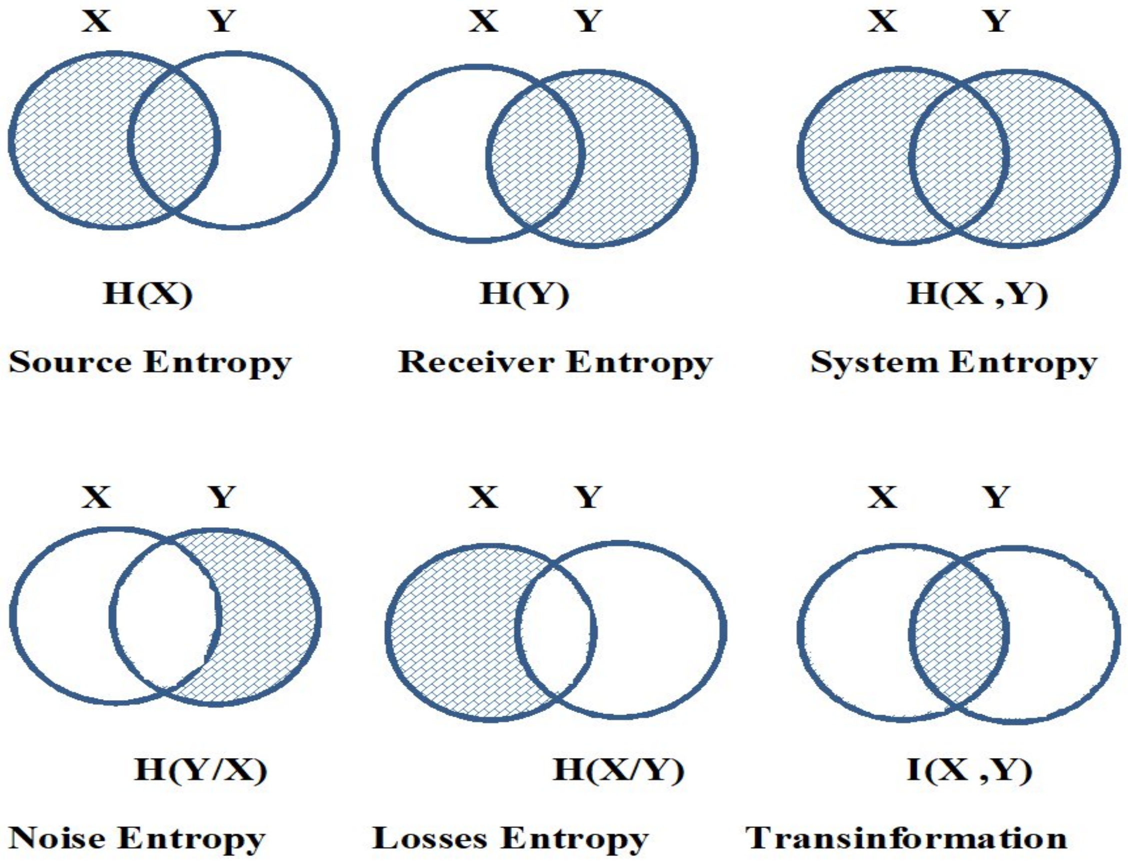
Definitions:

Source efficiency = $\boldsymbol{\eta_s}$ = $\boldsymbol{H(X)/H(X)|_{max}}$= $\mathbf{H(X)}/\log_2 \boldsymbol{n}$

Source redundancy=$\mathbf{1-}$ $\boldsymbol{\eta_s}$ $=1-$ [$\boldsymbol{H(X)}/\log_2 \boldsymbol{n}$]

Note above can also be given as percentage.

Venn Diagram representations of entropies:



X=Transmitter Tx : Y= Receiver Rx

Discrete channel is a channel whose input X is discrete. The channel capacity is defined as the maximum of **I(X,Y):**

C = channel capacity = max [I(X,Y)] *bits/symbol.*

Physically it is the maximum amount of information each symbol can carry to the receiver. Sometimes this capacity is also expressed in bits/sec if related to the rate of producing symbols ***r***:

C = r × max[I(X,Y)] bits/sec,

where ***r*** is the number of symbols produced per second.

