Introduction to the channel coding

Part 0: Intro

From Source Coding to Channel coding

- In data compression (source coding), we try to remove “redundancy”

- In channel coding, we try to add “redundancy” in a clever way.

From Source Coding to Channel Coding

- The reasons (to add redundancy) are

① reliability ("fiabilidad") in the communication (detect and correct errors; prob. Error Pe)

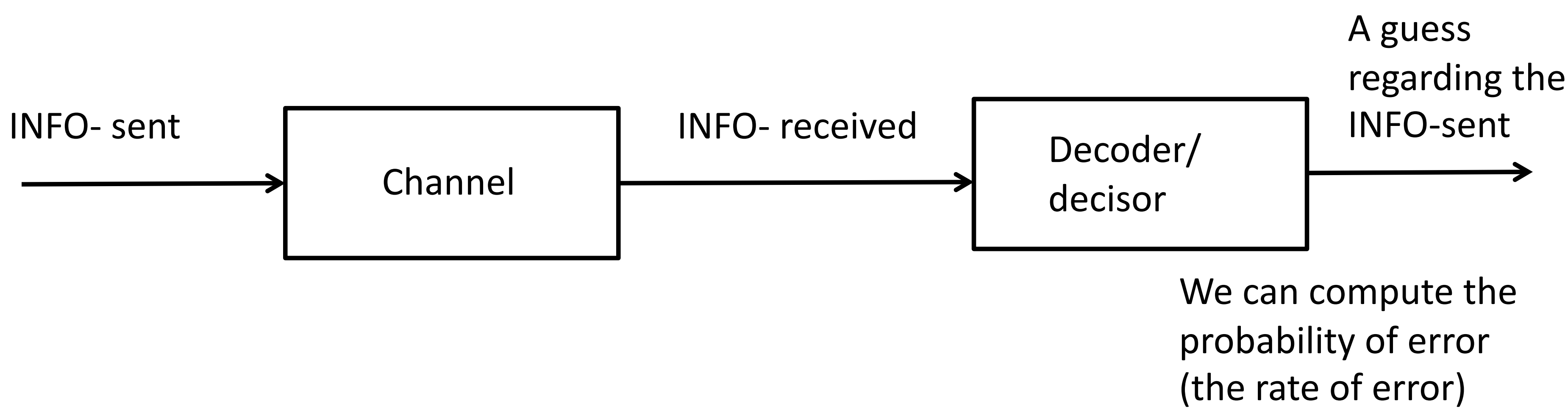
② velocity/speed in the transmission (R)

There is a trade-off **reliability-velocity**

Reliability

- The main goal when we transmit something is "reliability" ("fiabilidad")

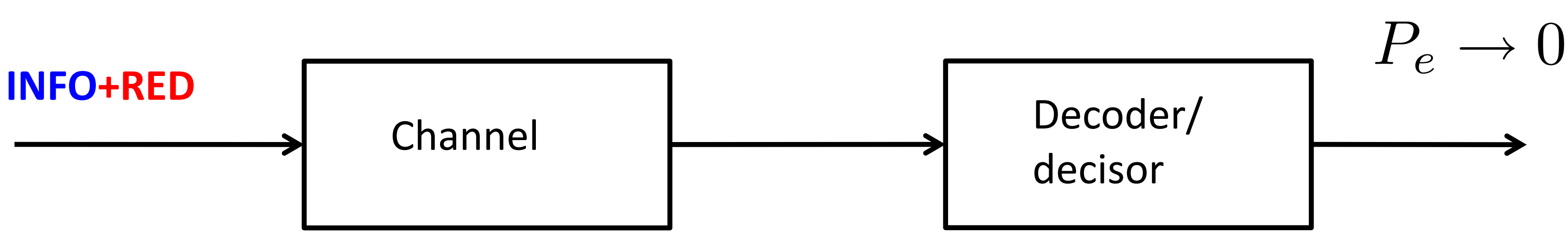
- The reliability is measured by the **Probability of error, Pe**, in detection.



