6th SEM ECE ITC I-II-UNIT

Ms. SHRUTI AWASTHI

40/-

Random Variables

Variables that take on values determined by probability distributions. They may be discrete or continuous, in either domain or their range.

Ex: Analog speech, a stream of ASCII encoded text characters in a transmitted message.

Probability & Baye's Theorem

$P(A, B)$ = Joint probability of both A and B

$= P(A|B)\, P(B)$

or $P(B|A)\, P(A)$

In case A and B are independent events

$P(A|B) = P(A)$

& $P(B|A) = P(B)$

$P(A, B) = P(A)\, P(B)$

Bayes Theorem

$$P(B|A) = \frac{P(A|B)\, P(B)}{P(A)}$$

Information Theory deals with

1) Measure of source information
2) Information capacity of the channel.
3) Coding

If the rate of information from a source does not exceed the capacity of the channel, then there exist a coding scheme such that ~~the~~ comm. information can be transmitted over with arbitary small amount of errors despite the presence of Noise.

Information Measure

→ used to determine the "information rate" of discrete sources

Consider 2 messages

A Dog bites a man → High Probability → Less Info

A man bites a dog → Less " → More "

∴ Information $\propto \dfrac{1}{(\text{Prob. of occurrence})}$

Rule 1 :- I approaches to 0 as P_k approaches ∞.

eg: Sun rises in East

$I = 0$ as $P_k \to 1$

Rule 2 :- Information content I must be non-negative. It may be zero

eg: Sun rises in west.

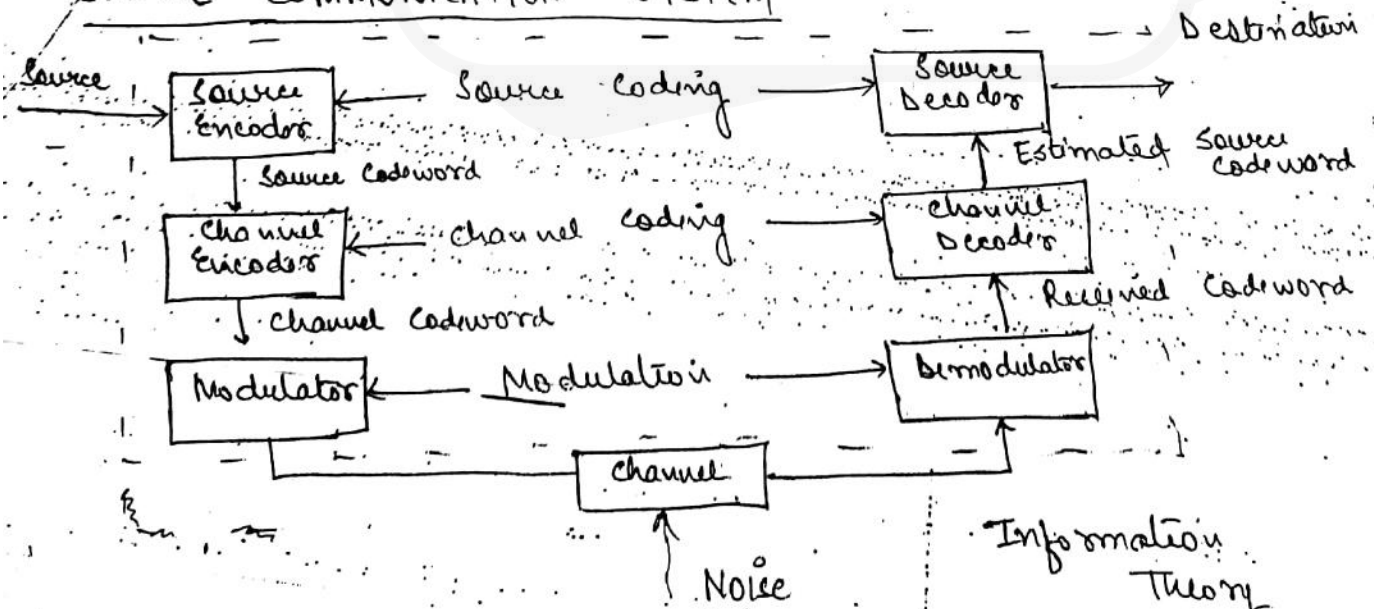
$I \geq 0$ as $0 \leq P_k \leq 1$.

Rule 3 :- Information content of message having higher probability is less than the information content of message having lower probability.

* $I(m_k \text{ and } m_i) = I(m_k\, m_i)$
$= I(m_k) + I(m_i)$

* $\log_2 x = \dfrac{\ln x}{\ln 2} = \dfrac{\log_{10} x}{\log_{10} 2}$

DIGITAL COMMUNICATION SYSTEM



Information Theory

Entropy :- Average Information content over the whole alphabet of symbols.

$$H_r(S) = \sum_{i=1}^{q} P_i \log_r \frac{1}{P_i} \qquad \begin{Bmatrix} s_1 & s_2 & \cdots & s_q \\ P_1 & P_2 & \cdots & P_q \end{Bmatrix}$$

* Entropy function involves only the distribution of the probabilities.

Ex :- Weather of Shimla

$X = \{\text{Rain, fine, cloudy, snow}\} = \{R, f, C, S\}$

$P(R) = 1/4$, $P(F) = 1/2$, $P(C) = 1/4$, $P(S) = 0$

$H_2(X) = 1.5$ bits/symbol

$r = 2$ [no. of bit assigned to symbol]

$S = X$ [name of source]

ITC

* In all modes of communication, communication is not error free. But we can improve the accuracy of transmission by reducing P_e.

P_e as $e^{-K E_b}$

Let say probability of error is denoted by P_e, but in general it is not possible to have a comm. which is completely error free if noise exists.

* P_e varies as $e^{-K E_b}$.

It means if I can increase energy / bit then I can reduce P_e.

* Now signal power is

$P_i = E_b R_b$ → bit rate

which implies if I need large value of E_b then I should increase my P_i for a given R_b or if P_i is fixed then I should reduce R_b to get large value of E_b.

But in general case, ~~reducing~~ increasing P_i beyond a certain limit is not feasible. So, if want P_e tends to 0. which is attained by increasing E_b then for a fixed P_i, R_b has to be reduced.

OR, to reduce P_e, R_b has to be reduced.

Which implies in a comm. channel where noise exists, it is not possible to have error free tx until value of R_b is v. low. B'coz to make $P_e = 0$, $R_b = 0$ for a fixed P_i. This is what all comm. engineers thought until the publication of a paper by ~~Shanon~~ Shannon in 1948.

Shannon showed in this paper is that as long as the rate of tx is less than a particular limit called as 'channel capacity' then it is possible to have error free tx. It is not necessary for R_b tends to 0 for $P_e \to 0$.

OR What he showed a is if I have my R_b less than channel capacity C, it is still possible to have error free tx.

OR we can say that disturbance which occur on a Comm. channel do not limit the accuracy of tx what it limits is the rate of tx of information.

Now, I have been using rate of tx of info. which means I should be able to quantify (measure) my information.

& this is what exactly 'information theory' does!

Information
- Synctatic ✓
- Symentic
- Pragmatic

1) Abraham was dropped to airport by taxi
ii) Taxi brought Abraham to airport.
iii) There is traffic jam on NH5 b/w M & Pune in India
iv) " " " on NH5 in India

i) & ii) carry same info – synctatic but syntax is diff. → Also symentaic carry same info.

iii) & iv) Not Symentaic (extra info)
Not Synctatic
(Diff info + Diff syntax)

iii) Pragmatic → gives extra info (NH5)

If I am interested in transmitting 1, & 2 both carry same information but from Comm. point of view, Symbols used for transmitting msg 1 & msg 2 could be different, & we are concerned with these symbols only.

Let us take

Source — Channel — Destination

Now, the o/p of source may not be really tuned to the way I should send it to the channel so, for that I have to convert this source o/p to a form suitable for tx on channel. So, usually we have 2 blocks b/w source & channel.

1st block – Source Encoder
2nd block – channel Encoder.

Source Encoder is based on the characteristics of (2). Source. It is possible that o/p of Source encoder is ready for tx on channel but b/coz of characteristics of channel you still need to modify o/p of source encoder :, channel encoder is used whose function is to take i/p from source encoder & gets o/p suitable for tx taking channel characteristics into account.

At Destination, for channel encoder, channel decoder will be, then followed by source decoder.

Random Variables :- Variables that take on values determined by probability distributions.

Consider variables like X, Y, Z

Say X = no. of heads

Y = no. of cell phones

Z = no. of movies

Thus, in basic maths, a variable is an alphabetical character that represents an unknown no.

But in probability, we also have variables, but we refer to them as random variables, which

A R.V is a variable that subject to randomness, means it can take different values.

As in maths, variables represent something & we denote them with x, y, x_2 etc. but in statistic it is we used X to represent R.V.

Each value of R.V has % or probability associated with it.

Suppose two coins are tossed & we have to find the probability of having heads.

So, there is a possibility of no heads, x = 0

1 head, x = 1

2 head, x = 2.

So, we define X (random variable) to be the no. of heads that I could get.

Now, these R.V has probability distribution associated with it.

Discrete R.V

When variable represents isolated points on the no. line such as with 0, 1, 2 we call it is a Discrete R.V

These are found by counting like no. of students in a class, marbles in a jar.

Now, tossing two coins, to get heads is a discrete value as it takes values as 0, 1, 2

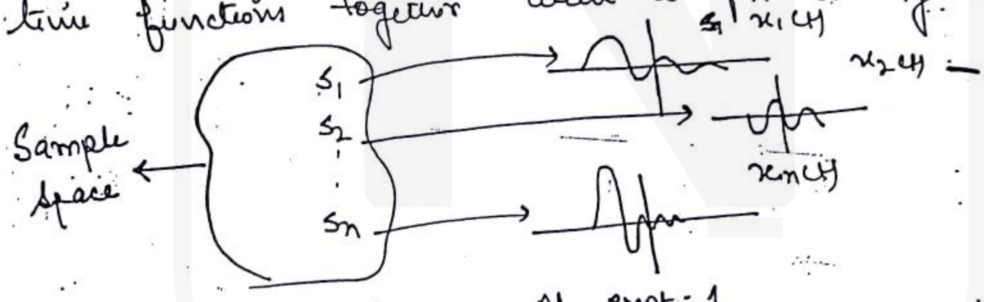
Now, with each value of variable, it has probability associated with it.

No. of heads	Prob
0	0.25
1	0.5
2	0.25

This is called probability Distribution.

Probability Distribution has all the possible values of the R.V & the associated probabilities.

A Random Process :- It is defined as ensemble of time functions together with a probability rule.



$x_1(t)$ – outcome of expt - 1

$x_2(t)$ – " " expt - 2

Each sample point in S is associated with a sample function $x(t)$

$X(t, s)$ is a R.P.

* Ensemble of ~~all~~ time fun's together with a probability rule.

Stationary Random Process :- A random process is said to be stationary if its characterization is independent of observation interval at which process is initiated.

$$F_X(t_1+T) \ldots X(t_k+T) = F_X(t_1) \ldots X(t_k)$$

$$\text{or} \quad F_X(t_1) = F_X(t_1+T)$$

Now, consider the event $S = s_k$, which emits s_k symbols by source with probability p_k.

Before, event occurs, there is amt. of uncertainity what would be the information.

When event occurs, there is a amount of surprise like what next information will occur.

When event is complete, there is a gain of information

∴ we define the amount of information gained after observing the even $S = s_k$, which occurs with probability p_k as the logarithmic fn.

$$I(s_k) = \log\left(\frac{1}{p_k}\right)$$

i) If $p_k = 1$, $I(s_k) = 0$

- which is obvious, if we are certain of the outcome of an event before it occurs, there is no information gained.

As we know $\sum_{0}^{k=1} p_k = 1$ when we have the knowledge of about all the probabilities of possible symbols emitted.

ii) $I(s_k) \geq 0$ for $0 \leq p_k < 1$

Means occurrence of events either provide some inf or no inf. but never brings about a loss of inf.

iii) $I(s_k) > I(s_i)$ for $p_k < p_i$

That is, less probable an event is, more inf. we gain when it occurs.

iv) $I(s_k s_l) = I(s_k) + I(s_l)$ if s_k & s_l are statistically independent.

If 2 bits are tx.

$$I(s_k) = \log_2\left(\frac{1}{p_k}\right) \text{ bits}$$

Discrete Memoryless Source :- which emits symbols at any time independent of previous choices.

$I(s_k)$ is discrete R.V that takes on values $I(s_0), I(s_1) \ldots I(s_{k-1})$ with probabilities $p_0, p_1, \ldots p_{k-1}$ respectively.

The mean value of $I(s_k)$

$$H = E[I(s_k)]$$

ensemble mean collection

$$= \sum_{k=0}^{K-1} p_k I(s_k)$$

$$= \sum_{k=0}^{K-1} p_k \log_2\left(\frac{1}{p_k}\right)$$

H is called entropy of a discrete memoryless source. It is a measure of av. information content per source symbol.

Properties of Entropy

The entropy of source is as follows:-

$$0 \le H \le \log_2 K$$

$K \rightarrow$ no. of symbols.

1> $H = 0$, if $p_k = 1$ for some k and the remaining probabilities in the set are all zero. This corresponds to no uncertainty.

$\begin{pmatrix} 00 - \cancel{0.25}\ 0.5 \\ 01 - \cancel{0.25}\ 0.5 \\ 11 - 0 \\ 10 - 0 \end{pmatrix}$

2> $H = \log_2 K$, if $p_k = 1$ for all K. This corresponds to max. uncertainty.

$\begin{pmatrix} 00 - 0.25 \\ 01 - 0.25 \\ 11 - 0.25 \\ 10 - 0.25 \end{pmatrix}$

—— x —— x —— x —— x —— x —— x ——

Qu 1. A source is generating 3 possible symbols with probabilities of 1/4, 1/4, 1/2 respectively. Find the information associated with each symbol.

