University of Diyala

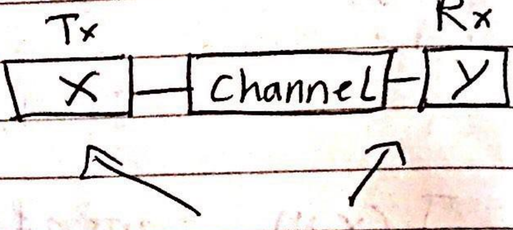
Information Theory

Lecturer Marwa Mohammed

Communication department / Engineering collage

* Murginal Entropies:

A term usually used to denote both source entropy $H(X)$ as defined before, and the receiver entropy $H(Y)$



$$H(Y) = -\sum_{j=1}^{m} p(y_j) \log_2 p(y_j)$$

margins of the channel

* Joint and conditional entropies:

The avarage amount of information associated with the pair (x_i, y_j) is called joint entropy $H(X,Y)$.

$$H(X,Y) \equiv H(XY) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(x_i, y_j)$$

bits/symbol

The average amount of information associeted with the pairs (y_j/x_i) & (x_i/y_j) are called Conditional entropies $H(Y/X)$ and $H(X/Y)$ they are given by:

$$H(Y/X) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(y_j/x_i) \quad \text{bits/symbol}$$

noise entropy

and

$$H(X/Y) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(x_i/y_j) \quad \text{bits/symbol}$$

losses entropy

Ex: Show that $H(X,Y) = H(X) + H(Y/X)$

Sol: we know that

$$H(X,Y) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(x_i; y_j) \quad — ①$$

$$p(x_i, y_j) = p(x_i)\, p(y_j/x_i) \quad — ②$$

put ② inside the log term only

$$H(X,Y) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(x_i) \cdot p(y_j / x_i)$$

$$= -\left[\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(x_i) + \sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log_2 p(y_j / x_i)\right]$$

$$= -\sum\sum p(x_i, y_j) \log_2 p(x_i) - \sum\sum p(x_i, y_j) \log_2 p(y_j / x_i)$$

if we reverse the first sum for i and j then

$$\sum_j p(x_i, y_j) = p(x_i)$$

then

$$H(X,Y) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) - \sum_j\sum_i p(x_i, y_j) \log_2 p(y_j / x_i)$$

$$\Rightarrow H(X,Y) = H(X) + H(Y/X)$$

H.w: show that $H(X,Y) = H(Y) + H(X/Y)$

Ex: Show that $I(X,Y) = H(X) - H(X/Y)$

Sol: We know that

$$I(X,Y) = \sum_j\sum_i p(x_i, y_j) \log_2 \frac{p(x_i / y_j)}{p(x_i)}$$

$$= \sum_j\sum_i p(x_i, y_j) \log_2 p(x_i / y_j) - \sum\sum p(x_i, y_j) \log_2 p(x_i)$$

now: reverse the order of the 2nd sum, then

$$I(X,Y) = H(X) - H(X/Y)$$

Note: Above identity indicates that the transinformation $I(X,Y)$ is the net average information gained at the Rx which is difference between the originel average of information produced by the source $H(X)$ and that information lost in the channel $H(X/Y)$ (losses entropy) due to noise and jamming.

H.W: show that $I(X,Y) = H(Y) - H(Y/X)$.

Ex: The joint Probability is given by

$$P(X,Y) = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} & \begin{bmatrix} 0.5 & 0.25 \\ 0 & 0.125 \\ 0.0625 & 0.0625 \end{bmatrix} \end{array}$$

find ① marginal entropies, ② joint (system) entropy
③ noise & losses entropy, ④ mutual information between x_1 & y_2
⑤ Transinformation, ⑥ Draw the channel model.