Reliability

- We can always obtain a smaller Pe if we add bits of "redundancy":

$$\text{Num. of bits of Red.} \rightarrow \infty$$
$$P_e \rightarrow 0$$

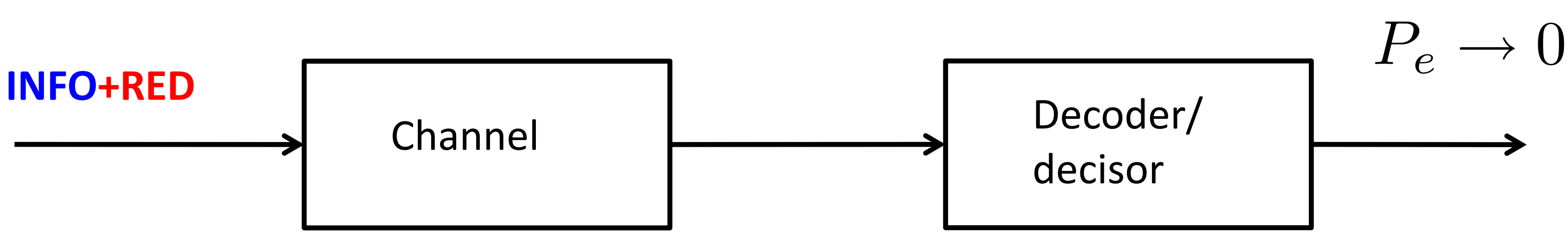


Reliability

- We will see that, in certain cases, we can also have (stronger result) if (R<C; a certain value C)

$$\text{Num of Bits of the codewords (INFO+RED)} \rightarrow \infty$$
$$P_e \rightarrow 0$$

$$\text{Keeping constant } R \text{ !!!}$$



Trade-off **reliability-velocity**

- Add more redundancy, reduce the speed/rate of transmission (of the “info”)

- With smart/clever coding algorithms, we can obtain small Pe and high speed.

- However, each channel has a characteristic maximum speed/rate…

Channel Capacity (Shannon)

- However, each channel has a characteristic maximum speed/rate, that is called **Channel Capacity (C)**,

- Capacity: “ highest information rate (i.e., speed) that can be achieved (in a channel) with arbitrarily small error probability (Pe)”

Channel Capacity

- Consider a binary channel with C=0.7 (just an example).
- Bits of info: k
- Bits of red: m
- Bits of codeword: n=k+m
- Velocity/Speed/Rate: R= k/n

$$n \to \infty \Longrightarrow P_e \to 0 \quad \text{if } R < C \text{ !!}$$
$$\text{with constant } R \text{ !!!}$$

Channel Capacity

- Shannon says: “ if R<C, we could find a code with rate R such that Pe vanishes to zero. For R>C, it is not possible.”

First part: Channels and Channel Capacity

Discrete Memoryless Channel (DMC)

- AS random variables, in formula

$$Y(t) = X(t) + E(t)$$

Streams of bits

t= time
Y(t)= received observation at time t
X(t)= trasmitted information at time t
E(t)= noise perturbation

- Likelihood (in our case, CHANNEL MATRIX)

$$p(y_t \mid x_t)$$

$$p(y \mid x)$$

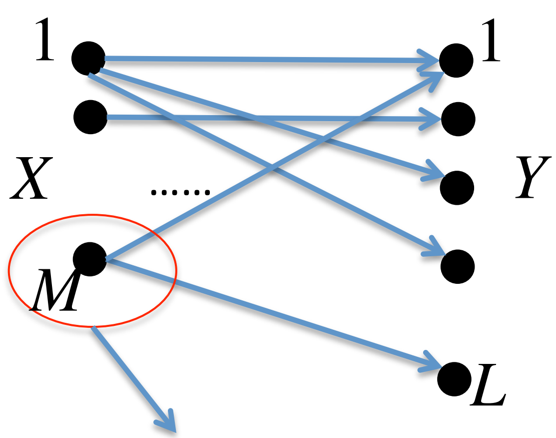
Discrete Memoryless Channel (DMC)

- $M \times L$ CHANNEL MATRIX

$$p(y|x) = \begin{bmatrix} p(y=1|x=1) & \ldots\ldots & p(y=L|x=1) \\ \ldots & \ldots\ldots & \ldots \\ \ldots & \ldots\ldots & \ldots \\ p(y=1|x=M) & \ldots\ldots & p(y=L|x=M) \end{bmatrix}$$

Rows must sum 1.

