Information Theory & Coding
(5CS3-01)

Unit-1&2 Notes

Vision of the Institute

To become a renowned center of outcome based learning and work towards academic, professional, cultural and social enrichment of the lives of individuals and communities.

Mission of the Institute

M1- Focus on evaluation of learning outcomes and motivate students to inculcate research aptitude by project based learning.
M2- Identify, based on informed perception of Indian, regional and global needs, the areas of focus and provide platform to gain knowledge and solutions.
M3- Offer opportunities for interaction between academia and industry.
M4- Develop human potential to its fullest extent so that intellectually capable and imaginatively gifted leaders can emerge in a range of professions.

Vision of the Department

To become renowned Centre of excellence in computer science and engineering and make competent engineers & professionals with high ethical values prepared for lifelong learning.

Mission of the Department

M1-To impart outcome based education for emerging technologies in the field of computer science and engineering.
M2-To provide opportunities for interaction between academia and industry.
M3- To provide platform for lifelong learning by accepting the change in technologies
M4- To develop aptitude of fulfilling social responsibilities.

Program Outcomes (PO)

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis**: Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/development of solutions**: Design solutions for complex engineering problemsand design system components or processes that meet thespecified needs with appropriate consideration for the public health and safety, andthe cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issuesand the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Educational Objectives (PEO)

1. To provide students with the fundamentals of Engineering Sciences with more emphasis in **Computer Science &Engineering** by way of analyzing and exploiting engineering challenges.
2. To train students with good scientific and engineering knowledge so as to comprehend, analyze, design, and create novel products and solutions for the real life problems.
3. To inculcate professional and ethical attitude, effective communication skills, teamwork skills, multidisciplinary approach, entrepreneurial thinking and an ability to relate engineering issues with social issues.
4. To provide students with an academic environment aware of excellence, leadership, written ethical codes and guidelines, and the self-motivated life-long learning needed for a successful professional career.
5. To prepare students to excel in Industry and Higher education by Educating Students along with High moral values and Knowledge

Program Specific Outcomes (PSO)

PSO1: Ability to interpret and analyze network specific and cyber security issues, automation in real word environment.

PSO2: Ability to Design and Develop Mobile and Web-based applications under realistic constraints.

Course Outcome:

CO1: Apply the fundamental concepts of information theory viz. entropy, mutual information and channel capacity in communication system.

CO2: Examine the principles of source coding and data transmission.

CO3: Analyze linear block code, cyclic code and Convolution code.

CO4: Evaluate information theoretic methods to novel settings of encoding and decoding techniques.

CO-PO Mapping:

CO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO1 0	PO1 1	PO1 2
Apply the fundamental concepts of information theory viz. entropy, mutual information and channel capacity in communication system.	3	2	2	2	1	1	1	1	1	1	1	3
Examine the principles of source coding and data transmission.	3	3	3	3	2	1	1	1	1	1	1	3
Analyze linear block code, cyclic code and Convolution code.	3	3	3	2	1	1	1	1	1	1	1	3
Evaluate information theoretic methods to novel settings of encoding and decoding techniques.	3	3	3	2	1	1	1	1	1	1	1	3

SYLLABUS:

RAJASTHAN TECHNICAL UNIVERSITY, KOTA

Syllabus

III Year-V Semester: B.Tech. Computer Science and Engineering

5CS3-01: Information Theory & Coding

Credit: 2 | **Max. Marks: 100(IA:20, ETE:80)**

2L+0T+0P | **End Term Exam: 2 Hours**

SN	Contents	Hours
1	**Introduction:**Objective, scope and outcome of the course.	01
2	**Introduction to information theory:** Uncertainty, Information and Entropy, Information measures for continuous random variables, source coding theorem. Discrete Memory less channels, Mutual information, Conditional entropy.	05
3	**Source coding schemes for data compaction:** Prefix code, Huffman code, Shanon-Fane code &Hempel-Ziv coding channel capacity. Channel coding theorem. Shannon limit.	05
4	**Linear Block Code:** Introduction to error connecting codes, coding & decoding of linear block code, minimum distance consideration, conversion of non-systematic form of matrices into systematic form.	05
5	**Cyclic Code:** Code Algebra, Basic properties of Galois fields (GF) polynomial operations over Galois fields, generating cyclic code by generating polynomial, parity check polynomial. Encoder & decoder for cyclic codes.	06
6	**Convolutional Code:** Convolutional encoders of different rates. Code Tree, Trllis and state diagram. Maximum likelihood decoding of convolutional code: The viterbi Algorithm fee distance of a convolutional code.	06
	Total	**28**

LECTURE PLAN:

Unit No./ Total lec. Req.	Topics	Lect. Req.
	Objective, Scope & Outcome of the Course	1
Unit-1	Introduction to information theory, Uncertainty, Entropy	1
	Information measures for continuous random variables	1
	Numerical problem on entropy	1
	Source coding theorem, Discrete memory less channels	1
	Mutual information, Conditional entropy	1
Unit-2	Prefix code, Huffman coding	1
	Shannon – fanon coding	1
	Numerical on haffman and shanon fano coding	1
	Hempel-Ziv coding	1
	Channel capacity, Channel coding theorem, Shannon limit	1
Unit-3	Introduction to error correcting codes	1
	Coding and decoding of linear block code	1
	Numerical problem on Linear block code	1
	Error correcting codes, Minimum distance consideration	1
	Conversion of non symmetric form of matrix into symmetric form	1
Unit-4	Code algebra	1
	Basic properties of Galois Field(GF)	1
	Polynomial operation over Galois field	1
	Generating cyclic code by generating polynomial	1
	Numerical Problems on generator polynomial	1
	Parity check polynomial , Encoder and decoder for cyclic codes	1
Unit-5	Convolutional encoders of different rates	1
	Code tree	1
	Trellis diagram	1
	state diagram	1
	Maximum likelihood decoding of convolution code	1
	Viterbi algorithm, Free distance of convolution codes	1

Information theory & Coding:-

Information theory:- is branch of probability theory, which can be applied to the study of the communication system.