Soln.

$\boxed{Tx} \longrightarrow x_1, x_2, x_3$ -- symbols

$P[x_1] = 1/4, \quad P[x_2] = 1/4, \quad P[x_3] = 1/2$

$$I[x_1] = \log_2 \frac{1}{P[x_1]}$$
$$= \log_2 4 = 2 \text{ bits}$$
$$I[x_2] = 2 \text{ bits}$$
$$I[x_3] = 1 \text{ bits}$$

Note: If the probability of occurrence of a symbol is more then the information associated with that symbol is less & vice-versa.

Ex: Probability of x_1, x_2 is less than x_3 that means information contained in x_1 & x_2 is more importance than x_3.
No. of bits assigned to $x_1\ x_2$ is more as compared to x_3.

Entropy :- It gives the measure of Uncertainity.

$$H = \sum_i I(x_i)\, P(x_i)$$

$$H = \sum P[x_i] \log_2 \frac{1}{P[x_i]} \text{ bits/symbol.}$$

i) If tx / source emit symbol with equal probabilities
$P[x_1] = P[x_2] = 1/2$

$$H = \sum_{i=1}^{2} P[x_i] \log_2 \frac{1}{P[x_i]}$$
$$= \frac{1}{2}\log_2 2 + \frac{1}{2}\log 2$$
$$H_{max} = 1 \text{ bit/symbol.}$$

ii) $P[x_1] = 1$ & $P[x_2] = 0$

$$H = \sum_{i=1}^{2} P[x_i] \log_2 \frac{1}{P[x_i]} = -\sum_{i=1}^{2} p(x_i) \log_2 p[x_i]$$
$$= -1 \log_2 1 + 0 \log_2 0$$
$$H_{min} = 0 \text{ bit/symbol.}$$

* If all the symbols have same probability to occur then Uncertainity will be max. & when only on symbol have prob. of occurence, there will be no. Uncertainity.

Now, for the case

$\boxed{Tx}$ $x_1 \; x_2 \; \ldots \; x_m$

$P[x_1] = P[x_2] \ldots P[x_m] = 1/M$

$$H_{max} = \sum^{M} \frac{1}{M} \log_2 1/1/M$$

$$\boxed{H_{max} = \log_2 M} \text{ bits / symbol.}$$

Information Rate

$$R = H \times r$$

$$\frac{\text{bits}}{\text{sym}} \times \frac{\text{Symbol}}{\text{Sec}}$$

$$\boxed{R = \text{bits / Sec}}$$

Que A source is generating 4 possible symbols with probabilities of 1/8, 1/8, 1/4, 1/2 resp. Find Entropy & I. Rate. If source is generating 1 symbol / m sec.

Soln.

$$r = 1000 \frac{\text{Symbol}}{\text{Sec}}$$

$$H = \sum_{i=1}^{4} P[x_i] \log_2 \frac{1}{P[x_i]}$$

$$= 1/8 \log_2 1/1/8 + \frac{1}{4} \log_2 4 + \frac{1}{2} \log_2 2$$

$$= \frac{15}{8} = 1.75 \text{ bit / symbol.}$$

$$R = r \times H$$

$$= 1000 \times 1.75$$

$$R = 1.75 \text{ Kbps}$$

Conditional Entropy

It defines the uncertainty at the output when diff. inputs are tx.

OR

It measures the information needed to describe the outcome of a R.V (Y) given that the value of another R.V (X) is known.

Represented as:-

$$H\left(\frac{Y}{X}\right) =$$

x_1, x_2, x_3, ⋮, x_n —— y_1, y_2, y_3, ⋮, y_n

$$H\left(\frac{Y}{X}\right) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 P\left(\frac{y_i}{x_i}\right)$$

↳ Uncertainty about the receiver Y when X is transmitting (given)

For ex:

(i) $P\left(\frac{y_1}{x_1}\right) = 0.5$, $P\left(\frac{y_2}{x_2}\right) = 0.5$

$P\left(\frac{y_1}{x_2}\right) = 0.49$, $P\left(\frac{y_2}{x_2}\right) = 0.51$

(ii) $P\left(\frac{y_1}{x_1}\right) = 1$, $P\left(\frac{y_2}{x_2}\right) = 0$

$P\left(\frac{y_1}{x_2}\right) = 0.99$, $P\left(\frac{y_2}{x_2}\right) = 0.01$

From this 2 ex:- There is max. uncertainty about rx when X is tx in case of 1 as compared to 2 as $H(Y/X)$ is high as compared to 2.

To calculate entropy of Binary Symmetric channel.

Consider a binary memoryless source
- binary → emits 2 symbols
  - 0 with probability p ~~&~~ p
  - 1 with probability ~~$1-p$~~ ~~$1-\alpha$~~ $1-p$
- memoryless → whose current value is independent of previous values.

Symmetric b'coz probability of receiving 0 when 1 is send is same as probability of receiving 1 when 0 is send.

X	T_x		R_x	
	x_1	0	$y_1 =$	0
	x_2	1	y_2	1

$0 \xrightarrow{1-p} 0$, $0 \xrightarrow{p} 1$, $1 \xrightarrow{p} 0$, $1 \xrightarrow{1-p} 1$

$$H\left(\frac{y}{x}\right) = -\sum_{i=1}^{2} \sum_{j=1}^{2} P[x_i, y_j] \log_2 P\left(\frac{y_j}{x_i}\right)$$

$$P\left[\left(\frac{y}{x}\right)\right] = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \end{bmatrix}$$

$$P\left[\frac{y}{x}\right] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

$$[P(x, y)] = P[x]_d \cdot P\left(\frac{y}{x}\right)$$

$$[P(x)]_d = \begin{bmatrix} P(x_1) & 0 \\ 0 & P(x_2) \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & 1-\alpha \end{bmatrix}$$

Let $P(x_1) = \alpha$
$P(x_2) = 1-\alpha$

$$[P[x, y]] = \begin{bmatrix} \alpha & 0 \\ 0 & 1-\alpha \end{bmatrix} \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

$$= \begin{bmatrix} \alpha(1-p) & \alpha p \\ p(1-\alpha) & (1-\alpha)(1-p) \end{bmatrix}$$

$$[P[x, y]] = \begin{bmatrix} P(x_1, y_1) & P(x_1, y_2) \\ P(x_2, y_1) & P(x_2, y_2) \end{bmatrix}$$

Now

$$H\left(\frac{y}{x}\right) = -\Big\{ P(x_1,y_1)\log_2\left(\frac{y_1}{x_1}\right) + P(x_1,y_2)\log_2\left(\frac{y_2}{x_1}\right) + P(x_2,y_1)\log_2\left(\frac{y_1}{x_2}\right) + P(x_2,y_2)\log_2\left(\frac{y_2}{x_2}\right)\Big\}$$

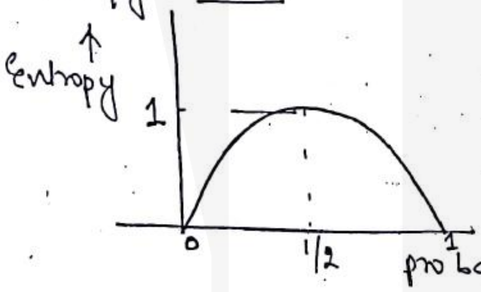
$$= -\Big\{ \alpha(1-p)\log_2(1-p) + \alpha p\log_2 p + (1-\alpha)p\log_2 p + (1-\alpha)(1-p)\log_2(1-p)\Big\}$$

$$= -\Big\{ p\log_2 p + (1-p)\log_2(1-p)\Big\}$$

Imp

$$\boxed{H\left(\frac{y}{x}\right) = p\log_2\frac{1}{p} + (1-p)\log_2\frac{1}{(1-p)}}$$

○ Entropy curve



$H = \sum_i p(x_i)\log_2\frac{1}{p(x_i)}$

If $p(x) = 0$
$H = 0$
If $p(x) = 1$
$H = 0$
If $p(x) = 1/2$
$H = \frac{1}{2}\log_2 2$
$= 0.5$

Entropy of n discrete Memoryless Source

We generally consider block of symbols originating from n sources.
So Entropy of n sources = entropy of original source

$$H(f^n) = nH(f)$$

Source Coding Theorem

Source coding is the process for representation of data.
Now we can provide short codes to ~~data~~ frequent source symbols and long codes to rare ones.
For ex: In Morse code, dots are used for freq. symbol 'E' / dash for rare one like 'Q'. represented as ... ▷ Q as dash (-).

This coding is done by source encoder. that must satisfy 2 conds :-

i) code words should be in decimal forms decodable

ii) code should be uniquely reconstructed at i.e. it can be easily Rx. end.

Consider a discrete memoryless source

$$S = \{ s_0, s_1 - - s_k \}$$

S emits $s_0, s_1, - - s_k$ symbols

Let corresponding probabilities be

$$\{ p_0, p_1 - - p_k \}$$

& code lengths be

$$\{ l_0, l_1, - - l_k \}$$

Then, average code length (av. no. of bits per symbol) of the source is defined as :-

$$\bar{L} = \sum_{k=0}^{K-1} p_k l_k$$

If L_{min} = min. possible value of $\bar{L}$, then coding efficiency is

$$\eta = \frac{L_{min}}{\bar{L}}$$

for efficient coding, η approaches Unity.

Now, question arises what is the smallest av. codelength that is possible?

Answer :- Shannon's source coding theorem

which states

Given a discrete memoryless source of entropy $H(s)$, av. codeword length $\bar{L}$ for any distortionless source encoding scheme is bounded by

$$\boxed{\bar{L} \geq H(s)}$$

∴ Since entropy limit the no. of bits / symbol ∴ Av. code word length can not be made smaller than entropy.

$$L_{min} = H(s)$$

$$\eta_1 = \frac{H(s)}{\bar{L}}$$

Data Compaction

1) Removal of redundant information prior to tx.
2) Lossless Data compaction i.e. no information is lost.
3) A source code should be Uniquely decodable

Now,
Source Coding Scheme for Data Compaction

Prefix Coding

* It is a variable length source coding scheme where no ~~code~~ code is the prefix of any other code.
* Prefix code is a Uniquely decodable code.
* But all Uniquely decodable codes are may not be prefix codes.

For ex: Consider a sequence, or set of codes

$P = \{01, 010, 10\}$

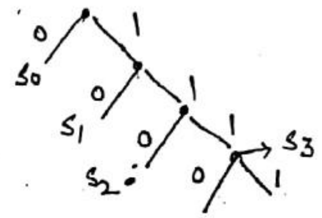
Now this '01' in 010 that mean this is not a prefix code b'coz '01' is present as a prefix in other code.

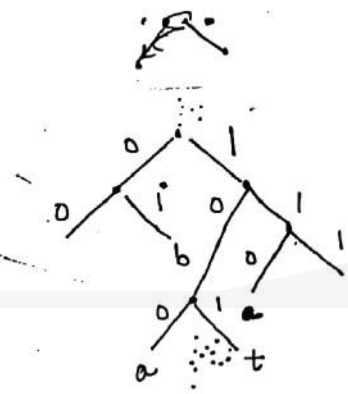
Now, Consider a code / Sequence

$P = \{01, 100, 101\} \rightarrow$ Prefix code.

$$P = \{ \underset{b}{01}, \underset{a}{100}, \underset{t}{101} \}$$

This can be represented in a form of tree





Prefix code also satisfies Kraft - Mc Millan Inequality given by

$$\sum_{k=0}^{K-1} 2^{-l_k} \leq 1$$

Given a discrete memory less source of entropy $H(S)$, a prefix code can be constructed with an average code word length $\bar{l}$, which is bounded by as follows:-

$$H(S) \leq (\bar{L}) < H(S) + 1$$

Left Hand equality is satisfied owing to the cond^n that any symbol S_k is emitted with probability

$$p_k = 2^{-l_k}$$

l_k = length of codword assigned to symbol S_K

Hence,

$$\sum_{k=0}^{k-1} 2^{-l_k} = \sum_{k=0}^{K-1} p_k = 1$$

$-l_k \log 2 = \log p_k$

$l = -\frac{\log p_k}{\log 2}$

$= -\log_2 p_k$

With this condition, Kraft - Mc Millan inequality tells that a prefix code can be constructed such that length of codeword assigned to source symbol $S_k = -\log_2 p_k$...

I, Average codeword length is given by

$$\bar{L} = \sum_{k=0}^{k-1} \frac{l_k}{2^{l_k}} \qquad \left\{ \bar{L} = \sum_{k=0}^{k-1} p_k l_k \right\}$$

& corresponding entropy

$$H(S) = \sum_{k=0}^{k-1} \frac{1}{2^{l_k}} \log_2 (2^{l_k})$$

Hence, left side inequality is proved.

for inequality, let $\bar{L}_n$ denotes the av. code word length of extended prefix code.

Huffman Coding

Huffman coding is a prefix code.

* The length of codeword for each symbol is roughly equal to the amount of information conveyed.
* Uses Data compression

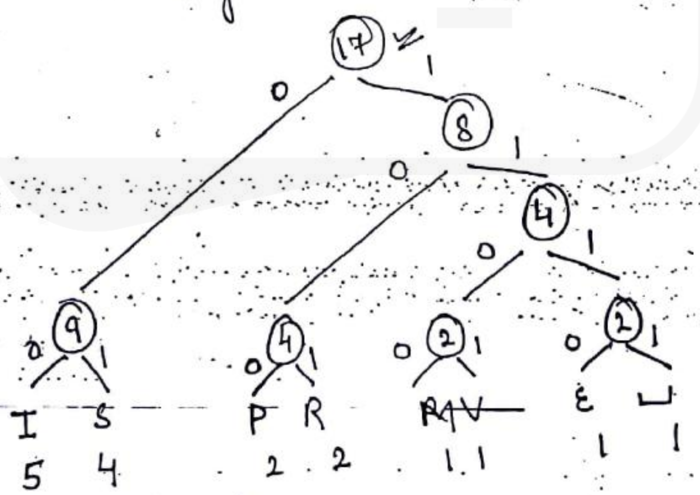
More Common Letters = fewer bits
Less Common Letters = More bits } To reduce code length

Consider An example

MISSISSIPPI ␣ RIVER = 17 characters . 8 bits = 136 bits

starting with Lowest bit symbol.