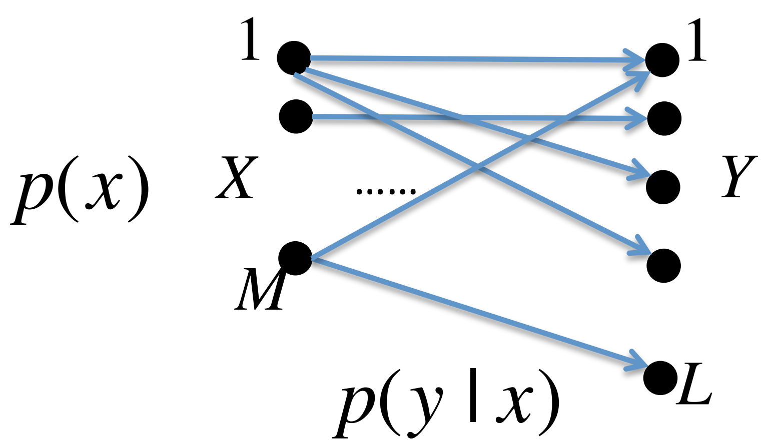
- Graphically,



All the branches have no zero probabilities. The probabilities of the branches which go out from a input node, must sum 1.

Discrete Memoryless Channel (DMC)

- We usually know also p(x)

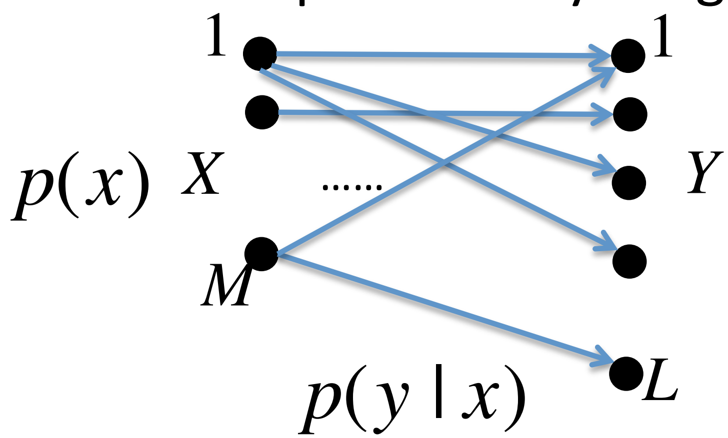


- With p(y|x), we have all the statistical information:

$$p(x,y) = p(y \mid x)\, p(x)$$

Discrete Memoryless Channel (DMC)

- We can compute "everything"



$$p(y) = \sum_{i=1}^{M} p(x=i, y) = \sum_{i=1}^{M} p(y \mid x=i)\, p(x=i)$$

$$p(x,y) = p(y \mid x)\, p(x)$$

$$p(x \mid y) = \frac{p(x,y)}{p(y)} = \frac{p(y \mid x)\, p(x)}{p(y)} = \frac{p(y \mid x)\, p(x)}{\sum_{i=1}^{M} p(y \mid x=i)\, p(x=i)}$$

- We have the 5 elements

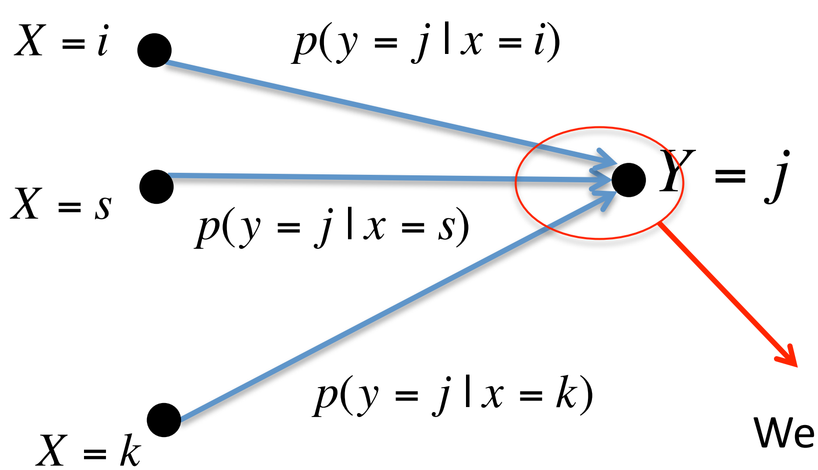
$$p(y,x) \qquad p(y \mid x) \qquad p(x \mid y) \qquad p(x) \qquad p(y)$$

Discrete Memoryless Channel (DMC)

- This formula

$$p(y=j)=\sum_{i=1}^{M}p(x=i,y=j)=\sum_{i=1}^{M}p(y=j\mid x=i)\,p(x=i)$$

- can be recall graphically



The other branches (that do not appear) have null probabilities

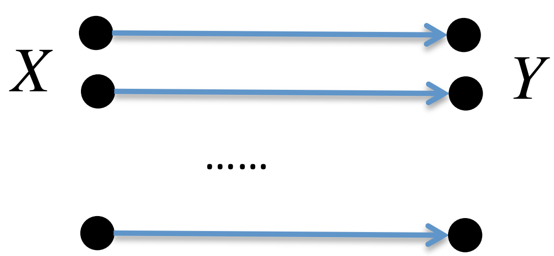
$$p(y=j\mid x)=0$$

We have to consider all branches which go in the j-th output node.

