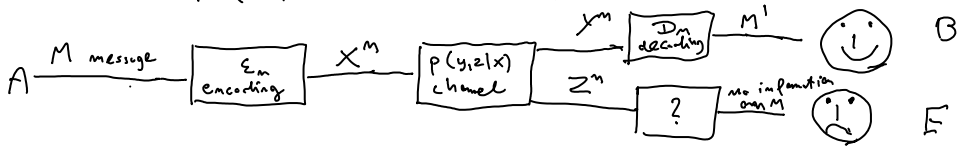
4 - Secret capacity

20 października 2010
15:11

4.1 Setup (confidential communication)



We want $M' = M$, and be sure that E has no information on message M.
This means A and B can communicate secretly.

Intuition: this is possible provided channel $X \to Y$ "is better" than $X \to Z$

4.2 Csiszár – Körner theorem

Def R is an achievable rate of confidential communication if there exist $(2^{mR}, n)$ codes such that probability of error $\lambda_n \xrightarrow{n\to\infty} 0$ $(M \neq M')$ and assuming M are uniformly distributed:

$\frac{1}{n} H(M|Z^n) \xrightarrow{n\to\infty} R \left[\frac{1}{n} H(M)\right]$ (no knowledge about M for somebody knowing Z^n)

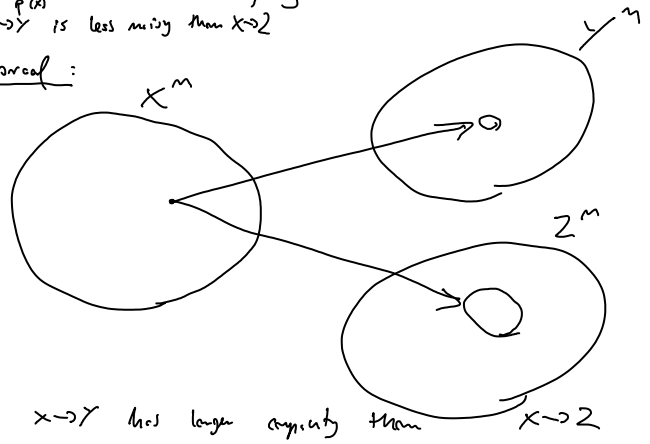
Theorem:
R is achievable iff:

$$R < C_s = \max_{p(x)} \left[I(X:Y) - I(X:Z) \right]$$

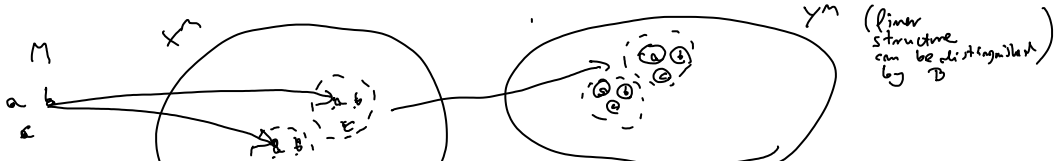
secret capacity

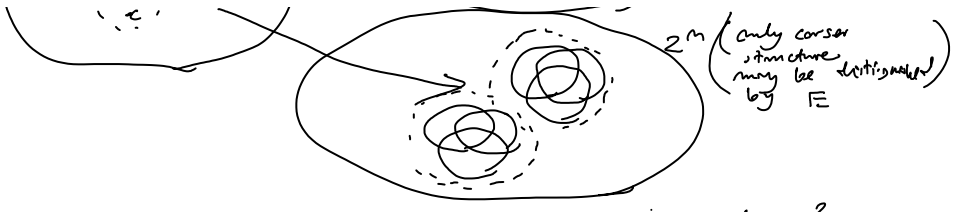
[provided $\forall_{p(x)}$ $I(X:Y) \geq I(X:Z)$]
channel $X \to Y$ is less noisy than $X \to Z$

Intuitive proof:



channel $X \to Y$ has longer capacity than $X \to Z$





How many messages can we encode into the finer structure?