C is also expressed as: **C=max[R(X,Y)],** where: R(X,Y)=rate of information transmission :

R(X,Y)= ***r*** × I(X,Y) bits/sec or

$$R(X,Y) = \frac{I(X,Y)}{\overline{\tau}},$$

$$\overline{\tau} = \sum_{i=1}^{n} \tau_i P(x_i)$$

The maximization of I(X,Y) is done with respect to input prob P(X) or output prob P(Y) for a constant channel conditions, i.e. with p(Y/X) being a constant.

Definition:- A more general definition of symmetric channel is that channel where:

1- ***n* = *m*** , equal number of symbols in X & Y, i.e. P(Y/X) is a square matrix.

2- Any row in P(Y/X) matrix comes from some permutation of other rows.

<u>Examples:</u>

1- $p(Y/X)=\begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$ is a BSC where n=m=2 and the 1st row is the permutation of the 2nd row.

2- $p(Y/X)=\begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.05 & 0.9 & 0.05 \\ 0.05 & 0.05 & 0.9 \end{bmatrix}$ is TSC where n=m=3 and each row is a permutation of the others.(same numbers appear)

3- $p(Y/X)=\begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \end{bmatrix}$ is a nonsymmetric since n≠m(not square) although the 1st row is permutation of the 2nd row.

4- $p(Y/X)=\begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$ is a nonsymmetric since the 2nd row is not a permutation of some other row, although n=m=3.

Channel capacity of Discrete Symmetric Channels:

Channel capacity of such symmetric channel is easy to find using the following derivation: To find **max**[I(X,Y)], then:

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

$$I(X,Y)=H(Y)+\sum_{j=1}^{m}\sum_{i=1}^{n}P(xi\ ,\ yj)\ log_2 P(yj/xi)$$

$$I(X,Y)=H(Y)+\sum_{j=1}^{m}\sum_{i=1}^{n}P(xi)P(yj\ /xi\ )log_2 P(yj/xi)$$

If the channel is symmetric then above expression for I(X,Y) is reduced to H(Y) + $\sum_{j=1}^{m}$ P(yj /xi) log_2P(yj/xi) (***n*=*m***) since $\sum_{i=1}^{n} P(x_i)$ which is equal to unity, and the 2nd sum over j is independent of i (independent of the row number i because the channel is symmetric and each row is the permutation of the other rows). Hence the quantity $\sum_{j=1}^{m}$ P(yj /xi) log_2P(yj/xi) is a constant = ***K*** (***K*** is –Ve and calculated for one row only since it is the same for any row***). Hence for symmetric channels only*** :

I(X,Y)=H(Y) – H(Y/X)= H(Y) + $\sum_{j=1}^{m}$ P(yj /xi) log_2P(yj/xi) $= \boldsymbol{H(Y)+K}$

And since **max**[H(Y)]=$\log_2 m$ when **Y** has equiprobable symbols, then:

$$\boxed{C = \log_2 m + K} \quad \text{bits/symbol}$$

$$K=\sum_{j=1}^{m}P(yj\ /xi)\ log_2 P(yj/xi)\ \textit{for one row only}$$

Channel Efficiency and Redundancy

Channel efficiency: $\eta c = \frac{I(X,Y)}{C}$

Channel redundancy $R = 1 - \eta c = 1 - \frac{I(X,Y)}{C}$

Notes:

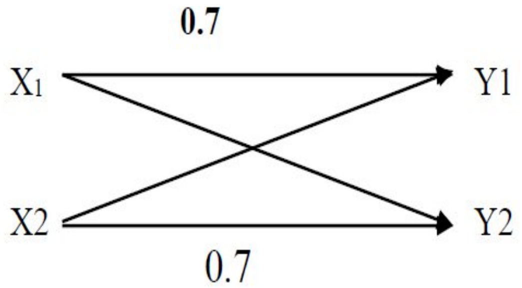
1-I(X,Y) becomes maximum equals **C** only if the condition for maximization is satisfied, i.e. only if Y has equiprobable symbols.