IDEAL CHANNEL

- clearly, we would like *X=Y* (ideal case).
- In this case, we obtain maxima mutual information ($I_{XY} = H_X = H_Y$).
- zero loss of information ($H_{X|Y} = 0$):

Namely, if I know *Y* I have no uncertainty with respect to *X* !!!



IDEAL CHANNEL

The worst case: X and Y are independent

- Information about Y, give me no information about X.
- In this case we have $I_{XY} = 0$ (minimum).
- Maximum of the loss of information $H_{X|Y} = H_X$
- knowing Y, the uncertainty over X does not decrease.

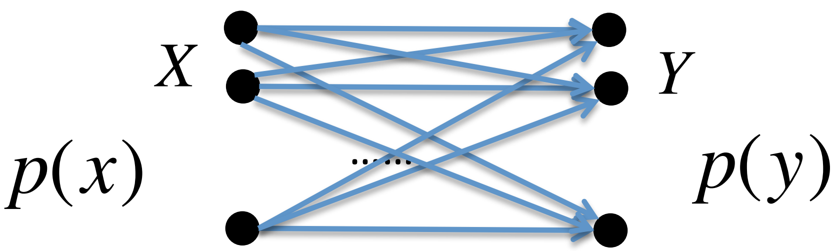
$$p(x,y) = p(x)p(y)$$

$$p(y|x) = \frac{p(x,y)}{p(x)} = \frac{p(y)p(x)}{p(x)} = p(y)$$

The worst case: X and Y are independent

- When does it happen?

$$p(y \mid x) = \frac{1}{\mathrm{L}} = \frac{1}{\text{num. of branches going out from one input (and num. of outputs)}}$$



The worst possible channel

All the inputs have the same number of arrows with the same probabilities (B=3, the probabilities are 1/3, 1/3,1/3).

$$p(x,y) = p(y \mid x)p(x) = \frac{1}{L}p(x),$$

$$p(y) = \sum_i p(x_i,y) = \frac{1}{L}\sum_i p(x_i) = \frac{1}{L} \Rightarrow p(y) = p(y \mid x)$$

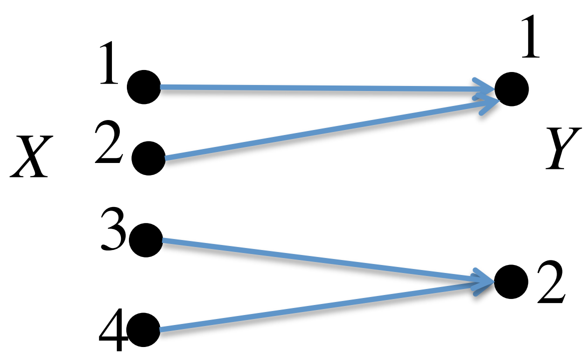
$$p(x,y) = p(y)p(x)!!!$$

Other interesting cases

- When Y=X we have

$$I_{XY} = H_X = H_Y \qquad H_{X|Y} = 0 \quad H_{Y|X} = 0$$ Both!!

- One conditional entropy can be zero and the other one no, even when $Y \neq X$. For instance,