$2^{nI(X:Y)}$ - number of messages that can be distinguished by B

$2^{nI(X:Z)}$ - " " E

$\frac{2^{nI(X:Y)}}{2^{I(X:Z)}}$ - number of messages encoded into the finer structure

$2^{n(I(X:Y)-I(X:Z))}$

Proof:

($\Rightarrow$) Provided $R < C_S$ we construct the code explicitly

$|M| = 2^{nR}$ $\quad M \xrightarrow{\text{privacy amplification}} (J, M) \xrightarrow{\text{coding}} X^n \to (Y^n, X^n)$

J - uniformly distributed random variable $j = 1 \dots |J|$

$m \xrightarrow{\mathcal{E}} x^n(j,m)$ (coarser structure, finer structure)

Our code: we generate $|J| \cdot |M|$ random sequences $x^n(j,m)$

From Shannon theorem we know we can take $|J| \cdot |M| = 2^{n(I(X:Y)-\varepsilon)}$ to have error-free communication. We take $|J| = 2^{n(I(X:Z)-\varepsilon)}$ and want to show that:

$$\frac{1}{n} H(M|Z^n) \to \frac{1}{n} H(M) = \frac{1}{n} \log |M| = R$$

calculate $H(M|Z^n)$

$$p(z^n|m) = \frac{1}{|J|} \sum_j p(z^n | x^n(j,m))$$

$$H(M|Z^n) = H(M, Z^n) - H(Z^n) = H(M, X^n, Z^n) - H(X^n|M, Z^n) - H(Z^n)$$

$$= H(M, X^n) + H(Z^n|MX^n) - H(X^n|MZ^n) - H(Z^n) \geq$$

$$\geq \underbrace{H(X^n)}_{(a)} + \underbrace{H(Z^n|X^n)}_{(b)} - \underbrace{H(X^n|MZ^n)}_{(c)} - \underbrace{H(Z^n)}_{(d)}$$

a) For each codebook we have $|J||M|$ sequences $x^n(j,m)$ provided they are all different we would have $H(X^n) = \log |J||M|$
We have to control somehow the problem of choosing randomly the same sequences

For given code $\mathcal{E}$

$$H(X^n) = -\sum_{x^n} p(x^n) \log p(x^n) \geq -\log \sum_{x^n} p(x^n)^2 \qquad p(x^n) = \sum_{j,m} \delta(x^n, \mathcal{E}(j,m)) \frac{1}{|J||M|}$$

$\sum_{x^n} p(x^n)^2$ - collision probability $\Pr[X^n = \bar{x}^n]$

$$\overline{H}(X^n) = -\sum_{\mathcal{E}} p(\mathcal{E}) H_{\mathcal{E}}(X^n) \geq -\sum_{\mathcal{E}} p_{\mathcal{E}} \log\left(\sum_{x^n} p_{\mathcal{E}}(x^n)^2\right) \geq -\log \sum_{\mathcal{E}} p_{\mathcal{E}} \sum_{x^n} p_{\mathcal{E}}(x^n)^2 =$$

$$= -\log \sum_{\mathcal{E}} p_{\mathcal{E}} \sum_{x^n} \sum_{j,m} \delta(x^n, \mathcal{E}(j,m)) \delta(x^n, \mathcal{E}(\bar{j},\bar{m})) \frac{1}{|J||M|} = -\log \sum_{\mathcal{E}} p_{\mathcal{E}} \sum_{j,m} \delta(\mathcal{E}(j,m), \mathcal{E}(\bar{j},\bar{m})) \frac{1}{|J||M|} =$$