M — 1
I — 5
S — 4
P — 2
R — 2
V — 1
E — 1
␣ — 1



Code for I = 00 P = 100
S = 01 R = 101

MISSISSIPPI RIVER = 46 bits

1100 00 — — —

$\frac{46}{136}$ $\frac{136}{46}$ = 33% of original

means = 67% bits are saved from original.

Variance :- It is a measure of the variability in codeword lengths of a source code.

It is defined as

$$\sigma^2 = \sum_{k=0}^{K-1} p_k (l_k - \bar{L})^2$$

l_k → codeword length of K^{th} symbol

p_k → prob. of K^{th} symbol

$\bar{L}$ is av. codeword length.

→ It is reasonable to choose huffman tree which gives greater variance.

Drawbacks

1) Require proper statistics. (i.e. identification of more frequent words / less frequent words, their freq.)

2) Do not consider redundancy of the language.

Lempel Ziv Coding

1) overcomes drawbacks of Huffman Coding.

2) Simple Encoding Scheme.

3) Encodes pattern in text.

4) It compresses by building a dictionary of previously seen string. It codes group of characters of varying lengths.

Example

A|A B|A B B|B|A B A|A B A B|B B|A B B A|B B
1, 2, 3, 4, 5, 6, 7, 8, 9

Position	1	2	3	4	5	6	7	8	9
Sequence	A	AB	ABB	B	ABA	ABAB	BB	ABBA	BB
Numerical representation	ΦA	1B	2B	ΦB	2A	5B	4B	3A	7
Code	0	1, 1	10, 1	00, 1	010, 0	101, 1	100, 1	011, 1	0111

A → 0
B → 1

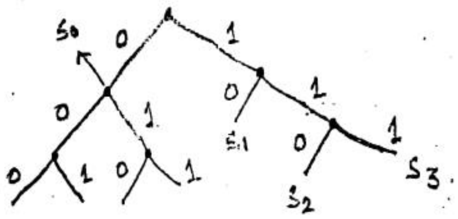
Algorithm

i) The source sequence is sequentially parsed into strings that have not appeared so far.
ii) After every separation, we look along input sequence until we come to the shortest string that has not been marked off before.
iii) We code this phrase by giving the location of the prefix & value of last bit.
iv) The code length will be max. if the string has more ~~randomness~~ randomness.

Discrete/ Memoryless channel :-

Another example of a Prefix code

Source symbol	Probabilities of occurrence	Code I	Code II	Code III
s_0	0.5	0	0	0
s_1	0.25	1	10	01
s_2	0.125	00	110	011
s_3	0.125	11	111	0111
		Not a prefix code	Prefix code	Not a prefix code



Decision Tree for Code III.

In order to decode a sequence of code-words generated from a prefix source code, the source decoder simply starts at the beginning of the sequence and decodes one code-word at a time, which is called as decision tree.

The tree has an initial state and 4 terminal states corresponding to source symbol s_0, s_1, s_2 and s_3. The decoder always starts in the initial state. The first received bit moves the decoder to the terminal state s_0 if it is 0, or else to a second decision point if it is 1.

* Although all prefix codes are Uniquely decodable, the converse is not true.
For ex:- Code 3 is Uniquely decodable but it is not a prefix code.

* Prefix codes also referred as Instantaneous code b'coz decoding of a prefix code can be accomplished as soon as the binary sequence representing a source symbol is fully received.

—x——x——x——x——x——x—

Huffman Encoding Algorithm :-

1) The source symbol is listed in order of decreasing probability. The two source symbols of lowest probabilities are assigned as 0 and 1. (Splitting stage)

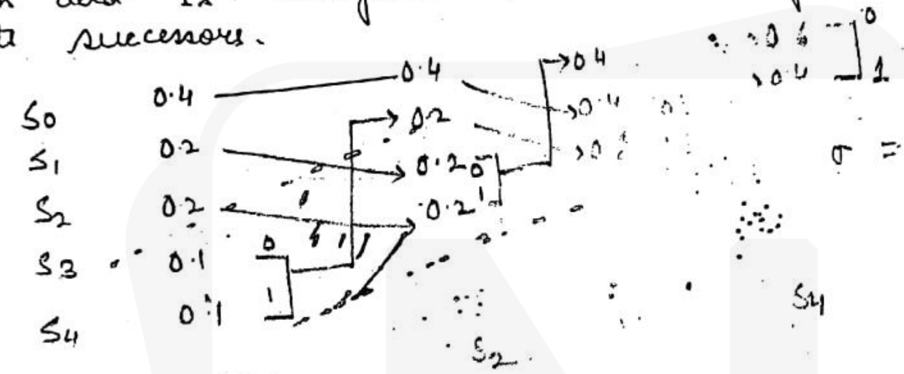
2) These 2 source symbols are regarded as being combined into a new source symbol with probability equal to the sum of two original probabilities.

The probability of the new symbol is placed in the list in accordance with its value.

3) The procedure is repeated until we are left with a final list of source statistics of only two for which a 0 and 1 are assigned.

The code for which each source is found by working backward and tracing the sequence of 0s and 1s assigned to that symbol as well as its successors.

i)



S_0	00
S_1	10
S_2	11
S_3	010
S_4	011

$\sigma = 0.4(2-2.2)^2 + 0.2(2-2.2)^2 \ldots$

$= 0.16$

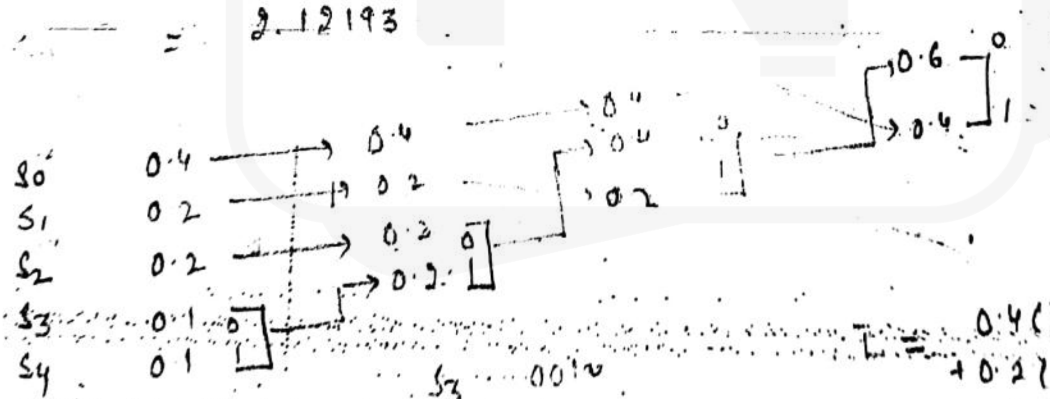
$\bar{L} = 0.4(2) + 0.2(2) + 0.2(2) + 0.1(3) + 0.1(3)$

$= 2.2$

$H(\mathcal{S}) = 0.4 \log_2\left(\frac{1}{0.4}\right) + 0.2 \log_2\left(\frac{1}{0.2}\right) \ldots$

$= 2.12193$

ii)



S_0	1
S_1	01
S_2	000
S_3	0010
S_4	0011

$\bar{L} = 0.4(1) + 0.2(2) \ldots$

$= 2.2$

$\sigma^2 = 1.36$

$= 0.4(1-2.2)^2 + 0.2(2-2.2)^2 \ldots$

$= 1.36$

These results confirm that min. variance Huffman code is obtained by moving the probability of a combined symbol as high as possible.

—— x —— x —— x —— x —— x

DISCRETE MEMORYLESS CHANNELS

A discrete memoryless channel is a statistical model with an i/p X and o/p Y that is a noisy version of X.

Every unit of time, channel accepts an input symbol X selected from alphabet $\mathcal{X}$ (collection of i/p samples)

and in response, it emits output symbol Y selected from an alphabet $\mathcal{Y}$.

$$\mathcal{X} = \{x_0, x_1, \ldots x_{j-1}\} \qquad J = \text{size of } \mathcal{X}$$

$$\mathcal{Y} = \{y_0, y_1, \ldots y_{k-1}\} \qquad K = \text{size of } \mathcal{Y}.$$

$\mathcal{X}$ & $\mathcal{Y}$ need not to have same size.

$K \geq J$ if size of $\mathcal{X}$ (K) larger than size of $\mathcal{Y}$ (J)

$K \leq J$ when channel emits the same symbol when either one of two input symbols is sent.

Transition probability $p(y_k|x_j)$ is the conditional probability that the channel output $Y = y_k$, given that the channel i/p $X = x_j$.

When $k = j$, the transition probability $p(y_k|x_j)$ represents a conditional probability of correct reception i.e. no error.

When $k \neq j$, conditional probability of error.

Discrete Memoryless channel can be represented as

$$P = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) & \cdots & p(y_{k-1}|x_0) \\ p(y_0|x_1) & p(y_1|x_1) & & \vdots \\ \vdots & \vdots & & \vdots \\ p(y_0|x_{j-1}) & p(y_1, x_{j-1}) & \cdots & p(y_k|x_{j-1}) \end{bmatrix}$$

The J-by-K Matrix P is called the channel Matrix.

Each row corresponds to a fixed channel i/p

Each column " " " fixed channel o/p.

$$\sum_{K=0}^{K-1} p(y_K | x_j) = 1 \quad \text{for all } j$$

→ Fundamental Property of channel Matrix P.-

JOINT PROBABILITY DISTRIBUTION

JOINT ENTROPY AND CONDITIONAL ENTROPY

Let channel i/p

$$[X] = [x_1, x_2, \dots x_m]$$

channel o/p

$$[Y] = [y_1, y_2, \dots y_n]$$

$$[XY] = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ \vdots & & & \\ x_my_1 & x_my_2 & \dots & x_my_n \end{bmatrix}$$

∴ we have three probabilities and three entropies

$$P(X) = [P(x_j)]$$

$$P(Y) = [P(y_k)]$$

$$P(X,Y) = [P(x_j, y_k)]$$

$$H(X) = -\sum_{j=1}^{m} P(x_j) \log P(x_j) \longrightarrow \text{Marginal entropy of } X$$

where

$$P(x_j) = \sum_{k=1}^{n} P(x_j, y_k)$$

$$H(Y) = -\sum_{k=1}^{n} P(y_k) \log P(y_k) \longrightarrow \text{Marginal entropy of } Y$$

where

$$P(y_k) = \sum_{j=1}^{m} P(x_j, y_k)$$

$$H(X,Y) = -\sum_{j=1}^{m} \sum_{k=1}^{n} P(x_j, y_k) \log P(x_j, y_k)$$

→ Joint Entropy of X and Y.

H(X) :- Average Uncertainity of channel i/p.

H(Y) :- Average Uncertainity of channel o/p.

$H(X,Y)$:- Average Uncertainity of the comm. channel as a whole.

The Conditional Probability

$$P(X/Y) = \frac{P(X,Y)}{P(Y)} \quad \text{--- } (1)$$

As y_k can occur in conjunction with $x_1, x_2, \ldots x_m$ we have

$$[X/y_k] = \left[\frac{x_1}{y_k} \; \frac{x_2}{y_k} \cdots \frac{x_m}{y_k}\right]$$

$$P[X/y_k] = [P(x_1/y_k) \; P(x_2/y_k) \cdots P(x_m/y_k)]$$

$$= \left[\frac{P(x_1,y_k)}{P(y_k)} \; \frac{P(x_2,y_k)}{P(y_k)} \cdots \frac{P(x_m,y_k)}{P(y_k)}\right]$$

— from (1)

Now,

$$= P(x_1,y_k) + P(x_2,y_k) + \cdots + P(x_m,y_k) = P(y_k)$$

Thus,

$$\sum_{j=1}^{m} P(x_j/y_k) = 1$$

$$H(\cancel{X/y_k}) = -\sum_{j=1}^{m} \frac{P(x_j,y_k)}{P(y_k)} \log \frac{P(x_j,y_k)}{P(y_k)}$$

$$= -\sum_{j=1}^{m} P(x_j/y_k) \log P(x_j/y_k)$$

Taking average of this conditional entropy for all admissible values of y_k

$$\underbrace{H(X/Y)}_{\text{av. Conditional Entropy of System}} = \overline{H(X/y_k)}$$

$$= \sum_{k=1}^{n} P(y_k) \; H(X/y_k)$$

$$= -\sum_{k=1}^{n} P(y_k) \sum_{j=1}^{m} P(x_j/y_k) \log P(x_j/y_k)$$

$$= -\sum_{k=1}^{n} \sum_{j=1}^{m} P(y_k) \, P(x_j/y_k) \log P(x_j/y_k)$$

$$= -\sum_{j=1}^{m} \sum_{k=1}^{n} P(x_j,y_k) \log P(x_j/y_k)$$

Similarly,

$$H(Y/X) = -\sum_{j=1}^{m}\sum_{k=1}^{n} P(x_j, y_k)\log P(y_k/x_j)$$

$H(X/Y)$:- Measure of average Uncertainty remaining about the channel i/p after the channel o/p has been observed.

~~$H(Y/X)$~~ * It describes how well we can recover the transmitted symbol from the received ~~transmitted~~ symbol.

$H(Y/X)$:- Measure of average Uncertainty of the channel o/p when X was transmitted.

* It describes how well we can recover the received symbols from the transmitted symbols.

