2016. november 16., szerda 16:25

By what data-speed can we communicate over an unreliable channel reliably?

$$H(Y|X) = \sum_x p(x) \sum_y p(y|x) \, ld\left(\frac{1}{y(x)}\right) = \sum_x \sum_y \underbrace{p(x)p(y|x)}_{p(x,y)} ld \frac{1}{y(x)} = \sum_x \sum_y p(x,y) \, ld \frac{1}{y(x)}$$

$$H(X,Y) = H(X) + H(Y|X)$$

$$\sum_x \sum_y p(x,y) \, ld \frac{1}{p(x)} + \sum_x \sum_y p(x,y) \, ld \frac{1}{p(y|x)} = H(X) + H(Y|X)$$

Joint ??

$$I(X,Y) = D(p(x,y) \| p(x)p(y)) = \sum_x \sum_y p(x,y) \, ld \frac{p(x,y)}{p(x)\,p(y)} = \sum_x \sum_y p(x,y) \, ld \frac{p(y|x)\cancel{p(x)}}{\cancel{p(x)}\,p(y)} =$$

$$= \sum_x \sum_y p(x,y) \, ld \, p(y|x) + \sum_x \underbrace{\sum_y p(x,y)}_{p(y)} ld \frac{1}{p(y)} = H(Y) - H(Y|X) = H(X) - H(X|Y)$$

Typical set:

$$\lim_{n\to\infty} \frac{1}{n} ld \frac{1}{p(x_1 \dots x_n)} - H(X)$$

why? $\to$ $$\frac{1}{n} ld \frac{1}{p(x_1 \dots x_n)} = \frac{1}{n} ld\left(\prod_{i=1}^{n} \frac{1}{p(x_i)}\right) = \frac{1}{n} \sum_{i=1}^{n} \left(ld \frac{1}{p(x_i)}\right) = \frac{1}{n} \sum_{i=1}^{n} I(x_i) \Rightarrow$$

$$\Rightarrow \mathbb{E}(I(x)) = \sum_x p(x) \, ld \frac{1}{p(x)} - H(X)$$

$$\mathcal{A} := \{(x_1 \dots x_n) : 2^{-nH(X)-\varepsilon} \le p(x_1 \dots x_n) < 2^{-nH(X)+\varepsilon}$$

Properties:

1) $P(\mathcal{A}) \ge 1-\varepsilon$ Proof: $P(\mathcal{A}) = P\left(\left|H(X) - \frac{1}{n} ld \frac{1}{p(x_1 \dots x_n)}\right| < \varepsilon\right) \ge 1-\varepsilon$

2) $|\mathcal{A}| < 2^{nH(X)+\varepsilon}$ Proof: $1 = \sum_x p(\underline{x}) \ge \sum_{\underline{x}:\underline{x}\in\mathcal{A}} p(\underline{x}) \ge \sum_{\underline{x}:\underline{x}\in\mathcal{A}} 2^{-nH(X)-\varepsilon} = |\mathcal{A}| \cdot 2^{-nH(X)-\varepsilon} \le 1$

3) $|\mathcal{A}| \ge (1-\varepsilon) 2^{nH(X)+\varepsilon}$ Proof: $(1-\varepsilon) \le P(\mathcal{A}) = \sum_{\underline{x}:\underline{x}\in\mathcal{A}} p(\underline{x}) \le \sum_{\underline{x}:\underline{x}\in\mathcal{A}} 2^{-nH(X)+\varepsilon} = |\mathcal{A}| = 2^{-nH(X)+\varepsilon} \ge 1-\varepsilon$

$$\Rightarrow |\mathcal{A}| \ge (1-\varepsilon) 2^{nH(X)-\varepsilon}$$

Conclusion:

$|\mathcal{A}| \approx 2^{nH(X)}$

$P(\mathcal{A}) \approx 1$

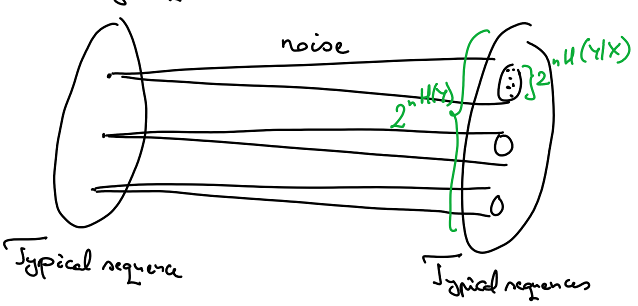
$p(x_1 \dots x_n) \approx 2^{-nH(X)}$

} source is emitting "only" typical sequence which occurs with uniform distribution

How to communicate (at what speed) reliably over an unreliable channel?

$$\text{Source} \xrightarrow{X} \boxed{\text{unreliable channel}} \xrightarrow{Y} \text{recieved info}$$

"k" long sequences Sequences of recieved symbols of length "n"



Non overlapping is the criterion for reliable communication:

$$\frac{2^{nH(Y)}}{2^{nH(Y|X)}} \ge 2^k$$

$$2^{n\{H(Y)-H(Y|X)\}} \ge 2^k$$

$$\frac{k}{n} \le \max_{p(x)} H(Y) - H(Y|X)$$

$$\frac{k}{n} \le \max_{p(x)} I(X,Y)$$

(speed) $\overbrace{}^{p(x)}$

data rate $\leq C$: channel capacity

Channel coding theorem: $\boxed{\frac{k}{n} = R \leq C} \Rightarrow C := \max_{p(x)} I(X, Y)$