$$= -\log \sum_\varepsilon p_\varepsilon \sum_{x^n} \sum_{\substack{j,m \\ j',m'}} \delta(x^n, \varepsilon(j,m))\,\delta(x^n, \varepsilon(j',m'))\,\frac{1}{|J|M|} = -\log \sum_\varepsilon p_\varepsilon \sum_{\substack{j,m \\ j',m'}} \delta(\varepsilon(j,m), \varepsilon(j',m'))\,\frac{1}{|J|M|^2} =$$

$$= -\log \sum_\varepsilon p_\varepsilon \frac{1}{|J|M|} \sum_{j,m} \delta(\varepsilon(j,m), \varepsilon(0,0)) = -\log \sum_\varepsilon \frac{p_\varepsilon}{|J|M|} \left(1 + \sum_{j,m \neq 0,0} \delta(\varepsilon(j,m), \varepsilon(0,0))\right) \geq$$

$$\geq -\log\left(1 + (|J|M|-1)\,2^{-nH(X)}\right) + \log|J|M|$$

provided $(|J|M|-1) \leq 2^{-n(H(X)+\varepsilon)}$ $\quad \overline{H}(X^n) \xrightarrow{n\to\infty} \log|J|M|$

this is satisfied since $J(X:Y) \leq J(X)$

b) $H(Z^n|X^n) = -\sum_{x^n z^n} p(x^n) p(z^n|x^n) \log \underbrace{p(z^n|x^n)}_{p(z_1|x_1)\cdots p(z_n|x_n)} = \sum_{x^n} p(x^n) \sum_i H(Z|X=x_i)$

$$\overline{H}(Z^n|X^n) = \sum_\varepsilon \sum_{x^n} p_\varepsilon(x^n) \sum_i H(Z|X=x_i) = \sum_i \sum_{x_i} p(x_i)\, H(Z|X=x_i) = n H(Z|X)$$

c) $H(X^n|MZ^n)$ knowing the message M and his sequence Z^n how well can E guess X^n. E can simply use the strategy look for $\tilde{x}^n$ such that $\tilde{x}^n = \varepsilon(j,m')$ (unknown) and $(\tilde{x}^n, z^n)$ is (known) jointly typical

since $|J| = 2^{n(I(X:Z)-\varepsilon)}$ this procedure is asymptotically error free

so X^n is decoded correctly: $H(X^n|MZ^n) \to 0$

d) $H(Z^n) \leq n H(Z)$

$$\frac{1}{n}\overline{H}(M|Z^n) \geq \frac{1}{n}\left(\log|J|M| + n H(Z|X) - n H(Z) - \varepsilon\right) =$$

$$= I(X:Y) - I(X:Z) - \varepsilon = \frac{1}{n}\log|M| - \varepsilon = R - \varepsilon$$

So we proved that averaged over codes $\frac{1}{n}\overline{H}(M|Z^n) \xrightarrow{n\to\infty} \frac{1}{n}\log|M|$

But this means there exists at least one code for which $\frac{1}{n}H(M|Z^n) \xrightarrow{n\to\infty} \frac{1}{n}\log|M|$ ∎

($\Leftarrow$) $M \to X^n \to Y^n Z^n \qquad |M| = 2^{nR}$

Take ie:

$\frac{1}{n}H(M|Y^n) \to 0$ tzn. że B może bez błędu odczytać wiadomość

a $\frac{1}{n}H(M|Z^n) \to R$ tzn że E nie wie

to $R < \max_{p(x)} \left(I(X:Y) - I(X:Z)\right)$

$$H(M|Z^n) = H(M) - I(M:Z^n) = I(M:Y^n) + H(M|Y^n) - I(M:Z^n)$$

$$I(M:Y^n) \overset{\text{chain rule}}{=} \sum_i I(M:Y_i|\underbrace{Y_1 \ldots Y_{i-1}}_{Y^{i-1}}) \qquad I(M:Z^n) = \sum_i I(M:Z_i|\underbrace{Z_{i+1}\ldots Z_n}_{\tilde{Z}^{i+1}})$$