Que 1 Show $H(X,Y) = H(X/Y) + H(Y)$

Sol^n:

$$H(X,Y) = -\sum_{j=1}^{m}\sum_{k=1}^{n} P(x_j, y_k)\log P(x_j, y_k)$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} P(x_j, y_k)\log\left[P(x_j/y_k)\,P(y_k)\right]$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} P(x_j, y_k)\left[\log P(x_j/y_k) + \log P(y_k)\right]$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n}\left[P(x_j, y_k)\log P(x_j/y_k) + P(x_j, y_k)\log P(y_k)\right]$$

$$= H(X/Y) - \sum_{j=1}^{m}\sum_{k=1}^{n}\left[P(x_j, y_k)\log P(y_k)\right]$$

$$= H(X/Y) - \sum_{j=1}^{m} P(x_j, y_k)\sum_{k=1}^{n}\log P(y_k)$$

$$= H(X/Y) - P(y_k)\sum_{k=1}^{n}\log P(y_k)$$

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

Mutual Information

Consider a comm. channel with an input X and an output Y.

$I(x_j ; y_k)$ = initial Uncertainity − final Uncertainity

Since $H(X)$ represents our uncertainity about the channel i/p before observing the channel o/p.

$H(X/Y)$ represents uncertainity about channel i/p after observing the channel o/p.

The difference $H(X) - H(X/Y)$ represents our uncertainity about channel i/p that can be resolved by observing the channel o/p. This important quantity is called 'mutual information' of the channel.

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

Properties of Mutual Information

i) The mutual information of a channel is symmetric, i.e.

$$I(X,Y) = I(Y,X)$$

$I(X,Y)$:- Measure of the uncertainity about the channel i/p that is resolved by observing channel o/p.

$I(Y,X)$:- Measure of uncertainity about the channel o/p that is resolved by sending the channel i/p.

$$H(X) = \sum_{j=0}^{j-1} p(x_j) \log_2 \frac{1}{p(x_j)}$$

$$= \sum_{j=0}^{j-1} p(x_j) \log_2 \frac{1}{p(x_j)} \underbrace{\sum_{k=0}^{K-1} p(y_k/x_j)}_{1}$$

$$= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(y_k/x_j) \cdot p(x_j) \log_2 \left[\frac{1}{p(x_j)}\right]$$

$$= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j)}\right]$$

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

$$= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{p(x_j/y_k)}{p(x_j)}\right]$$

$$\Rightarrow \left[H(X/Y) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j/y_k)}\right] \right]$$

from Baye's rule for conditional Probabilities,

$$\frac{p(x_j/y_k)}{p(x_j)} = \frac{p(y_k, x_j)}{p(y_k)}$$

$$\therefore \quad I(X;Y) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{p(y_k/x_j)}{p(y_k)}\right]$$

$$\Rightarrow \quad I(Y;X) = I(X;Y)$$

ii). Mutual Information is always non-negative

i.e.

$$I(X;Y) \geq 0$$

$$p(x_j/y_k) = \frac{p(x_j, y_k)}{p(y_k)}$$

$$I(X;Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{p(x_j, y_k)}{p(x_j) \cdot p(y_k)}\right)$$

$$I(X;Y) \geq 0$$

& Equal if

$$p(x_j, y_k) = p(x_j)\, p(y_k)$$

⇒ we cannot lose information, on the average, by observing the o/p of the channel.

iii) The mutual information of a channel may be expressed in terms of the entropy of the channel o/p.

$$I(X;Y) = H(Y) - \underbrace{H(Y/X)}_{\text{conditional Entropy}}$$

iv) Mutual information of a channel is related to the joint entropy of channel i/p & channel o/p by

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$H(X,Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \frac{1}{(p(x_j, y_k))}$$

$$H(X,Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{p(x_j)\,p(y_k)}{p(x_j), y_k}\right]$$

$$+ \underbrace{\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j)\,p(y_k)}\right]}$$

$$H(X,Y) = -I + \quad ''$$

$$\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j)\,p(y_k)}\right]$$

$$= \sum_{j=0}^{J-1} \log_2 \left[\frac{1}{p(x_j)}\right] \sum_{k=0}^{K-1} p(x_j, y_k)$$

$$+ \sum_{k=0}^{K-1} \log_2 \left[\frac{1}{p(y_k)}\right] \sum_{j=0}^{J-1} p(x_j, y_k)$$

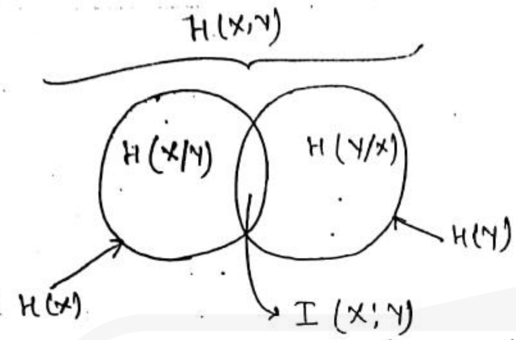
$$= \sum_{j=0}^{J-1} p(x_j) \log_2 \left[\frac{1}{p(x_j)}\right]$$

$$+ \sum_{k=0}^{K-1} p(y_k) \log_2 \left[\frac{1}{p(y_k)}\right]$$

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

$$H(X,Y) = -I(X;Y) + H(X) + H(Y)$$

EEE



CHANNEL CAPACITY

— x — x — x — x — x — x

Various Entropies of Tx and Rx & their Properties :

$\boxed{Tx}$ ~~~ $\boxed{Rx}$

$A = \{a_1, a_2, \ldots a_m\}$ $\qquad$ $B \in \{b_1, b_2, \ldots b_n\}$

$\sum_{i=1}^{m} p(a_i) = 1$ $\qquad$ $\sum_{j=1}^{n} p(b_j) = 1$

Property 1 :- Self entropy of the tx is

$$H(A) = \sum_{i=1}^{m} P(a_i) \log_2 (1/a_i) \text{ bits/symbol}$$

Property 2 : Self entropy of Rx is

$$H(B) = \sum_{j=1}^{n} p(b_j) \log_2 (1/b_j) \text{ bits/symbol}$$

Property 3 :- Conditional Entropy of Tx and Rx

$$H(A/B) = -\sum_{i=1}^{m} \sum_{j=1}^{n} P(a_i, b_j) \log_2 P(a_i/b_j) \text{ bits/symbol}$$

$$H(B/A) = -\sum_{i=1}^{m} \sum_{j=1}^{n} P(a_i, b_j) \log_2 P(b_j/a_i) \text{ bits/symbol}$$

Property 4 :- Joint entropy of Tx & Rx is

$$H(A,B) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P(a_i, b_j)\log_2 P(a_i, b_j) \text{ bits/symbol.}$$

Property 5 :-

$$H(A,B) = H(A) + H(B/A)$$

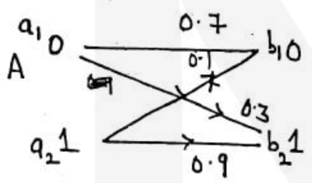
$$H(A,B) = H(A/B) + H(B)$$

Ques1 A binary channel is described by fig

~~P(A) =~~

$P(A=0) = p(a_1) = 0.7$ $\quad P(A=1) = p(a_2) = 0.3$

~~P(B=0) = p(b.~~



find

a) $H(A)$ b) $H(B)$ c) $H(A,B)$ d) $H(A/B)$ e) $H(B/A)$

Soln: channel Matrix

$$P(B/A) = \begin{array}{c} \\ a_1 \\ a_2 \end{array}\begin{array}{c} \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{bmatrix} 0.7 & 0.3 \\ 0.1 & 0.9 \end{bmatrix} \end{array}$$

a) $H(A) = P(a_1)\log_2(1/P(a_1)) + P(a_2)\log_2(1/P(a_2))$

$= 0.7\log_2\left(\frac{1}{0.7}\right) + 0.3\log_2\left(\frac{1}{0.3}\right)$

$= 0.88125$ b/symbol

b) $H(B) =$

$P(b_1) = P(a_1)P(b_1/a_1) + P(a_2)P(b_1/a_2)$

$= 0.7 \cdot 0.7 + 0.3 \times 0.1$

$= 0.52$

$\left[\begin{array}{l}\text{Total Prob. Theorem} \\ P(B) = \sum_{j=1} P(A)\, P(B/A)\end{array}\right]$

$P(b_2) = P(a_1)P(b_2/a_1) + P(a_2)P(b_2/a_2)$

$= 0.7 \times 0.3 + 0.3 \times 0.9 = 0.48$

EE

* $P(b_1) + P(b_2) = 1$

$$H(B) = P(b_1) \log_2 (1/P(b_1)) + P(b_2) \log_2 (1/P(b_2))$$

$$= 0.52 \log_2 (1/0.52) + 0.48 \log_2 (1/0.48)$$

$$= 0.998 \text{ b/s}$$

c) $H(A, B) \rightarrow$ JPM

$$P(b_1/a_1) = \frac{P(a_1, b_1)}{P(a_1)}$$

$$P(a_1, b_1) = 0.7 \times 0.7 = 0.49$$

$$P(a_1, b_2) = P(b_2/a_1) \cdot P(a_1) = 0.3 \times 0.7 = 0.21$$

$$P(a_2, b_1) = P(b_1/a_2) \; P(a_2) = 0.1 \times 0.3 = 0.03$$

$$P(a_2, b_2) = P(b_2/a_2) \; P(a_2) = 0.9 \times 0.3 = 0.27$$

$$P(A, B) = \begin{array}{c|cc} & b_1 & b_2 \\ \hline a_1 & 0.49 & 0.21 \\ a_2 & 0.03 & 0.27 \end{array} \quad \left\{ \begin{bmatrix} P(a_1, b_1) & P(a_1, b_2) \\ P(a_2, b_1) & P(a_2, b_2) \end{bmatrix} \right.$$

$$\Sigma P(a_i) = 1 = \Sigma P(a_2) = 1$$

d) $$H(B/A) = - \sum_{i=1}^{m} \sum_{j=1}^{n} P(a_i, b_j) \log_2 P(b_j/a_i)$$

$$= - \{ 0.49 \log_2 0.7 + 0.21 \log_2 0.3 + 0.03 \log_2 0.1 + 0.27 \log_2 0.9 \}$$

$$= 0.7576 \text{ b/s}$$

e) $$H(A, B) = - \sum_{i=1}^{m} \sum_{j=1}^{n} P(a_i, b_j) \log_2 P(a_i, b_j)$$

$$= - \{ 0.49 \log_2 0.49 + 0.21 \log_2 0.21 + 0.03 \log_2 0.03 + 0.27 \log_2 0.27 \}$$

$$= 1.6388 \text{ b/symbol}$$

To cross-check

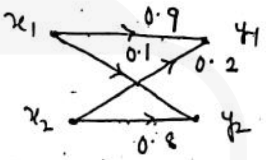
$$H(A,B) = H(A) + H(B/A)$$
$$= 1.6388.$$

d) $H(A,B) = H(A/B) + H(B)$

$$H(A/B) = H(A,B) - H(B)$$
$$= 1.6388 - 0.9987$$
$$= 0.6401 \text{ b/sy}$$

Qus

$$P(X=0) = P(x_1) = 0.7$$
$$P(X=1) = P(x_2) = 0.3$$



find i) Probability of o/p symbol

ii) find conditional probability of o/p symbols

iii) find joint probabilities of i/p & o/p symbols..

Soln: channel Matrix

$$P(Y/X) = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

a) $P(y_1) = P(x_1)\,P(y_1/x_1) + P(x_2)\,P(y_1/x_2)$

$$= 0.7 \times 0.9 + 0.3 \times 0.2 = 0.69$$

$P(y_2) = P(x_1)\,P(y_2/x_1) + P(x_2)\,P(y_2/x_2)$

$$= 0.7 \times 0.1 + 0.3 \times 0.8 = 0.31$$

$P(y_1) + P(y_2) = 1$

b) $P(x_1/y_1) = \dfrac{P(x_1, y_1)}{P(y_1)}$, $\dfrac{P(x_1/y_1)}{P(y_1/x_1)} = \dfrac{P(x_1)}{P(y_1)}$

$P(y_1/x_1) = \dfrac{P(x_1, y_1)}{P(x_1)}$ $\quad P(x_1/y_1) = \dfrac{P(y_1/x_1)\,P(x_1)}{P(y_1)}$

$$= \frac{0.9 \times 0.7}{0.69}$$
$$\simeq 0.913$$

Similarly,

$$P(x_1|y_2) = \frac{P(y_2|x_1)\,P(x_1)}{P(y_2)}$$

$$= 0.2258$$

$$P(x_2|y_1) = \frac{P(y_1|x_2)\,P(x_2)}{P(y_1)} = 0.0869$$

$$P(x_2|y_2) = \frac{P(y_2|x_2)\,P(x_2)}{P(y_2)} = 0.774$$

c) $P(x_1, y_1) = P(y_1|x_1)\,P(x_1)$ or $P(x_1|y_1)\,P(y_1)$

$$= 0.9 \times 0.7$$

$$P(x_1, y_2) = P(y_2|x_1)\,P(x_1)$$

$$= 0.1 \times 0.7 = 0.07$$

$$P(x_2, y_1) = P(y_1|x_2)\,P(x_2)$$

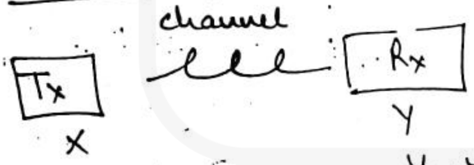
$$= 0.2 \times 0.3 = 0.06$$

$$P(x_2, y_2) = P(y_2|x_2) \cdot P(x_2)$$

$$= 0.8 \times 0.3 = 0.24$$

— x — x — x — x — x — x —

Mutual Information



$x_1, x_2, \ldots x_m$ $\qquad$ $y_1, y_2, \ldots y_n$

$\sum_{i=1}^{m} P(x_i) = 1$ $\qquad$ $\sum_{j=1}^{n} P(y_j) = 1$

If the channel is noise free, the average amount of info would be given as $H(X)$ bits/symbol. B'coz of noise, we receive on an average $H(Y/X)$ bits of info/symbol. ∴ In this transaction, amt of information R_x recieves is on the avg $I(X, Y)$ bits/symbol.