This condition yields that X has also equiprobable symbols since if the output of a symmetric channel is equiprobable, then its input X is also symmetric.

2-For symmetric channel only, and to ease calculations, we can use the formula

I(X,Y)=H(Y)+K.

Example: For the BSC shown:
Find the channel capacity and efficiency if $I(x_1)$=2bits



Solution: First we write p(Y/X), as: $p(y/x)=\begin{bmatrix}0.7 & 0.3\\ 0.3 & 0.7\end{bmatrix}$ and since channel is symmetric, then $C=\log_2 \boldsymbol{m}+\boldsymbol{K}$, (n=m=2) and
$\boldsymbol{K}= 0.7\log_2 0.7 + 0.3\log_2 0.3 = -0.88129$ then
$C=\log_2 2 + K=1-0.88129=0.1187$ bits/symbol.

To find the channel efficiency, then we must find I(X,Y). First, we find $p(x_1)$ from $I(x_1)= -\log_2 p(x_1)=2$, giving $p(x_1)=2^{-2}=0.25$, then : p(X)=[0.25 0.75], multiplying with p(Y/X) get:

$$p(x\ ,\ y)=p(x)\ .\ p(y/x)=\begin{bmatrix}0.7\times 0.25 & 0.3\times 0.25\\ 0.3\times 0.75 & 0.7\times 0.75\end{bmatrix}=\begin{bmatrix}0.175 & 0.075\\ 0.225 & 0.525\end{bmatrix}$$

then summing the columns to give p(Y)=[0.4 0.6], from which H(Y)=0.97095 bits/symbol.
Then: I(X,Y)=H(Y)+K=0.97095-0.88129=0.0896 bits/symbol.

Then: ηc =I(X,Y)/C=0.0896/0.1187=75.5%.

HW: repeat previous example for the channel having transition prob. If $P(x_1)= P(x_2)=0.25$

$$p(Y/X)=\begin{bmatrix}0.7 & 0.2 & 0.1\\ 0.1 & 0.7 & 0.2\\ 0.2 & 0.1 & 0.7\end{bmatrix}$$

Procedure:

1- First, we find I(X,Y) as a function of input prob:

$I(X,Y)=\boldsymbol{f}\,(p(x_1), p(x_2), \ldots .p(x_n))$. subject to the constraint: $\sum P(x_i) = 1$. i.e. use this constraint to reduce the number of variables by 1.

2- Partial differentiate I(X,Y) with respect to the (***n-1***) input prob., then equate these partial derivatives to zero.

3-Solve the (n-1) equations simultaneously then find $p(x1),p(x2),\ldots\ldots,p(x_n)$ that gives maximum I(X,Y).

Note that the condition here is not necessarily equiprobable since the channel is not symmetric.

4-put resulted values of input prob. in the function ***f*** given in step 1 above to find **C= max[I(X,Y)].**

Example : Find the channel capacity for the channel having transitional prob.

$$p(Y/X)=\begin{bmatrix} 0.7 & 0.3 \\ 0.1 & 0.9 \end{bmatrix}$$

Solution :

Note that the channel is not symmetric since the 1st row is not a permutation of the 2nd row.

Now let $p(x_1)=\mathbf{p}$, then $p(x_2)=1-\mathbf{p}$, hence instead of having two variables, we will have only one variable **p**.

Next we find p(X,Y) by multiplying the rows of p(Y/X) by p(X):

$$p(X,Y)=\begin{bmatrix} 0.7p & 0.3p \\ 0.1(1-p) & 0.9(1-p) \end{bmatrix}$$

then $P(Y)=[0.1+0.6\boldsymbol{p} \qquad 0.9-0.6\boldsymbol{p}]$ and:

H(Y/X) $= -\sum_{j=1}^{m}\sum_{i=1}^{n}$ P(xi , yj) log_2P(yj/xi)