$$I(M:Y_i|Y^{i-1}) = I(M,\tilde{Z}^{i+1}:Y_i|Y^{i-1}) - I(Y_i:\tilde{Z}^{i+1}|Y^{i-1}M) =$$

$$= I(\tilde{Z}^{i+1}:Y_i|Y^{i-1}) + I(M:Y_i|Y^{i-1}\tilde{Z}^{i+1}) - I(Y_i:\tilde{Z}^{i+1}|Y^{i-1}M)$$

similarly:

$$I(M:Z_i|\tilde{Z}^{i+1}) = I(M:Z_i|\tilde{Z}^{i+1}Y^{i-1}) + I(Y^{i-1}:Z_i|\tilde{Z}^{i+1}) - I(Z_i:Y^{i-1}|\tilde{Z}^{i+1}M)$$

we have:

$$H(M|Z^n) = I(M:Y^n) - I(M:Z^n) = \sum_i I(M:Y_i|Y^{i-1}\tilde{Z}^{i+1}) - I(M:Z_i|\tilde{Z}^{i+1}Y^{i-1})$$

$$+\underbrace{\sum_i \left[I(\tilde{Z}^{i+1}; Y_i | Y^{i-1}) - I(Y^{i-1}; Z_i | \tilde{Z}^{i+1}) \right]}_{A} - \underbrace{\sum_i \left[I(Y_i; \tilde{Z}^{i+1} | Y^{i-1} M) - I(Z_i; Y^{i-1} | \tilde{Z}^{i+1} M) \right]}_{B}$$

Fact: $A = B = 0$

$$A = \sum_i \sum_{j=i+1}^{\tilde{n}} I(Z_j; Y_i | Y^{i-1} \tilde{Z}^{j+1}) - \underbrace{\sum_i \sum_{j=1}^{i-1} I(Y^i; Z_i | \tilde{Z}^{i+1} Y^{j-1})}_{i \leftrightarrow j} = 0$$

$B = 0$ analogously

$$H(M|Z^n) = \sum_i I(M; Y_i | U_i) - I(M; Z_i | U_i) \leq \left\{ \text{where } U_i = Y^{i-1} \tilde{Z}^{i+1} \right.$$

$$\leq n \max_i \left[I(M; Y_i | U_i) - I(M; Z_i | U_i) \right]$$

{ Fact: $I(AB; C|A) = I(B; C|A)$

$$= n \cdot \max_i \left[I(\underbrace{MU_i}_{V_i}; Y_i | U_i) - I(\underbrace{MU_i}_{V_i}; Z_i | U_i) \right]$$

we may think: $U_i \to V_i \to X_i \to Y_i$

$$= n \max_{U \to V \to X} \left[I(V; Y|U) - I(V; Z|U) \right]$$

conditioning on U can only lower the mutual value

$$\left\{ \max_{U \to XY} p(X, Y|U) = \sum_u p(u)\, p(X_u, Y_u) \leq \max_{XY} p(X, Y) \right.$$

$$\leq n \max_{V \to X} \left(I(V; Y) - I(V; Z) \right)$$

$$\left\{ I(V; Y) = I(XV; Y) - I(X; Y|V) \right. \quad \text{since } V \to X \to Y,\ I(XV; Y) = I(X; Y)$$

$$\geq I(X; Y) - I(X; Y|V)$$

$$= n \max_{V \to X} \left[I(X; Y) - I(X; Z) - \underbrace{\left(I(X; Y|V) - I(X; Z|V) \right)}_{C} \right]$$

Since channel $X \to Y$ is less noisy than $X \to Z$, $C \geq 0$

hence

$$\leq n \max_X \left(I(X; Y) - I(X; Z) \right) \quad \blacksquare$$

Ciszar-Körner theorem can be view as error-correction + privacy amplification.
We use codes which allow A and B to communicate without errors, but introduce additional randomization such that E knowledge is reduced to zero.