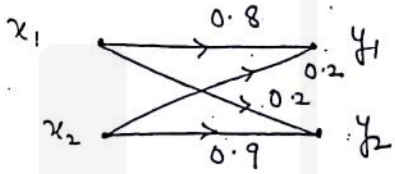
where $I(X, Y) = H(X) - H(X|Y)$ bits/symbol.

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

Properties of Mutual Information

1) M.I is always +ve i.e. $I(X,Y) \geq 0$
2) $I(X,Y) = I(Y,X) \Rightarrow$ symmetrical property of propagation, means interchanging the role of Tx & Rx does not change Mutual Info.
3) $I(X,Y) = 0$ if x_i & y_j are statistically independent.
4) $I(X,Y) = H(X) + H(Y) - H(X,Y)$

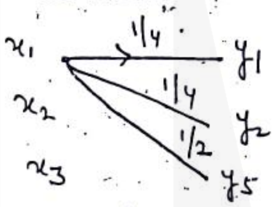
Que $P(x_1) = 0.6$, $P(x_2) = 0.4$



find $H(X)$, $H(Y)$, $H(X,Y)$, $H(Y,X)$, $H(X/Y)$, $H(Y/X)$, $I(X,Y)$

Lossless channel :- only one Non-zero element in each column

$$\begin{array}{c|ccccc} & y_1 & y_2 & y_3 & y_4 & y_5 \\ \hline x_1 & 1/4 & 1/4 & 0 & 0 & 1/2 \\ x_2 & 0 & 0 & 1 & 0 & 0 \\ x_3 & 0 & 0 & 0 & 1 & 0 \end{array}$$

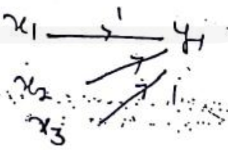


State transition diag

$\Rightarrow H(X|Y) = 0$

Deterministic channel :- Non zero element in each row.

ex:- $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$



$\Rightarrow H(Y/X) = 0$

Noiseless channel :- Both Lossless & Deterministic

e.x: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

JOINT ENTROPY

Entropy of Random Variable X over $\mathcal{X}$

$$H(X) = -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x)$$

for $X \sim p$, $H(X) = -\sum_x p(x) \log p(x)$

* The Joint Entropy – entropy of two random variables jointly distributed with $(X \& Y)$

$(X,Y) \sim p$ is

$$H(X,Y) = -\sum_{x,y} p(x,y) \log p(x,y)$$

This is simply the entropy of rv $Z = (X,Y)$

Example:

X \ Y	0	1
0	1/4	1/4
1	1/2	0

$$H(X,Y) = -\frac{1}{2}\log\frac{1}{1/2} - \frac{1}{4}\log\frac{1}{1/4} - \frac{1}{4}\log 1/1/4$$

$$= -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log 1/4 - \frac{1}{4}\log 1/4$$

$$= -\frac{1}{2}(\log 1 - \log 2) - \frac{1}{4}(\log 1 - \log 1/4) - \frac{1}{4}(\log 1 - \log 4)$$

$$= \frac{1}{2}\times 1 + \frac{1}{4}\times 2 + \frac{1}{4}\times 2$$

$$= \frac{3}{2}$$

Conditional Entropy [$H(Y/X)$]

Measures uncertainity in Y given X.

Uniform / Symmetric channel :-

ex :-
$$\begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

Binary Symm channel :- 2 i/p , 2 o/p

$$P(Y/X) = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

Channel Capacity

1) channel capacity per symbol :-

$$C_s \triangleq \max_{\{p(x_i)\}} I(X,Y)$$

2) channel capacity per second :- If r_s symbols are tx of information / sec tx / sec the max rate of tx is C_s.

$$C = C_s \cdot r_s \text{ bits/sec.}$$

channel capacity of special channels

a) Lossless :-
$$C_s = \max_{\{p(x_i)\}} I(X,Y)$$
$$= \max [H(X) - \cancel{H(X/Y)}^{\,0}]$$
$$= \log_2 M \text{ bits / symbol}$$
if M symbols are emitted at i/p

b) Deterministic :-
$$C_s = \max_{\{p(x_i)\}} I(X,Y)$$
$$= \max [H(Y) - \cancel{H(Y/X)}^{\,0}]$$
$$= \log_2 n$$
n symbols at o/p

c) Symmetric :-
$$C_s = \max I(X,Y)$$
$$C_s = \max [H(Y) - H(Y/X)]$$

$$C_s = \log_2 n + \max_{\{p(x_i)\}} \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) \log_2 p(y_j/x_i) \quad \text{—— (1)}$$

$$p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}$$

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

(2) in (1)

$$C_s = \log_2 n + \max \underbrace{\sum_{i=1}^{m} p(x_i)}_{1} \left[\sum_{j=1}^{n} p(y_j|x_i) \log_2 p(y_j/x_i) \right]$$

$$= \log_2 n + \sum_{j=1}^{n} p(y_j|x_i) \log_2 p(y_j|x_i) \quad b/s$$

Que Find Capacity of Binary Symmetric channel

Soln:

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

p = transition probability

$$C_s = \log_2 n + \sum_{j=1} p(y_j|x_i) \log_2 p(y_j|x_i)$$

$$n = 2$$

$$C_s = \log_2 2 + (1-p)\log_2(1-p) + p\log_2 p$$

$$C_s = 1 + (1-p)\log_2(1-p) + p\log_2 p \quad \text{bit/symbol.}$$

Que

$$p(Y/X) = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \qquad p(x_1) = 0.6,\quad p(x_2) = 0.4$$

find $I(X,Y)$ & C_s & η of channel.

$$\Rightarrow \text{channel Efficiency} = \frac{I(X,Y)}{C_s} \times 100$$

Differential Entropy

~~To distinguish Absolute entropy~~

So, far we considered discrete R.V, now considering continuous R.V, Entropy associated with continuous R.V — Differential Entropy to distinguish from absolute entropy.

$$h(X) = -\int_{-\infty}^{\infty} f_X(x) \log_2 \left[\frac{1}{f_X(x)}\right] dx$$

pdf.

By considering the continuous R.V. X as the limiting form of a discrete R.V that assumes the value $x_k = k\Delta x$, where $k = 0, \pm 1, \pm 2$ & Δx approaches to zero.

By definition, continuous R.V X assumes a value in the interval $[x_k, x_k + \Delta x]$ with probability $f_X(x_k)\Delta x$.

$$H(X) = \lim_{\Delta x \to 0} \sum_{k=-\infty}^{\infty} f_X(x_k)\Delta x \log_2 \left(\frac{1}{f_X(x_k)\Delta x}\right)$$

$$= \lim_{\Delta x \to 0} \left[\sum_{k=-\infty}^{\infty} f_X(x_k) \log_2 \left(\frac{1}{f_X(x_k)\Delta x}\right)\Delta x - \log_2 \Delta x \sum_{k=-\infty}^{\infty} f_X(x_k)\Delta x\right]$$

$$= \int_{-\infty}^{\infty} f_X(x) \log_2 \left(\frac{1}{f_X(x)}\right) dx - \lim_{\Delta x \to 0} \log_2 \Delta x \int_{-\infty}^{\infty} f_X(x)\,dx$$

$$= h(X) - \lim_{\Delta x \to 0} \log_2 \Delta x \qquad \left[\int_{-\infty}^{\infty} f_X(x)\,dx = 1\right]$$

In $\lim \Delta x \to 0$ log terms approaches ∞, means entropy of continuous R.V is infinitely large which is obvious as continuous R.V may assume a value b/w $[-\infty, \infty]$, Uncertainty associated with a variable is of order ∞. We avoid this problem by adopting $h(X)$ as a differential entropy with term $-\log_2 \Delta x$.

Qu Find the differential entropy of uniformly distributed random variable 'x' having the following pdf.

$$f_X(x) = \begin{cases} 1/a, & 0 \le x \le a \\ 0, & \text{otherwise} \end{cases}$$

Soln: $h(X) = -\int_{-\infty}^{\infty} f_X(x) \log_2 f_X(x)\, dx$

$$= -\int_0^a \frac{1}{a} \log_2\left(\frac{1}{a}\right) dx$$

$$= -\frac{1}{a} \log_2 \frac{1}{a} (x)_0^a$$

$$= -\frac{1}{a} \log_2\left(\frac{1}{a}\right) \cdot a - 0$$

$$= -\frac{1}{a} \cdot a \log_2 \frac{1}{a}$$

$$= \log_2 a \text{ bits / symbol}$$

SHANNON - HARTLEY LAW (CHANNEL CAPACITY THEOREM)

It is used for finding ~~the~~ capacity of a channel in which noise is white & gaussian. The capacity of such a channel is given by

$$C = B \log_2 [1 + S/N] \text{ -- bits / sec.}$$

where,

C = channel capacity

B = channel BW in Hz

N = noise power in watts $= \eta B$

S = avg. signal power in watts.

$\frac{\eta}{2} = \frac{N_0}{2}$ = noise power spectral density

II Unit

Source Coding Theorem

Shannon's Ist theorem :- By which we can measure the information emerging from a discrete memoryless source.
This theorem tells that we can make average no. of bits per source symbol as small as, but no smaller than entropy of the source measured in bits.

Channel Coding Theorem :-

Shannon's IInd theorem :- for BSC, the channel coding theorem tells us that for any code rate r less than or equal to the channel capacity C, codes do exist such that av. probability of error is as small as we want it.

BSC is simplest form of a Discrete memoryless channel. It is symmetric b'coz probability of receiving a 1 if 0 is sent is same as probability of receiving 0 if 1 is sent.

Channel capacity Theorem :- Max. rate at which information (bits) can be sent through a communication channel without error. Max. rate is called 'channel capacity'.

When system operates at a rate greater than channel capacity, it is subjected to high probability of Error, regardless of choice of signal used for transmission or Rx used for processing the received signal.

Shannon Hartley Law proof :-

$$C = B \log_2 \left(1+\frac{S}{N}\right)$$

$$= B \log_2 \left(1+\frac{S}{\eta B}\right) \quad \text{--- (1)}$$

Let $\lambda = \dfrac{S}{\eta B}$

$$B = \frac{S}{\eta \lambda}$$

eq^n (1) becomes

$$C = B \log_2 \left(1 + \frac{S\eta\lambda}{\eta S}\right)$$

$$= \frac{S}{\eta \lambda} \log_2 (1+\lambda)$$

$$= \frac{1}{\log_e 2} \frac{S}{\eta\lambda} \log_e (1+\lambda).$$

As $B \to \infty$, $\lambda \to 0$

Hence,

$$C_\infty = \lim_{\lambda \to 0} \frac{1.44 S}{\eta} \log_e \left(\frac{1+\lambda}{\lambda}\right)$$

Applying L Hospital's Rule

$$C_\infty = \lim_{\lambda \to 0} \frac{1.44 S}{\eta} \frac{\frac{1}{1+\lambda}}{1}$$

$$\boxed{C_\infty = \frac{1.44 S}{\eta}}$$

⇒ B.W increases as, so channel capacity increase & this is at the cost of (S/N) ratio b'coz $N = \eta B$. Hence, there must be trade off b/w channel BW & (S/N) ratio

In the limiting case, as $B \to \infty$, we get $C_\infty = \dfrac{1.44 S}{\eta}$

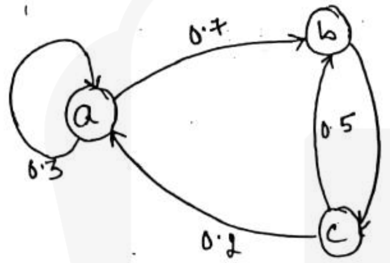
Two types of sources:-

1) The memoryless source:- Each symbol independent of the previous one

2) Markov source:- Each symbol depends on the previous one.

The Markov source

A symbol depends only on the previous symbol, so that source can be modelled by a state diagram.



A ternary source with alphabet $A = (a, b, c)$

Assume we are in state a ie $X_k = a$

$P(X_{k+1} = a \mid X_k = a) = 0.3$ $\quad P(X_{k+1} = c \mid X_k = a) = 0$

$P(X_{k+1} = b \mid X_k = a) = 0.7$

$P(X_n \mid X_{n-1})$

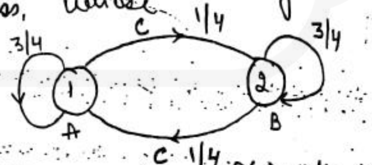
so if $\cancel{X_{k+1} = b}$ X_{k+2} will be c

$P(X_{k+2} = a \mid X_{k+1} = b) = 0$

$P(X_{k+2} = b \mid X_{k+1} = b) = 0$

$P(X_{k+2} = c \mid X_{k+1} = b) = 1$

Que Consider an information source modelled by a discrete Markov Random process, whose graph is shown:-



a) find entropy of each state

b) find entropy of source, H.

$P(1) = 1/2$

$P(2) = 1/2$

c) avg. information content per symbol in messages containing 1, 2 & 3 symbols.

find G_1, G_2 & G_3

a) To find Entropy of each state

$$H_i = \sum_{j=1}^{n} P_{ij} \log_2 \frac{1}{(P(ij))}$$

Where, P_{ij} = probability of symbols emitted by i^{th} state.

n = no. of symbols emitted in i^{th} state.

$$H_1 = \frac{3}{4} \log_2 \frac{4}{3} + \frac{1}{4} \log_2 4$$

$= 0.8113$ bits / symbol.

$$H_2 = \frac{3}{4} \log_2 \frac{4}{3} + \frac{1}{4} \log_2 4$$

$= 0.8113$ bits / symbol.

b) Entropy of source

$H(s) = H_1 + H_2$

$= 0.8113$ bits / symbol.

c) i) To find G_1

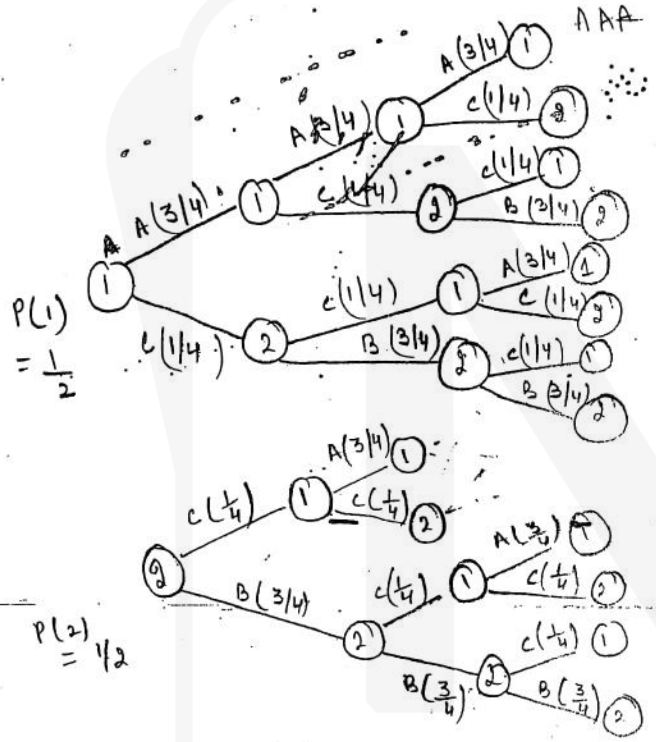
messages	Probabilities	
A	$\frac{3}{4} \times \frac{1}{2}$	$= 3/8$
B	$\frac{3}{4} \times \frac{1}{2}$	$= 3/8$
C	$\frac{1}{4} \times \frac{1}{2} + \frac{1}{4} \times \frac{1}{2}$	$= \frac{1}{4}$
	Total =	$\overline{1}$

$$G_1 = 2 \times \frac{3}{8} \log_2 (8/3) + \frac{1}{4} \log_2 4$$

$= 1.561$ bits / symbol

To find G_2

messages	Probabilities
AA	$\frac{1}{2} \times \frac{3}{4} \times \frac{3}{4} = 9/32$
AC	$\frac{1}{2} \times \frac{3}{4} \times \frac{1}{4} = 3/32$
CC	$\frac{1}{2} \times \frac{1}{4} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} \times \frac{1}{4} = \frac{2}{32}$
CB	
CA	
BC	
BB	



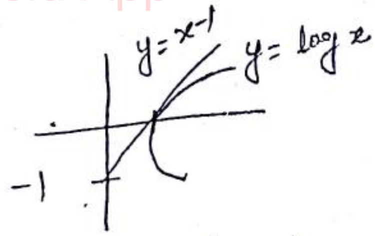
$$G_2 = 2 \times \frac{9}{32} \log_2\left(\frac{32}{9}\right) + \frac{2}{32} \log_2\left(\frac{32}{2}\right) + 4 \times \frac{3}{32} \log_2 \frac{32}{3}$$

$= 2.56$ bits/message

or $\frac{2.56}{2} = 1.28$ bits/symbol.

messages	probabilities
AAA	$\frac{1}{2} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = 27/128$
AAC	$\frac{1}{2} \times \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} = 9/128$
ACC	$\frac{1}{2} \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} = 3/128$

ACB
CCA
CCC
CBC
CBB
CAA
CAC
CCB
BCA
BCC
BBC
BBB



$\log x \leq x-1$

$$G_3 = 2 \times \frac{27}{128} \log_2 \frac{128}{27} + \frac{3\times 6}{128} \log_2 \left(\frac{128}{3}\right) + \frac{2}{128} \log_2 \left(\frac{128}{2}\right) + \frac{9}{128} \log_2 \left(\frac{128}{9}\right)$$

$= 3.418$ bits / message

or $\frac{3.418}{3} = 1.139$ bits / symbol

$G_1 \geq G_2 \geq G_3 \geq H$

Differential Entropy & Mutual Information for Continuous Ensembles

We define mutual information b/w the random variables X and Y as follows:

$$I(X;Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \log_2 \left[\frac{f_X(x/y)}{f_X(x)}\right] dx\, dy$$

Mutual information has following properties :-

1) $I(X;Y) = I(Y;X)$
2) $I(X;Y) \geq 0$
3) $I(X;Y) = h(X) - h(X/Y)$
$= h(Y) - h(Y/X)$

$h(X)$ = differential entropy of X
$h(Y)$ = " " of Y

Scanned by CamScanner

$h(X|Y)$:- Conditional differential entropy of X, given Y.

$$h(X|Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2\left[\frac{1}{f_X(x|y)}\right]dx\,dy$$

$h(Y/X)$: Conditional differential entropy of Y, given X.

Information Capacity of ~~Gaussian~~ Noise channel

Consider a zero-mean stationary process $X(t)$ that is Band-limited to B Hz.

Let X_K $(K=1,2,\dots K)$, denote the continuous R.V obtained by sampling at Nyquist rate of $2B$ samples/second. These samples are transmitted over, also in T seconds over a noisy channel band-limited to B Hz, hence no. of samples K is given by

$$K = 2BT$$

The channel o/p is perturbed by AWGN of zero mean & power spectral density $N_0/2$.

$$Y_K = X_K + N_K$$

N_K ↑ Noise with zero mean & variance

Model of discrete-time, memoryless Gaussian channel.

$$\sigma^2 = N_0 B$$

The transmitter is power limited; it is therefore reasonable to define the cost as

$$E[X_K^2] = P$$

P → Avg. transmitted power.

Information capacity,

$$C = \max_{f_{X_K}(x)}\left\{ I(X_K;Y_K) : E[X_K^2] = P \right\}$$

$$I(X_k; Y_k) = h(Y_k) - h(Y_k/X_k)$$

$$h(Y_k/X_k) = h(N_k) \quad (A) \qquad \because Y_k = X_k + N_k$$

$$I(X_k; Y_k) = h(Y_k) - h(N_k)$$

$$C = I(X_k; Y_k) : X_k \text{ Gaussian}, \quad E[X_k^2] = P$$

For evaluation of the information capacity C, we proceed in three stages:-

1) Variance of samples Y_k of the received signal equals $P + \sigma^2$.
Hence, the ~~use of~~ differential entropy of Y_k as

$$h(Y_k) = \frac{1}{2}\log_2\left[2\pi e(P+\sigma^2)\right] \quad -(1)$$

2) Variance of Noise sample N_k equals σ^2.

$$h(N_k) = \frac{1}{2}\log_2(2\pi e\sigma^2) \quad -(2)$$

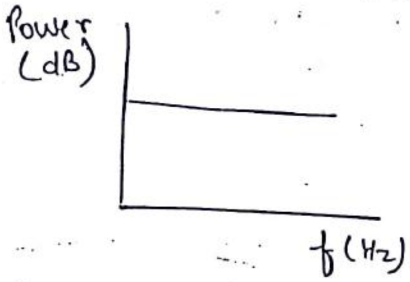
Substituting (1) & (2) in (A)

$$C = \frac{1}{2}\log_2\left(1+\frac{P}{\sigma^2}\right) \text{ bits per transmission}$$

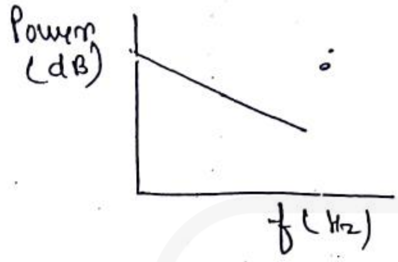
$$C = B\log_2\left(1+\frac{P}{N_0 B}\right) \text{ b/second}$$

$\Rightarrow$ It is easier to increase the information capacity of a communication channel by expanding its BW than increasing the transmitted power for a prescribed noise variance.

Information Capacity of Colored Noise channel :

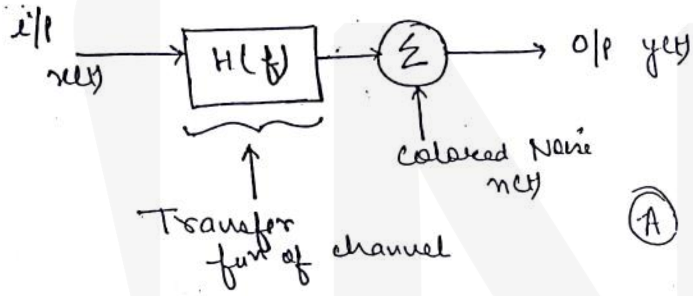


White Noise :- Noise signal whose power spectrum is flat.



Colored Noise :- Power concentration varies at different frequencies.

Consider the channel Model as shown in figure



n(t) :- channel noise, appears at o/p, modeled as sample fun of a stationary Gaussian process of zero mean & power spectral Density $S_N(f)$

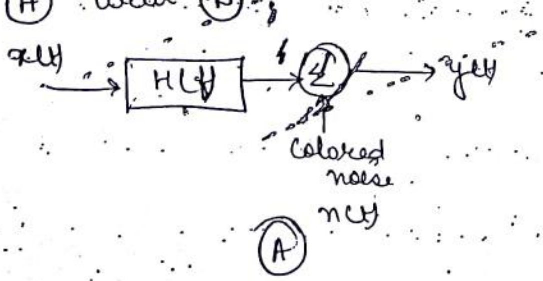
Two objectives :-

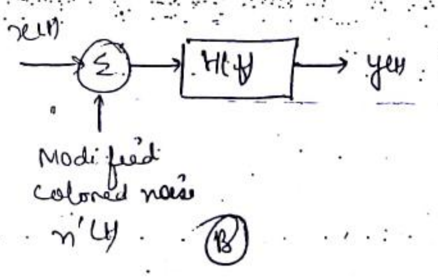
i) Find input ensemble, described by power spectral density ($S_X(f)$: power present in the signal as a function of frequency) that maximize the mutual information between the channel output y(t) and the channel i/p x(t).

ii) Determine the optimum information capacity of the channel.

To solve it,

→ Because the channel is linear, we may replace (A) with (B) ;





Two models are equivalent, provided that power spectral Density of the noise is expressed in terms of power spectral Density of noise $n(t)$

$$S_N'(f) = \frac{S_N(f)}{|H(f)|^2}$$

Now, channel is divided into a large number of frequency slots. The ~~better~~ smaller we make the incremental frequency interval Δf of each subchannel, the better is this approximation.

Model A is replaced by parallel combination of finite no. of subchannels N, each is corrupted by "band - limited white Gaussian Noise".

The k^{th} subchannel is represented by

$$y_k(t) = x_k(t) + n_k(t) \quad , \quad k = 1, 2, \ldots N$$

Average power of signal $x_k(t)$ is

$$P_k = \underbrace{S_x(f_k)}_{\text{Spectral Density of input at freq. } f_k} \Delta f \quad , \quad k = 1, 2, \ldots N \quad \text{—— (A)}$$

Variance of noise $n_k(t)$

$$\sigma^2_k = \frac{S_N(f_k)}{|H(f_k)|^2} \Delta f \quad , \quad k = 1, 2, \ldots N \quad \text{—— (B)}$$

Information capacity of k^{th} subchannel is

$$C_k = \frac{1}{2} \Delta f \log_2 \left(1 + \frac{P_k}{\sigma^2_k} \right) \quad \text{—— (E)}$$

$C = B \log_2 \left(1 + \frac{S}{N}\right)$

$S_p = P$

$\text{Noise} = \sigma^2$

$\frac{1}{2}$ accounts for the fact that Δf applies to both +ve & -ve frequencies.

All N subchannels are independent of one another. Hence, total capacity of the overall channel is

$$C \cong \sum_{k=1}^{N} C_k$$

$$= \frac{1}{2} \sum_{R=1}^{N} \Delta f \log_2 \left(1 + \frac{P_R}{\sigma^2_R}\right)$$

Summation of signal power from all frequencies

$$\sum_{R=1}^{N} P_R = P = \text{Constant}$$ [Constraint to maximize the overall information capacity]

$$J = \frac{1}{2} \sum_{R=1}^{N} \Delta f \log_2 \left(1 + \frac{P_R}{\sigma^2_R}\right) + \underbrace{\lambda}_{\text{lagrange Multiplier}} \left(P - \sum_{R=1}^{N} P_R\right)$$

$$\frac{dJ}{dP_K} = \frac{1}{2} \sum_{R=}^{N} \Delta f \log_2 \frac{1}{1 + \frac{P_K}{\sigma^2_K}} \times \frac{1}{\sigma^2_K} + \lambda(-1)$$

$$\frac{dJ}{dP_K} = \frac{\Delta f \log_2 e}{P_K + \sigma^2_K} - \lambda = 0$$

To find the solution

$$P_R + \sigma^2_R = K \Delta f \qquad R = 1, 2, \ldots N \quad (C)$$

Putting (A) & (B) in (C)

$$S_X(f_R)\Delta f + \frac{S_N(f_R)}{|H(f_R)|^2} \Delta f = K \Delta f$$

$$S_X(f_R) = K - \frac{S_N(f_R)}{|H(f_R)|^2}, \quad R = 1, 2, \ldots N$$

$$S_X(f_R) = \begin{cases} K - \dfrac{S_N(f)}{|H(f)|^2}, & \text{for } f \in f_A \\ 0, & \text{otherwise} \end{cases}$$

Let f_A denote the frequency range for which K satisfies the Condn

$$K \geq \frac{S_N(f)}{|H(f)|^2}$$

Average power of a random ~~process~~ is the total area under the PSD of the process i.e.

$$P = \int_{f \in F_A} \left(K - \frac{S_N(f)}{|H(f)|^2} \right) df \quad \text{—— (D)}$$

Substituting (D) in (E)

$$C \simeq \frac{1}{2} \log_2 \Delta f$$

$$C \simeq \frac{1}{2} \sum_{k=1}^{N} \Delta f \log_2 \left(K \frac{|H(f_k)|^2}{S_N(f_k)} \right),$$

$\Delta f \to 0$

then
$$C = \frac{1}{2} \int_{-\infty}^{\infty} \log_2 \left(K \frac{|H(f)|^2}{S_N(f)} \right) df$$

Que! A source produces symbols $x_1, x_2 - x_n$ with the probability of occurrence given by $P_1, P_2, \ldots P_n$. Prove that entropy of the source is atmost $\log_2 n$.

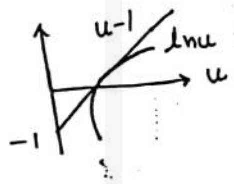
Solⁿ: Show that $H(X) - \log_2 n \leq 0$

Consider the LHS

$$H(X) - \log_2 n = \sum_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right) - \log_2 n \times 1 -$$

$$H(X) - \log_2 n = \sum_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right) - \log_2 n \sum_{i=1}^{n} p_i$$

$$= \sum_{i=1}^{n} p_i \log_2\left(1/np_i\right) = \frac{1}{\log_e 2} \sum_{i=1}^{n} p_i \log_e \frac{1}{np_i} \quad \text{--- (1)}$$



$$\ln u \leq u - 1$$

put $u = \frac{1}{np_i}$

$$\Rightarrow \ln\left(\frac{1}{np_i}\right) \leq \frac{1}{np_i} - 1 \quad \text{--- (2)}$$

Substituting eqⁿ (2) in (1), we get

$$H(X) - \log_2 n \leq \frac{1}{\log_e 2} \sum_{i=1}^{n} p_i \left(\frac{1}{np_i} - 1\right)$$

$$H(X) - \log_2 n \leq \frac{1}{\log_e 2}\left[\sum_{i=1}^{n} \frac{1}{n} - \sum_{i=1}^{n} p_i\right]$$

$$H(X) - \log_2 n \leq \frac{1}{\log_e 2}\left[n \times \frac{1}{n} - 1\right]$$

$$H(X) - \log_2 n \leq 0$$

Hence proved.

Que Prove $0 \leq H(X) \leq \log_2 n$

Solⁿ: Proof of upper bound is already given.
Proof for lower bound is.

$$p_i \leq 1$$

$$\Rightarrow \frac{1}{p_i} \geq 1$$

Taking $\log_2$ on both sides and multiplying the result by p_i

$$\sum_{i=1}^{n} p_i \log_2 (1/p_i) \geq 0$$

$$\Rightarrow H(X) = 0 \quad \text{or} \quad 0 \leq H(X)$$

Hence, entropy can not be negative.

Que Proof

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

Soln.

$$P(x_i, y_j) = P(y_j)\, P(x_i/y_j)$$

Taking summation on both sides wrt i

$$\sum_{i=1}^{M} P(x_i, y_j) = P(y_j) \sum_{i=1}^{M} P(x_i/y_j)$$

$$\sum_{i=1}^{M} P(x_i, y_j) = P(y_j) \times 1 = P(y_j) \quad \text{— (1)}$$

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

$$= -\sum_{i=1}^{M} \sum_{j=1}^{n} P(x_i, y_j) \log_2 \left[P(y_j) \cdot P(x_i/y_j) \right]$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{n} P(x_i, y_j) \log_2 P(x_i/y_j)$$

$$- \sum_{i=1}^{M} \sum_{j=1}^{n} P(x_i, y_j) \log_2 P(y_j)$$

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

$$= H(X/Y) - \sum_{j=1}^{n} P(y_j) \log_2 P(y_j)$$

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

Que Show that $I(X,Y) = I(Y,X)$

Soln:

$$I(X,Y) = \sum_{i=1}^{M}\sum_{j=1}^{n} P(x_i,y_j)\log_2\left[\frac{P(x_i|y_j)}{P(x_i)}\right] \quad \text{—} (1)$$

$$P(x_i|y_j) = \frac{P(x_i,y_j)}{P(y_j)}$$

$$P(y_j|x_i) = \frac{P(x_i,y_j)}{P(x_i)} \quad \frac{P(y_j,x_i)}{P(x_i)}$$

$$\frac{P(x_i;y_j)}{P(y_j)P(x_i)} = \frac{P(y_j,x_i)}{P(y_j)P(x_i)} = \frac{P(y_j|x_i)P(x_i)}{P(y_j)P(x_i)}$$

dividing above two equations, we get

$$\frac{P(x_i|y_j)}{P(y_j|x_i)} = \frac{P(x_i)}{P(y_j)} \quad \text{—} (2)$$

also, $P(x_i,y_j) = P(y_j,x_i) \longrightarrow (3)$

Using eqn (2) & (3) in (1)

$$I(X,Y) = \sum_{i=1}^{M}\sum_{j=1}^{n} P(y_j,x_i)\log_2\left[\frac{P(y_j/x_i)}{P(y_j)}\right]$$

$$I(X,Y) = I(Y,X)$$

Hence, proved.

Que Show that $I(X,Y) = \sum_{i=1}^{M}\sum_{j=1}^{n} P(x_i,y_j)\log_2\left[\frac{P(x_i|y_j)}{P(x_i)}\right]$

Soln:

$$P(y_j/x_i) = \frac{P(x_i,y_j)}{P(x_i)}$$

$$P(x_i,y_j) = P(y_j/x_i)\,P(x_i)$$

Taking summation on both sides w.r.t j we get

$$\sum_{j=1}^{n} P(x_i,y_j) = P(x_i)\sum_{j=1}^{n} P(y_j/x_i)$$

$$= P(x_i)\cdot 1$$

$$\sum_{j=1}^{n} P(x_i, y_j) = P(x_i) \quad — (1)$$

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

$$\Rightarrow I(X,Y) = -\sum_{i=1}^{m} P(x_i) \log_2 P(x_i) + \sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 P(x_i/y_j) \quad — (2)$$

making use of eqⁿ (1) in (2)

$$I(X,Y) = -\sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 P(x_i) + \sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 P(x_i/y_j)$$

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

$$I(X,Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 \left[\frac{P(x_i/y_j)}{P(x_i)}\right]$$

Qu show that $I(X,Y) \geq 0$ $\because$ mutual information can not be -ve.

Solⁿ: $$I(X,Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 \left[\frac{P(x_i/y_j)}{P(x_i)}\right] \quad — (1)$$

We know that,

$$P(x_i/y_j) = \frac{P(x_i, y_j)}{P(y_j)} \quad — (2)$$

Substituting (2) in (1)

$$I(X,Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_2 \left[\frac{P(x_i, y_j)}{P(x_i)P(y_j)}\right]$$

$$\Rightarrow -I(X,Y) = \frac{1}{\log_e 2}\sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i, y_j) \log_e \left[\frac{P(x_i)P(y_j)}{P(x_i, y_j)}\right] \quad — (3)$$

Scanned by CamScanner

We know that

$$\ln u \leq u-1$$

Letting $u = \dfrac{P(x_i) \cdot P(y_j)}{P(x_i, y_j)}$, we get

$$\ln \left[\frac{P(x_i) \cdot P(y_j)}{P(x_i, y_j)} \right] \leq \frac{P(x_i) \cdot P(y_j)}{P(x_i, y_j)} - 1 \quad \text{—④}$$

Substituting ④ in eqn ③, we get

$$-I(X,Y) \leq \frac{1}{\log_e 2} \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \left\{ \frac{P(x_i) P(y_j)}{P(x_i, y_j)} - 1 \right\}$$

$$-I(X,Y) \leq \frac{1}{\log_e 2} \left[\sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i) P(y_j) - \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \right]$$

$$-I(X,Y) \leq \frac{1}{\log_e 2} \left\{ \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i) P(y_j) - \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i) P(y_j/x_i) \right\}$$

$$-I(X,Y) \leq \frac{1}{\log_e 2} \left\{ \sum_{i=1}^{m} P(x_i) \sum_{j=1}^{n} P(y_j) - \sum_{i=1}^{m} P(x_i) \sum_{j=1}^{n} P(y_j/x_i) \right\}$$

$$\Rightarrow -I(X,Y) \leq \frac{1}{\log_e 2} \left\{ 1 \times 1 - 1 \times 1 \right\}$$

$$\Rightarrow -I(X,Y) \leq 0$$

or $I(X,Y) \geq 0$; Hence, proved.

Que Show that for a lossless channel, $H(X/Y) = 0$

Soln:

$$P(Y/X) = \begin{bmatrix} 1/4 & 1/4 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$P(x_1/y_1) = \frac{P(y_1/x_1)\, P(x_1)}{P(y_1)}$$

$$P(x_1|y_1) = \frac{P(y_1|x_1)\,P(x_1)}{P(x_1)\,P(y_1|x_1) + P(x_2)\,P(y_1|x_2) + P(x_3)\,P(y_1|x_3)}$$

$$P(x_1|y_1) = \frac{1/4\;P(x_1)}{P(x_1)\cdot 1/4 + 0 + 0} = 1$$

$$P(x_i|y_2) = \frac{P(y_2|x_1)\cdot P(x_1)}{P(y_2)}$$

$$= \frac{P(y_2|x_1)\cdot P(x_1)}{P(x_1)\,P(y_2|x_1) + P(x_2)\,P(y_2|x_2) + P(x_3)\,P(y_2|x_3)}$$

$$= \frac{1/4.\;P(x_1)}{P(x_1)\cdot\frac{1}{4} + 0 + 0} = 1$$

$$P(x_1|y_3) = \frac{P(y_3|x_1)\cdot P(x_1)}{P(y_3)}$$

$$= \frac{P(y_3|x_1)\;P(x_1)}{P(x_1)\cdot P(y_3|x_1) + P(x_2)\,P(y_3|x_2) + P(x_3)\,P(y_3|x_3)}$$

$$= 0$$

Summarizing for a lossless channel,
$P(x_i|y_i) = 0$ ~~for 1~~ or 1.

$$H(X/Y) = -\left[\sum_{i=1}^{m}\sum_{j=1}^{n} P(x_i|y_j)\log_2 P(x_i|y_j)\right] \quad \text{— (1)}$$

We know that

$$P(x_i|y_j) = \frac{P(x_i, y_j)}{P(y_j)}$$

$$\therefore\; P(x_i, y_j) = P(x_i|y_j)\,P(y_j) \quad \text{— (2)}$$

Substituting (2) in (1), we get

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

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

The term in the inner summation is of the form $0 \log_2 0$ or $1 \log_2 1$. hence

$$H(X/Y) = 0$$

Que For a noiseless channel show $H(X) = H(Y)$ & $H(Y/X) = 0$

Soln: Noiseless channel is characterised by a channel Matrix having 1's only along the diagonal.

$$P[Y/X] = \begin{array}{c c} & \begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ \vdots \end{array} & \left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & 1 \\ 0 & & & \end{array}\right] \end{array}$$

State Diagram

$$x_1 \xrightarrow{P(y_1/x_1)=1} y_1$$

$$x_2 \xrightarrow{P(y_2/x_2)=1} y_2$$

$$\vdots$$

$$x_n \xrightarrow{P(y_n/x_n)=1} y_n$$

Since, $P(y_j/x_i) = 1$, we get $P(y_j) = P(x_i)$ $(m = n)$

$i = 1, 2, \ldots m$; $j = 1, 2, \ldots n$

Hence
$$H(X) = \sum_{i=1}^{m} P(x_i) \log_2 1/P(x_i)$$

$$= \sum_{j=1}^{n} P(y_j) \log_2 1/P(y_j)$$

$$H(X) = H(Y)$$

For noiseless channel

$$P(y_j/x_i) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

We know that,

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

Also,

$$P(y_j/x_i) = \frac{P(x_i, y_j)}{P(x_i)}$$

$$\Rightarrow P(x_i, y_j) = P(x_i)\, P(y_j/x_i) \quad \text{— (2)}$$

Substituting (2) in (1), we get

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

$$H(Y/X) = -\sum_{i=1}^{M} P(x_i) \sum_{j=1}^{n} P(y_j/x_i) \log_2 P(y_j/x_i)$$

The term in inner summation is of form $0 \log_2 0$ or $1 \log_2 0$

$$H(Y/X) = 0$$

M-!-M

(2) Bi-Linear Transformation (one-to-one mapping)

Suppose, $H(s) = \frac{1}{s}$

$$\frac{Y(s)}{X(s)} = \frac{1}{s}$$

$$X(s) = sY(s)$$

taking inverse laplace

$$x(t) = \frac{d}{dt} y(t)$$

Integrating both sides with respect to t and taking limits from $(n-1)T$ to nT

$$\int_{(n-1)T}^{nT} x(t)\,dt = \int_{(n-1)T}^{nT} \frac{d}{dt} y(t)\,dt$$

$$y(t)\Big|_{(n-1)T}^{nT} = \int_{(n-1)T}^{nT} x(t)\,dt$$

$$\boxed{y(nT) - y(n-1)T = \frac{T}{2}\left(x(nT) + x(n-1)T\right)},$$

$$\left[\int_a^b x(t)\,dt = \frac{b-a}{2}\left(x(a) + x(b)\right)\right]$$

Trapezoidal rule

Normalizing,

$$y(n) - y(n-1) = \frac{T}{2}\left[x(n) + x(n-1)\right]$$

(1) Unit cap
Impulse
exp
Linear
Circular
(2) → DFT, IDFT
(4)(5) → sampling
(5) Butterworth low pass
(6) high pass
(7) Chebyshev low
high
(8) Inter conv
(9)

Taking z-transform

$$Y(z) - z^{-1}Y(z) = \frac{T}{2}\left[X(z) + z^{-1}X(z)\right]$$

$$\frac{Y(z)}{X(z)} = H(z) = \frac{T}{2}\left[\frac{1+z^{-1}}{1-z^{-1}}\right]$$

$$H(z) = \frac{T}{2}\left[\frac{z+1}{z-1}\right]$$

$$H(z) = \frac{1}{\frac{2}{T}\left[\frac{z-1}{z+1}\right]} \leftarrow S \quad (B)$$