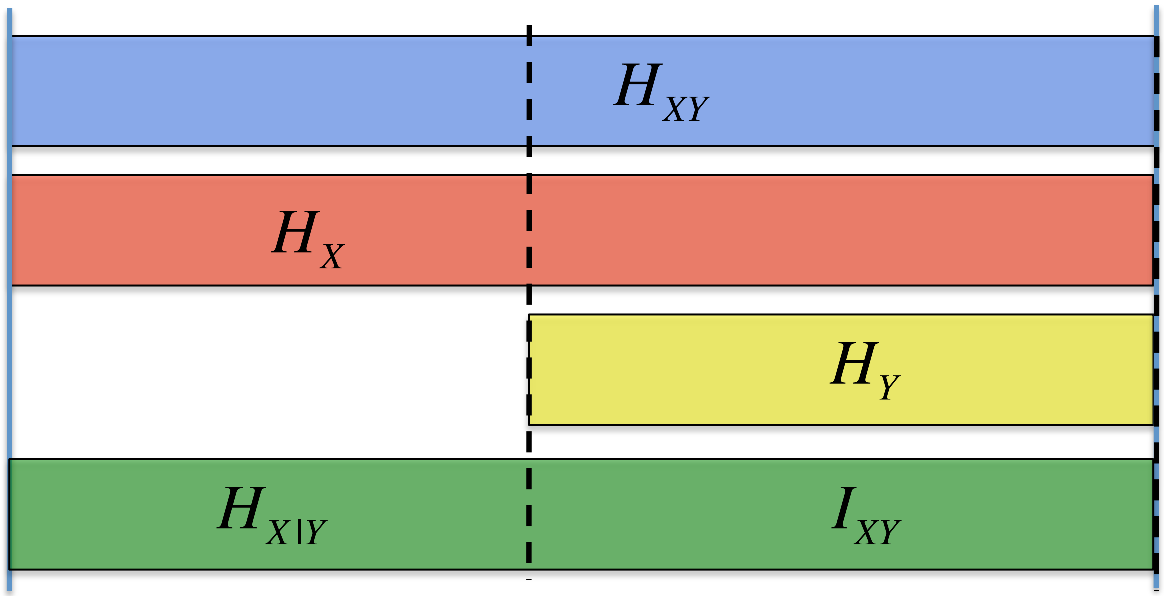
$$H_{Y|X} = 0$$

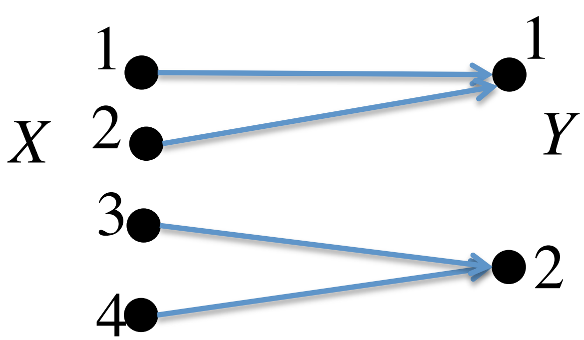
Uncertainty about Y if I know X? ZERO.

$$H_{X|Y} \neq 0$$

Other interesting cases

- The corresponding plot is:





$H_{Y|X} = 0$

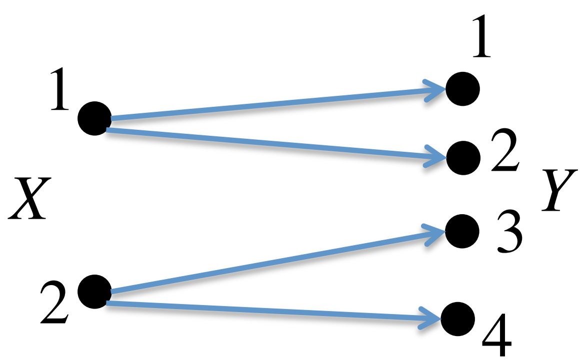
$H_{XY} = H_X$

$H_{X|Y} \neq 0$

$H_Y = I_{XY}$

Other interesting cases

- The symmetric case is



Info lost in the channel

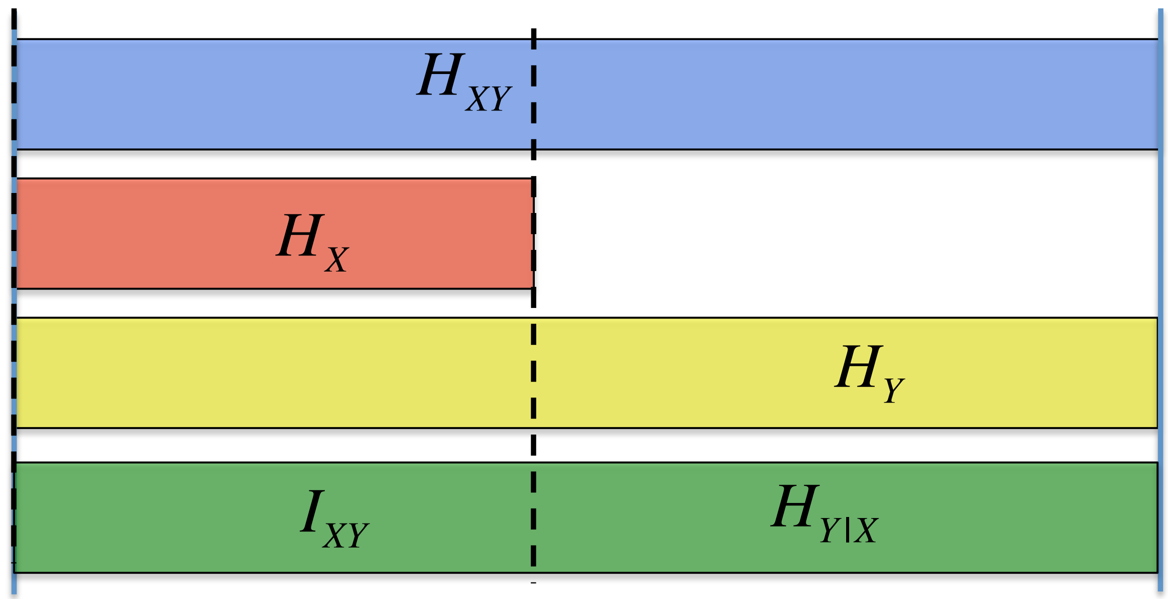
$H_{X|Y} = 0$ → Uncertainty about X if I know Y? ZERO.

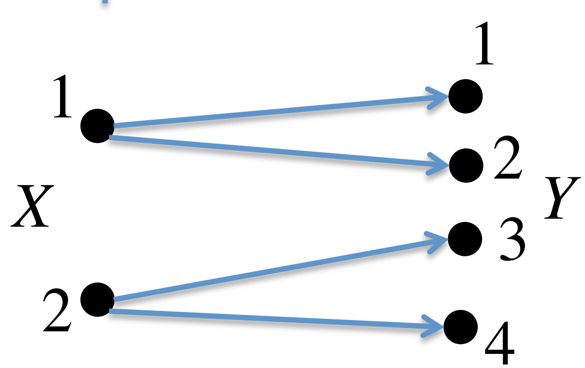
$H_{Y|X} \neq 0$

- Please note that $Y \neq X$, however the channel can be considered "ideal" since the info lost is zero!

Other interesting cases

- The corresponding plot is:





$H_{X|Y} = 0$

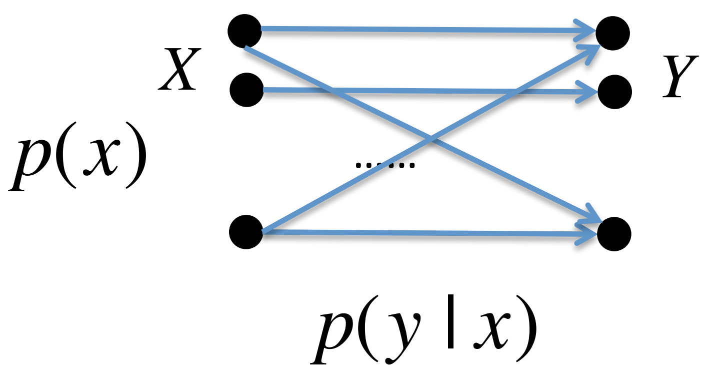
$H_{XY} = H_Y$

$H_{Y|X} \neq 0$

$I_{XY} = H_X$

Maximize I_{XY}

- We want to maximize the mutual information.
- The channel is given, we can only change p(x).
- **We look for the $p(x)$ in order to maximize I_{XY} .**



The channel matrix is given.

Channel Capacity

- In general, we cannot obtain $I_{XY} = H_X$, even trying to maximizing the mutual information.

- **The CHANNEL CAPACITY is a "feature" of the channel:**

$$C = \max_{p(x)} I_{XY}$$

Channel Capacity

- In order to obtain the capacity ,we can work with two expressions of the mutual information:

$$I_{XY} = H_X - H_{X|Y}$$

$H_X \downarrow$ $p(x)$

$H_{X|Y} \downarrow$

$$p(x|y) = \frac{p(x,y)}{p(y)} = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\sum_{i=1}^{M} p(y|x=i)p(x=i)}$$

$$p(x,y) = p(y|x)p(x)$$

$$I_{XY} = H_Y - H_{Y|X}$$

EASIER this one!

