C_B(N) = C(N)$.

It can ↓ coding / decoding complexity. eg. erasure channels