Sol: firstly we find $p(x), P(y)$ from $P(x,y)$

$$p(X) = \begin{array}{ccc} x_1 & x_2 & x_3 \\ \end{array}\left[\begin{array}{ccc} 0.75 & 0.125 & 0.125 \end{array}\right] ; \; p(x_i) = \sum_{j=1}^{2} p(x_i, y_j)$$

$$p(y) = \begin{array}{cc} y_1 & y_2 \end{array}\left[\begin{array}{cc} 0.5625 & 0.4375 \end{array}\right] ; \; P(y_j) = \sum_{i=1}^{3} P(x_i, y_j)$$

① $H(x) = -\sum_{i=1}^{3} p(x_i) \log_2 p(x_i)$

$= \frac{-1}{\ln 2} \left[0.75 \ln 0.75 + 2 * 0.125 \ln 0.125 \right]$

$= 1.06127$ bits/symbol.

$H(y) = -\sum_{j=1}^{2} p(y_j) \log_2 P(y_j)$

$= \frac{-1}{\ln} \left[0.5625 \ln 0.5625 + 0.4375 \ln 0.4375 \right]$

$= 0.9887$ bits/symbol

② joint entropy:

$H(X,Y) = -\sum_{j=1}^{3}\sum_{i=1}^{2} p(x_i, y_j) \log_2 P(x_i, y_j)$

$= -\left[0.5 \log_2 \frac{1}{2} + 0.25 \log_2 0.25 + 0.125 \log_2 0.125 + \right.$

$\left. 2 * 0.0625 \log_2 0.0625 \right]$

$= \left[0.5 + 0.25 * 2 + 0.125 * 3 + 0.125 * 4 \right]$

$= 1.875$ bits/symbol.

③ noise & losses entropies

$H(Y/X) = H(X,Y) - H(X)$

$= 1.875 - 1.06127 = 0.81373$ bits/symbol

$H(X/Y) = H(X,Y) - H(Y)$

$= 1.875 - 0.9887 = 0.8863$ bits/symbol

④ Mutual information between x_1 & y_2

$I(x_1, y_2) = \log_2 \frac{P(X_1/y_2)}{P(X_1)}$

but $p(X_1/y_2) = \frac{p(X_1, y_2)}{P(y_2)}$

$\Longrightarrow I(X_1, y_2) = \log_2 \; p(X_1, y_2) / p(X_1) \cdot P(y_2)$

$= \log_2 \frac{0.25}{0.75 * 0.4375} = -0.3923$ bits

That means y_2 gives ambiguity about x_i

⑤ Transinformation:

$I(X,Y) = H(X) - H(X/Y)$

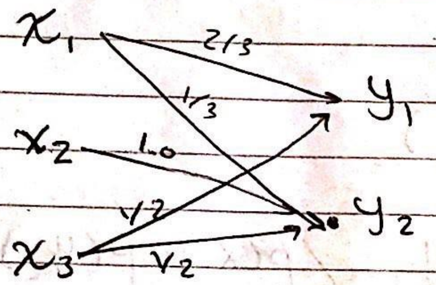
$= 1.06127 - 0.8863 = 0.17497$ bits/symbol

⑥ To draw channel model, we find $p(Y/X)$ matrix from $p(X,Y)$ matrix.

$$P(Y/X) = \frac{P(X,Y)}{P(X)} = \begin{bmatrix} 0.5/0.75 & 0.25/0.75 \\ 0/0.125 & 0.125/0.125 \\ 0.0625/0.125 & 0.0625/0.125 \end{bmatrix}$$

$$= \begin{bmatrix} 2/3 & 1/3 \\ 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

Note: sum of each row must be unity.



Note: sum of prob. branching out of each x_i is unity.

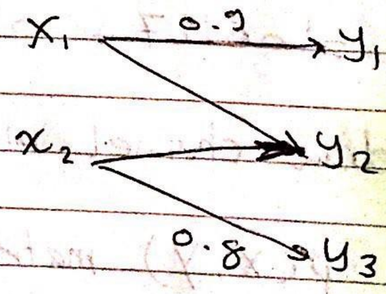
<u>H.w:</u> For the channel model shown, find

a. Source entropy rate if

$\tau_{x_1} = 1\,ms$, $\tau_{x_2} = 2\,ms$

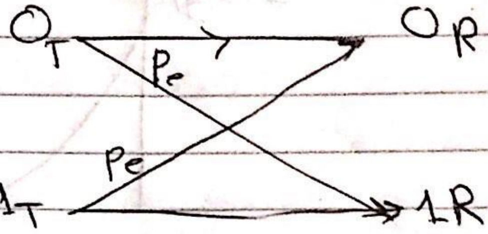
$I(x_1) = 2$ bits

b. Transinformation



Ex:- find and plot the transinformation for a binary symmetric channel (BSC) shown, if $P(0_T) = P(1_T) = \frac{1}{2}$.