* Communication system deals with the flow of some sort of information in some work.

I/P Message $\rightarrow$ I/P Transducer $\xrightarrow{\text{input Signal}}$ Transmitter $\xrightarrow{\text{Transmitted Signal}}$ $\boxed{\text{Channel}}$ $\xrightarrow{\text{Received Signal}}$ Receiver $\xrightarrow{\text{O/P Signal}}$ O/P Transducer $\rightarrow$ O/P Message.

noise

Information:-

Measurement of information:-

Let us suppose that a discrete information source emits m possible ~~diff.~~ messages $m_1, m_2, \ldots m_m$ with probability of occurrences $P_1, P_2, \ldots P_m$ where $P_1 + P_2 + \ldots P_m = 1$.

Let us assume that in a long time interval, L message have been generated. L is very large $L >> m$.

then, the number of messages

$$m_1 = P_1 L.$$

The amount of information in message $\boxed{m_1, I_{m_1} = \log\left(\frac{1}{P_1}\right)}$

Thus the total information in all m_1 message $= P_1 L \log\left(\frac{1}{P_1}\right)$

The total amount of information in all L message will then be

$$I_t = P_1 L \cdot \log\left(\frac{1}{P_1}\right) + P_2 L \log\left(\frac{1}{P_2}\right) + \cdots + P_m L \log\left(\frac{1}{P_m}\right)$$

The information content or amount of information in the k^{th} message denoted by I_K, must be inversely related to P_K.

i) $I(m_K) = 0$ as $P_K = \infty$

Q: A source produces one of four possible message during each interval having probabilities $P_1 = 1/2$, $P_2 = 1/4$, $P_3 = 1/8$, $P_4 = 1/8$. Obtain the information content of each of these messages.

2

Information content

$$I(m_k) = \log_2\left(\frac{1}{P_k}\right) \text{ bits}$$

$I(m_1) = \log_2(2) = 1$ bit.

$I(m_2) = \log_2(4) = \log_2(2)^2 = 2$ bits

$I(m_3) = \log_2(8) = \log_2(2)^3 = 3$ bits

$I(m_4) = \log_2(8) = \log_2(2)^3 = 3$ bits

Q: Calculate the amount of information if it is given that $P_K = \frac{1}{4}$.

amount of information $I(m_k) = \log_2\left(\frac{1}{P_k}\right)$

$$I(m_k) = \log_2(2)^2 \Rightarrow 2 \text{ bits}$$

$$I(m_k) = \frac{\log_{10}\left(\frac{1}{P_k}\right)}{\log_{10}(2)}$$

$$I(m_k) = \frac{\log_{10}(4)}{\log_{10}(2)} = \cancel{2 \text{ bits}} \Rightarrow \frac{\log_{10} 2^2}{\log_{10} 2} = 2 \text{ bits}$$

$$\log_2 4 = \frac{\log_{10} 4}{\log_{10} 2}$$

$\log_a b \Rightarrow \log_{10} b$.

$$\log_a b = \frac{\log_{10} b}{\log_{10} a}$$

Entropy :- Average information is termed as entropy.

Suppose we have M different and independent messages $m_1, m_2 \ldots$ with probabilities of occurrence $P_1, P_2 \ldots$ suppose for a long period of transmission of sequence of L messages have been generated thus if L is very large that we expect that in L message sequence we transmit $P_1 L$ messages of m_1, $P_2 L$ messages of m_2 etc

The I_{total} in such a sequence will be

$$I_{total} = P_1 L \log_2 \frac{1}{P_1} + P_2 L \log_2 \frac{1}{P_2} \cdots$$

The average information per message ~~interval~~ interval represented by symbol H will be

$$H = \frac{I_{total}}{L}$$

$$H = P_1 \log_2 \left(\frac{1}{P_1}\right) + P_2 \log_2 \left(\frac{1}{P_2}\right) + \cdots$$

$$H = \sum_{K=1}^{M} P_K \log_2 \left(\frac{1}{P_K}\right)$$

for example we have two messages with probabilities P and (1-P) Then avg information per message interval is

$$H = P \log_2 \left(\frac{1}{P}\right) + (1-P) \log_2 \left(\frac{1}{1-P}\right)$$

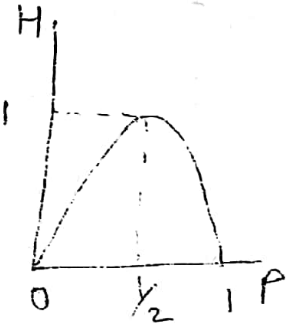
→ ~~Psychology:-~~

Average Information, