IV.4. Direct part of the coding theorem

Thm.: Let $p(y|x)$ with $x \in \mathcal{X}$, $y \in \mathcal{Y}$ describe a discrete memoryless channel. Every $R < \max_{p(x)} I(X;Y)$ is an achievable rate for it, if the mutual information is computed w.r.t to $p(x,y) := p(x)\,p(y|x)$.

proof:

- fix $\varepsilon > 0$, $n \in \mathbb{N}$, $p(x)$ and let $\mathcal{M} := \{1, \dots, 2^{nR}\}$
- produce "random" $(2^{nR}, n)$ code by generating 2^{nR} codewords in $\mathcal{X}^n$ independently according to $p(x^n) = \prod_{i=1}^{n} p(x_i)$
- use "typical-set decoding" $g : \mathcal{Y}^n \to \mathcal{M}$ defined by

$$g(y^n) = m \quad \text{if} \quad (x^n(m), y^n) \in B_\varepsilon^{(n)} \;\wedge\; \forall j \neq m : (x^n(j), y^n) \notin B_\varepsilon^{(n)}$$

$$g(y^n) = 1 \quad \text{otherwise.}$$

error analysis:

$$\hat{p} := \sum_c p(c)\, p_e^{(n)}(c) \quad \text{averaged over codes}$$

$$= \sum_c p(c)\, 2^{-nR} \sum_{m \in \mathcal{M}} \lambda_m(c) \quad \text{\& codewords}$$

$$= 2^{-nR} \sum_m \underbrace{\sum_c p(c)\, \lambda_m(c)}_{\text{independent of } m} = \sum_c p(c)\, \lambda_1(c)$$

two error types (assuming y^n is received upon sending $x^n(1)$):

(i) $(x^n(1), y^n) \notin B_\varepsilon^{(n)}$: this has prob. at most ε

(ii) $(x^n(j), y^n) \in B_\varepsilon^{(n)}$ for some $j \neq 1$

Since $X^n(1)$ and $X^n(j)$ are independent if $j \neq 1$, so are Y^n & $X^n(j)$

$\overset{\text{joint AEP}}{\Longrightarrow}$ prob. bounded by $2^{-n(I(X;Y)-3\varepsilon)}$ for each $j \neq 1$

$$\Rightarrow \quad \hat{p} \overset{(i)}{\leq} \varepsilon + (2^{nR}-1)\overbrace{2^{-n(I(X;Y)-3\varepsilon)}}^{(*)}$$

$$\leq \varepsilon + 2^{n(R-I(X;Y)+3\varepsilon)}$$

$$\leq 2\varepsilon \quad \text{if} \quad R < I(X;Y) - 3\varepsilon \ \& \ n \text{ suff. large}$$

Hence, if $R < I(X;Y)$ we can choose $\varepsilon > 0$ and $n \in \mathbb{N}$ accordingly and make $\hat{p}$ arbitrary small.

- $\hat{p} \leq 2\varepsilon \Rightarrow \exists c : p_e^{(n)}(c) \leq 2\varepsilon$
- modify this code by discarding the worst 50% codewords

  $\rightarrow$ $(2^{nR-1}, n)$ code $\tilde{c}$ for which the max. prob. of error is

  $$\lambda^{(n)}(\tilde{c}) \leq 2 p_e^{(n)}(c) \leq 4\varepsilon$$

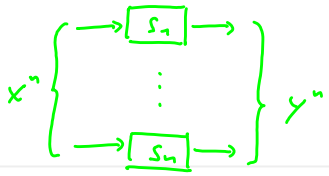
  The rate of $\tilde{c}$ is $\tilde{R} = R - \frac{1}{n}$ □

IV.5. Converse part of the coding theorem

Lemma: Let $x^n \in \mathcal{X}^n$ with distribution $p(x^n)$ be the input and $y^n \in \mathcal{Y}^n$ be the output of an n-fold product of discrete memoryless channels. For $p(x^n, y^n) := \prod_{i=1}^{n} p_i(y_i|x_i)\, p(x^n)$ we get

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

(note that the channels can be different)



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

chain rule ↓

$= H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_{i-1},\dots,Y_1,X^n)$

Y_i only depends on X_i ↓

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

subadditivity ↓

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

Thm.: (Shannon's noisy coding theorem - converse part)

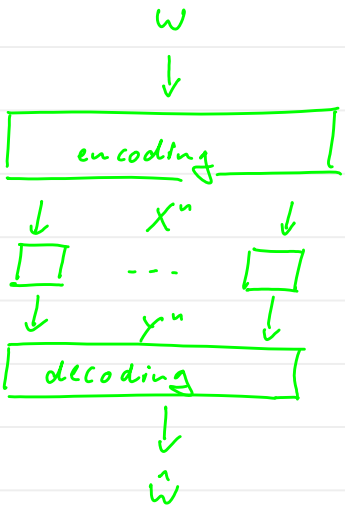
Any $(2^{nR}, n)$ code for a discrete memoryless channel satisfies

$$R \leq \frac{C}{1-p_e^{(n)}} \quad \text{where} \quad C := \max_{p(x)} I(X:Y) \quad \text{and}$$

$p_e^{(n)} := 2^{-nR} \sum_{m=1}^{2^{nR}} \lambda_m$ is the error prob. averaged over all codewords.

proof: Let W be a random variable assigned to uniform dist. of codewords.

That is, $\text{range}(W) = \{1,\dots,2^{nR}\}$



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

Fano's inequality → $\leq h(p_e^{(n)}) + p_e^{(n)} nR + I(W:\hat{W})$

data processing ineq. for Markov chain $W-X^n-Y^n-\hat{W}$ → $\leq h(p_e^{(n)}) + p_e^{(n)} nR + I(X^n:Y^n)$

previous Lemma → $\leq h(p_e^{(n)}) + p_e^{(n)} nR + \sum_{i=1}^{n} I(X_i:Y_i)$

$\leq h(p_e^{(n)}) + p_e^{(n)} nR + nC$

$\Rightarrow p_e^{(n)} \geq 1 - \frac{C}{R} - \frac{h(p_e^{(n)})}{nR}$

$\forall m \in \mathbb{N}$ we can construct a $(2^{nmR}, nm)$ code for which $p_e^{(nm)} = p_e^{(n)}$.

$$\Rightarrow p_e^{(n)} = p_e^{(nm)} \geq 1 - \frac{C}{R} - \frac{h(p_e^{(nm)})}{nmR} \xrightarrow{m \to \infty} 1 - \frac{C}{R}$$

$$\Rightarrow R \leq \frac{C}{1 - p_e^{(n)}}$$

□