<u>Sol:</u>



$0_T = x_1, \quad 0_R = y_1$

$1_T = x_2, \quad 1R = y_2$

$$P(Y/X) = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} 1-P_e & P_e \\ P_e & 1-P_e \end{bmatrix} \end{array}$$

$$P(X,Y) = P(Y/X) \cdot P(X)$$

$$P(X,Y) = \begin{bmatrix} \frac{1-P_e}{2} & \frac{P_e}{2} \\ \frac{P_e}{2} & \frac{1-P_e}{2} \end{bmatrix}$$

$$P(Y) = \begin{bmatrix} \underset{y_1}{1/2} & \underset{y_2}{1/2} \end{bmatrix}$$

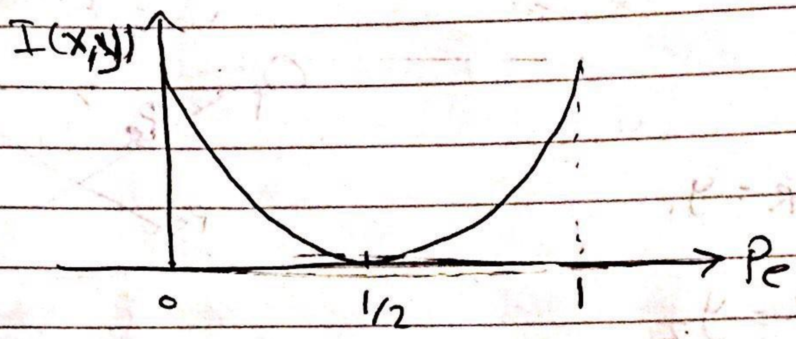
$$H(Y) = H(Y)\big|_{max} = 1 \text{ bit} = \log_2 2$$

$$H(Y/X) = -\left[\left[\left(\frac{1-P_e}{2}\right)\log_2 (1-P_e)\right] * 2 + \left[\frac{P_e}{2}\log_2 P_e\right] * 2\right]$$

$$= -\left[(1-P_e)\log_2 (1-P_e) + P_e \log_2 P_e\right]$$

$$I(X,Y) = H(Y) - H(Y/X)$$

$$= 1 + \left[(1-P_e)\log_2(1-P_e) + P_e \log_2 P_e\right]$$



<u>H.W:</u> A BSC has

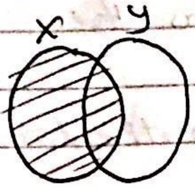
$$P(Y/X) = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

if $I(0_T) = 3$ bits

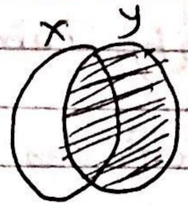
find system and losses entropies.

* Venn Diagram Representation of entropies :-

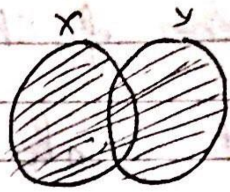
x: source y: receiver



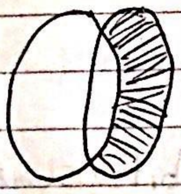
$H(X)$



$H(Y)$



$H(X,Y)$

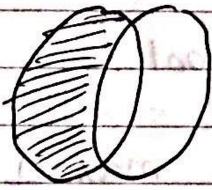


$H(Y/X)$

noise entropy

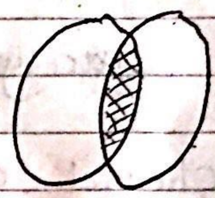
↑

$H(X,Y) - H(X)$



$H(X/Y)$

losses entropy



$I(X,Y)$

Transinformation