Comparing (A) & (B), we get,

$$\boxed{S = \frac{2}{T}\left[\frac{z-1}{z+1}\right]}$$ Conversion formula

putting $S = \sigma + j\Omega$

$z = re^{j\omega}$

$$\sigma + j\Omega = \frac{2}{T}\left[\frac{re^{j\omega}-1}{re^{j\omega}+1}\right]$$

$$\sigma + j\Omega = \frac{2}{T}\left[\frac{r\cos\omega - 1 + jr\sin\omega}{r\cos\omega + 1 + jr\sin\omega}\right]$$

$-j\Omega$

$z = re^{j\omega}$

rationalizing

$$\sigma + j\Omega = \frac{2}{T}\left[\frac{r\cos\omega - 1 + jr\sin\omega}{r\cos\omega + 1 + jr\sin\omega}\right] \times \left[\frac{r\cos\omega + 1 - jr\sin\omega}{r\cos\omega + 1 - jr\sin\omega}\right]$$

$$\sigma + j\Omega = \frac{2}{T}\left[\frac{r^2 - 1 + 2jr\sin\omega}{r^2 + 1 + 2r\cos\omega}\right]$$

Comparing LHS & RHS.

$$\sigma = \frac{2}{T}\left[\frac{r^2 - 1}{r^2 + 1 + 2r\cos\omega}\right]$$

$\boxed{r = e^{\sigma T}}$

$\boxed{\omega = \Omega T}$

$$\Omega = \frac{2}{T}\left[\frac{2r\sin\omega}{r^2 + 1 + 2r\cos\omega}\right]$$

Conditions for preservation of Characteristics:

(1) Marginally

if, $r = 1$

then $\sigma = \frac{2}{T}\left[\frac{0}{r^2+1+2r\cos\omega}\right] = 0$

Thus, mapped successfully

(2) Unstable

$r > 1$; $(r = 2)$

$$\sigma = \frac{2}{T}\left[\frac{4-1}{r^2+1+2r\cos\omega}\right]$$

$\Rightarrow \sigma > 0$

Thus, successfully mapped.

(3) Stable

$r < 1$; $(r = 0.2)$

$$\sigma = \frac{2}{T}\left[\frac{0.04-1}{r^2+1+2r\cos\omega}\right]$$

$\Rightarrow \sigma < 0$

Thus, successfully mapped.

For making a system stable :

put $r = 1$

$$\Omega = \frac{2}{T}\left[\frac{2r\sin\omega}{r^2+1+2r\cos\omega}\right]$$

$$= \frac{2}{T}\left[\frac{2\sin\omega}{2+2\cos\omega}\right]$$

$$= \frac{2}{T}\left[\frac{\sin\omega}{1+\cos\omega}\right]$$

$$\Omega = \frac{2}{T}\left[\frac{2\sin\omega/2 \cdot \tan\omega/2}{2\cos^2\omega/2}\right]$$

$$\boxed{\Omega = \frac{2}{T}\tan\frac{\omega}{2}}$$

$\boxed{\omega = \Omega T}$

or

$$\boxed{\omega = 2\tan^{-1}\frac{\Omega T}{2}}$$

frequency conversion.

Now if $\Omega = \infty$
$\omega = \pi$

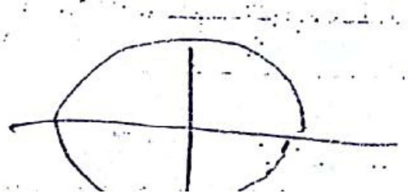
& $\Omega = -\infty$
$\omega = -\pi$

Thus, with this we obtain one to one mapping

⇒ for $-\infty$ (in analog domain) we have $-\pi$ (digital domain) that is half circle

⇒ for $+\infty$ (in analog domain) we have $+\pi$ (digital domain) that is rest of the half circle

Thus the limitation faced in IIM is overcome in this method BLT.



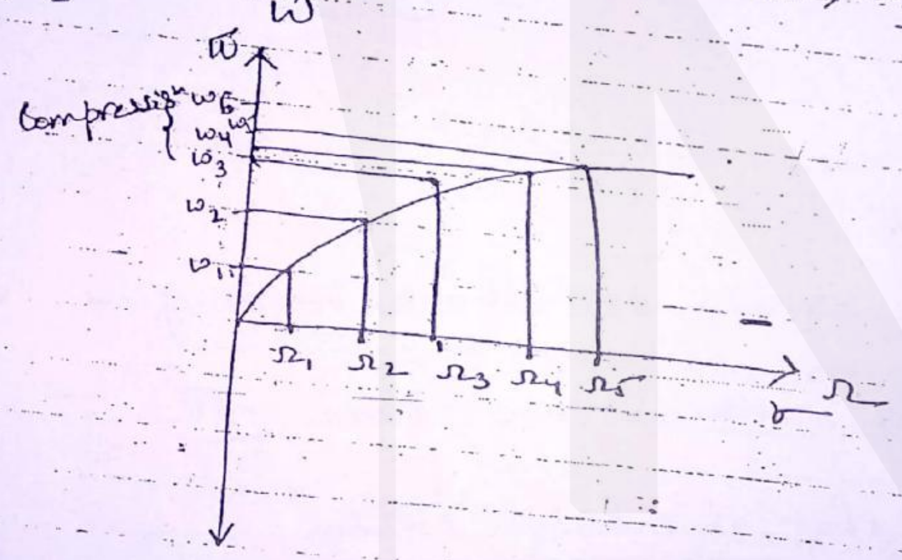
Thus, for every value in analog domain we have a value in digital domain. Thus, no overlapping of information.

Limitation of BLT.

frequency warping

$$\boxed{\omega = 2\tan^{-1}\frac{\Omega T}{2}}$$

plotting this equation, we get,



The mapping is non-linear and lower frequencies in analog domain are expanded in digital domain whereas higher frequencies are compressed. This is due to the non-linearity of the arc tangent function & is usually called as FREQUENCY WARPING

Ques Convert analog filter to a digital filter using Impulse Invariant Method, when the system function is

$$H(s) = \frac{s}{(s+1)(s+2)}, \quad T = 1\ sec.$$

Solⁿ: (1) Inverse laplace

$$\frac{s}{(s+1)(s+2)} = \frac{A}{(s+1)} + \frac{B}{(s+2)}$$ (partial fraction)

$$A = -1, \quad B = 2$$

Thus, $$H(s) = \frac{-1}{(s+1)} + \frac{2}{(s+2)}$$

$$h(t) = (-e^{-t} + 2e^{-2t})\, u(t)$$

(2) Sampling @ $t = nT$

$$h(nT) = (-e^{-nT} + 2e^{-2nT})\, u(nT)$$

(3) Z-Transform

$$H(z) = \sum_{n=-\infty}^{\infty} h(nT)\, z^{-n}$$

$$H(z) = \sum_{n=0}^{\infty} \left(e^{-nT} + 2e^{-2nT}\right) z^{-n}$$

$$= \sum_{n=0}^{\infty} e^{-nT} z^{-n} + 2\sum_{n=0}^{\infty} e^{-2nT} z^{-n}$$

$$= \frac{1}{1 - e^{-T}z^{-1}} + \frac{2}{1 - e^{-2T}z^{-1}}$$

Putting T = 1 sec

$$\boxed{H(z) = \frac{1}{1 - e^{-1}z^{-1}} + \frac{2}{1 - e^{-2}z^{-1}}}$$

Ques Convert the analog filter to digital using BLT when system function is.

$$H(s) = \frac{s}{s^2+2s+5}, \quad T = 1 \text{ sec}$$

Solⁿ:

$$H(z) = H(s)\Big|_{s = \frac{2}{T}\left[\frac{z-1}{z+1}\right]}$$

$$= \frac{\frac{2}{T}\left[\frac{z-1}{z+1}\right]}{\frac{4}{T^2}\left[\frac{z-1}{z+1}\right]^2 + \frac{4}{T}\left[\frac{z-1}{z+1}\right] + 5}$$

$$\boxed{H(z) = \frac{2z^2-2}{13z^2+2z+5}}$$

Scanned by CamScanner

Ques For the analog transfer function

$$H(s) = \frac{1}{(s+1)(s+2)}$$

determine $H(z)$ using impulse invariance Method. Assume $T = 1$ sec.

Sol^n: $H(s) = \frac{1}{(s+1)(s+2)}$

partial fractions

$$\frac{1}{(s+1)(s+2)} = \frac{A}{s+1} + \frac{B}{s+2}$$

$A = 1$
$B = -1$

$$H(s) = \frac{1}{s+1} - \frac{1}{s+2}$$

- taking inverse laplace

$$h(t) = (e^{-t} - e^{-2t})u(t)$$

- Sampling @ $t = nT$

$$h(nT) = (e^{-nT} - e^{-2nT})u(nT)$$

- taking Z-transf

$$H(z) = \sum_{n=-\infty}^{\infty} h(nT) z^{-n}$$

$$H(z) = \frac{1}{1-e^{-1}z^{-1}} - \frac{1}{1-e^{-2}z^{-1}}$$

<u>Ques</u> Determine $H(z)$ using IIM/IIT for the analog system function.

$$H(s) = \frac{1}{(s+0.5)(s^2+0.5s+2)}$$

<u>Soln</u> : partial fractions

$$\frac{1}{(s+0.5)(s^2+0.5s+2)} = \frac{A}{s+0.5} + \frac{Bs+C}{s^2+0.5s+2}$$

$$1 = A(s^2+0.5s+2) + (Bs+C)(s+0.5)$$

$$1 = As^2 + 0.5As + 2A + Bs^2 + 0.5Bs + Cs + 0.5C$$

$$A + B = 0$$

$$0.5A + 0.5B + C = 0$$

$$2A + 0.5C = 1$$

Solving the equations we get $A = 0.5$
$B = -0.5$
$C = 0$

$$\therefore H(s) = \frac{0.5}{s+0.5} - \frac{0.5s}{s^2+0.5s+2}$$

$$= \frac{0.5}{s+0.5} - 0.5\left(\frac{s}{s^2+0.5s+(0.25)^2+(2-(0.25)^2)}\right)$$ making perfect square

$$= \frac{0.5}{s+0.5} - 0.5\left[\frac{s}{(s+0.25)^2+(1.3919)^2}\right]$$

$$H(s) = \frac{0.5}{s+0.5} - 0.5\left[\frac{s+0.125 - 0.25}{(s+0.25)^2 + (1.3919)^2}\right]$$

$$= \frac{0.5}{s+0.5} - 0.5\left[\left(\frac{s+0.25}{(s+0.25)^2+(1.3919)^2}\right) - \left(\frac{0.25}{(s+0.25)^2+(1.3919)^4}\right)\right]$$

$$= \underbrace{\frac{0.5}{s+0.5}}_{\mathrm{I}} \underbrace{- 0.5\left(\frac{s+0.25}{(s+0.25)^2+(1.3919)^2}\right)}_{\mathrm{II}} + \underbrace{0.0898\left(\frac{1.3919}{(s+0.25)^2+(1.3919)^2}\right)}_{\mathrm{III}}$$

$$H(z) = \frac{1}{1-e^{-0.5T}z^{-1}} - 0.5\left[\frac{1-e^{-0.25}(\cos 1.3919)z^{-1}}{1-2e^{-0.25}(\cos 1.3919)z^{-1}+e^{-0.5}z^{-1}}\right] +$$

$$0.0898\left[\frac{e^{-0.25T}(\sin 1.3919)z^{-1}}{1-2e^{-0.25}(\cos 1.3919)z^{-1}+e^{-0.5}z^{-2}}\right]$$

By using formula

$$\boxed{\begin{aligned} \frac{s+a}{(s+a)^2+b^2} &\rightarrow \frac{1-e^{-aT}(\cos bT)z^{-1}}{1-2e^{-aT}(\cos bT)z^{-1}+e^{-2aT}z^{-2}} \\ \frac{b}{(s+a)^2+b^2} &\rightarrow \frac{e^{-aT}(\sin bT)z^{-1}}{1-2e^{-aT}(\cos bT)z^{-1}+e^{-2aT}z^{-2}} \end{aligned}}$$

$$\frac{1-e^{-aT}\cos bT\, z^{-1}}{1-2e^{-aT}\cos bT + e^{-2aT}z^{-2}}$$

$$\frac{e^{-aT}\sin bT\, z^{-1}}{1-2e^{-aT}\cos bT z^{-1} + e^{-2aT}z^{-2}}$$

Ques. Convert the analog filter with system function

$$H(s) = \frac{s+0.1}{(s+0.1)^2+9}$$

into a digital filter using bilinear Transformation. $T = 0.276$ sec.

Sol^n: $H(z) = H(s)\Big|_{s=\frac{2}{T}\left[\frac{z-1}{z+1}\right]}$

$$H(z) = \frac{\frac{2}{T}\left[\frac{z-1}{z+1}\right]+0.1}{\left(\frac{2}{T}\left[\frac{z-1}{z+1}\right]+0.1\right)^2+9}$$

$$H(z) = \frac{1+0.027z^{-1}-0.973z^{-2}}{8.572-11.84z^{-1}+8.177z^{-2}}$$

Ques. Apply BLT to $H(s) = \frac{2}{(s+1)(s+3)}$ with $T = 0.1s$.

Sol^n: $H(z) = H(s)\Big|_{s=\frac{2}{T}\left[\frac{z-1}{z+1}\right]}$

$$H(z) = \frac{2}{\left(\frac{2}{T}\left[\frac{z-1}{z+1}\right]+1\right)\left(\frac{2}{T}\left[\frac{z-1}{z+1}\right]+3\right)}$$

$$H(z) = \frac{2}{(21z-19)(23z-17)}$$

Scanned by CamScanner