$H_Y \downarrow$

$$p(y) = \sum_{i=1}^{M} p(x=i,y) = \sum_{i=1}^{M} p(y|x=i)p(x=i)$$

$H_{Y|X} \rightarrow$ $p(y|x)$

Channel Capacity

- In the "exercises", it is easy to use this one

$$I_{XY} = H_Y - H_{Y|X}$$

- But theoretically, the next one is more interesting:

$$I_{XY} = H_X - H_{X|Y}$$

INFO that PASSES through the channel (output INFO)

INPUT INFO

LOSS OF INFO through the channel

Channel Capacity

- VERY IMPORTANT OBSERVATION

2^c = number of inputs that I can use SIMULTANEOUSLY without having errors in detection

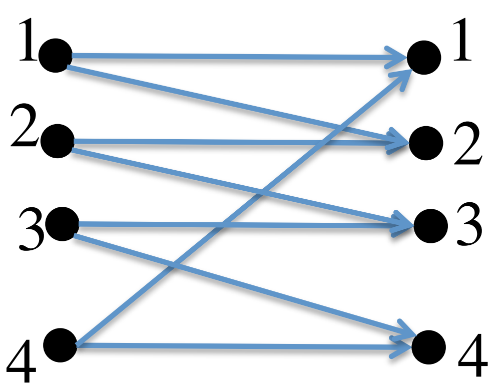
- The value $2^{capacity}$ can be interpreted as the number of entries (inputs/symbols) that can be used (simultaneously) without having error during the transmission.

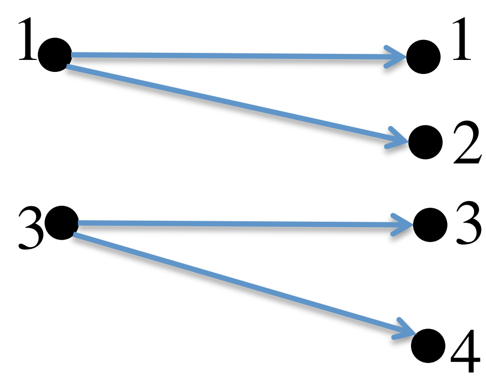
Channel Capacity

- Example:

For sure, I can use two entries without doing errors in detection. I can assert this sentence without knowing the values p(y|x)

Seguramente (cualquiera sea la matriz de canal) puedo utilizar 2 entradas sin equivocarme.





- Therefore I can write:

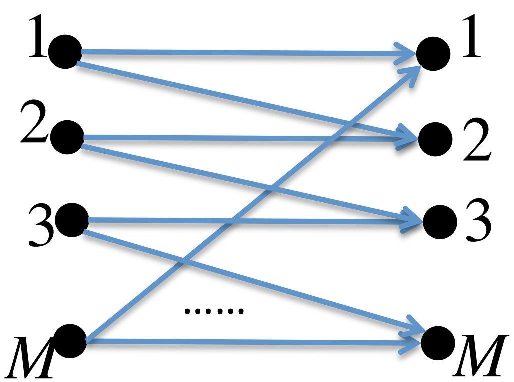
$$2^c \geq 2 \Rightarrow C \geq 1$$

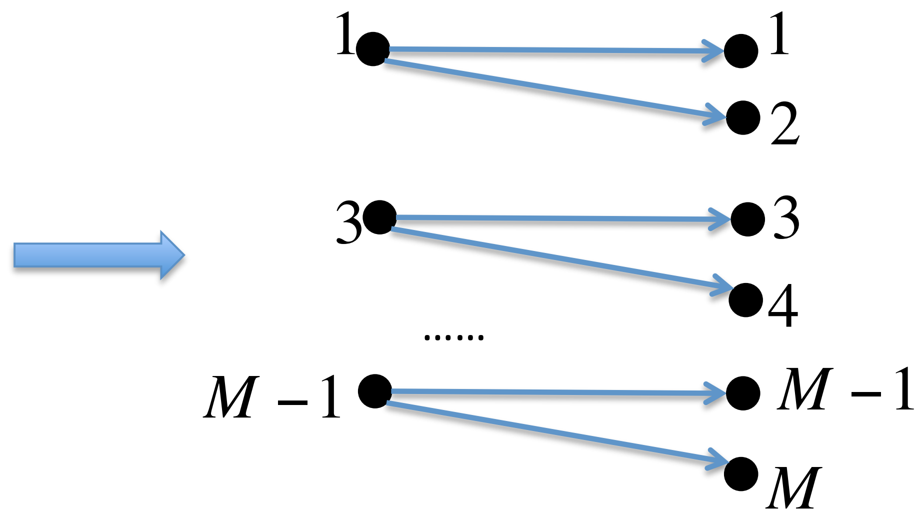
C=1 is the worst case.

C=1 es en el caso peor.

Channel Capacity

- When M is even:





I can use M/2 entries simultaneously, without errors in detection.