H(Y/X)= – [0.7***p*** ln 0.7 + 0.3***p*** ln 0.3 + 0.1(1-***p***) ln 0.1 + 0.9(1-***p***) ln 0.9]/ln(2)

$$\frac{\partial H(Y/X)}{dp} = \frac{d\ H(Y/X)}{dp} = - \ [\ 0.7\ln 0.7\ +\ 0.3\ln 0.3 - 0.1 \ln 0.1 - 0.9 \ln 0.9]/\ln(2) = -[-0.285781]\ /\ \ln(2)$$

Also: H(Y)= –[(0.1+0.6p) ln(0.1+0.6p) + (0.9-0.6p) ln(0.9-0.6p)]/ln2 Then:

$$\frac{\partial H(Y)}{dp} = -\frac{1}{\ln\ (2)}\ [(0.6ln(0.1+0.6p)+0.6)+(-0.6ln(0.9-0.6p)-0.6)] = -\frac{1}{\ln\ (2)}[0.6ln\frac{0.1+0.6p}{0.9-0.6p}]$$

$$\frac{\partial I(X,Y)}{dp} = \frac{\partial H(Y/X)}{dp} + \frac{\partial H(Y)}{dp} = 0 \ \Rightarrow\ 0.6ln\frac{0.1+0.6p}{0.9-0.6p} + 0.285781 = 0$$

Solving for p=0.47187=p(x_1), putting into H(Y) equation to get H(Y)=0.96021 bits/symbol,
and in H(Y/X) equation to get H(Y/X)=0.66354 bits/symbol. And finally, **C=max[I(X,Y)]**=0.96021-0.66354=0.29666 bits/symbol.

<u>Notes</u>

1-Previous example, shows that finding C for nonsymmetric channel is not so easy mathematically, specially if the number of symbols is greater, then we have to partial differentiate I(X,Y) to get a set of (n-1) nonlinear equations whose solution is not always easy.

2- Sometimes, we are asked to find C when the channel is not symmetric, but there are some similarities between some symbols(not all). In such a case, we can safely assume such similar symbols equiprobable, and then proceed as before. This note helps to more reduce number of variables.

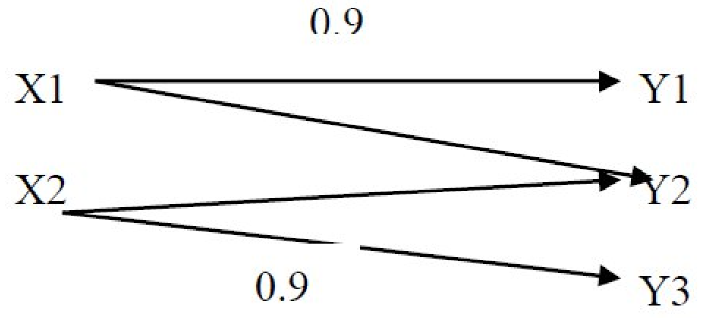
Notes: nonsymmetric channel

For example, the very practical ternary channel mentioned before having the transition prob:

$$p(Y/X)=\begin{array}{c} \\ x1 \\ x2 \\ x3 \end{array}\left[\begin{array}{ccc} y1 & y2 & y3 \\ 0.9 & 0.1 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0 & 0.1 & 0.9 \end{array}\right]$$

we note that x_1 acts very similar to x_3 in the channel, so that we safely assume $p(x_1)=p(x_3)=p$, so that $p(x_2)=1-2p$, then proceed as in previous example to find I(X,Y) as a function of only one variable p.

Also, take another example shown below, where the channel is not symmetric



since

$$p(Y/X)=\begin{array}{c} \\ x1 \\ x2 \end{array}\left[\begin{array}{ccc} y1 & y2 & y3 \\ 0.9 & 0.1 & 0 \\ 0 & 0.1 & 0.9 \end{array}\right]$$

is not a square matrix, but x_1 is similar to x_2, then we can assume that $p(x_1)=p(x_2)=0.5$, then use it to find I(X,Y) representing C. In such a case, no need for differentiation but be careful not to mix this special example with the symmetric case and the use of the formula $C=\log_2 m+K$ is not correct and gives wrong answer.