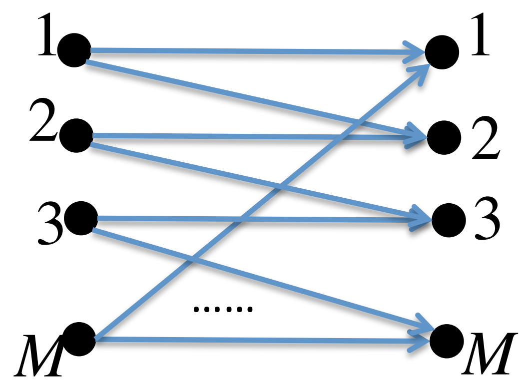
- For sure, I can write:

$$2^c \geq \frac{M}{2} \Rightarrow C \geq \log_2 \frac{M}{2} = \log_2 M - 1$$

Channel Capacity

- We found a lower bound for the capacity of this channel:

$$C \geq \log_2 \frac{M}{2}$$



- We can also obtain an upper bound. Indeed, in the ideal case, AT MOST we can use all the entries:

$$2^C \leq M$$

$$C \leq \log_2 M$$

THIS FORMULA IS ALWAYS TRUE!
NO JUST FOR THIS CHANNEL.

$$\log_2 \frac{M}{2} \leq C \leq \log_2 M$$

$$C \leq \log_2 L$$

In this case, L=M

also this inequality is valid

Upper Bound for Channel Capacity

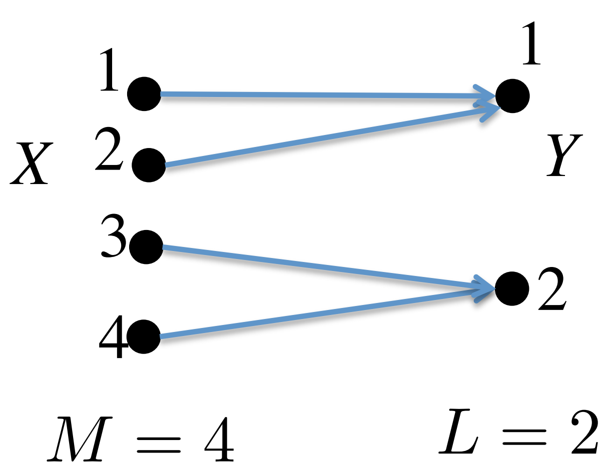
Always we have

$$C \leq \min(\log_2 M, \log_2 L)$$

You can see by formulas, or thinking on the number of entries that you can use simultaneously without having error in detection

$$I_{XY} = H_Y - H_{Y|X}$$

$$I_{XY} = H_X - H_{X|Y}$$



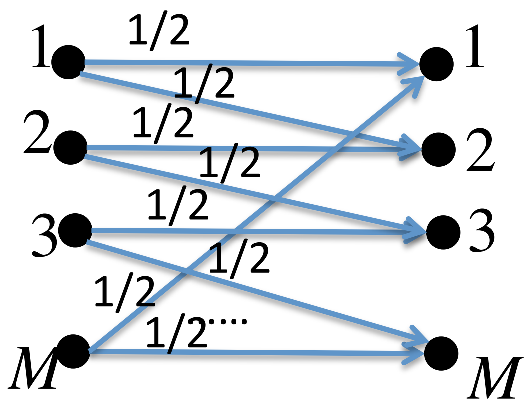
$$H_{Y|X} = 0$$

Uncertainty about Y if I know X? ZERO.

$$H_{X|Y} \neq 0$$

Channel Capacity

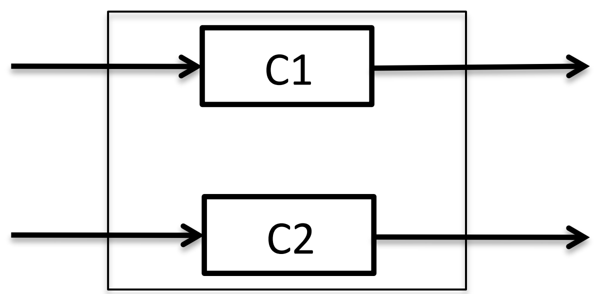
- In the special case of



- We can prove $C = \log_2 \frac{M}{2}$

Capacity of parallel channels

- Two channels in parallel:



$$2^{c_{tot}} = 2^{c_1} + 2^{c_2}$$

$$c_{tot} = \log_2\left(2^{c_1} + 2^{c_2}\right)$$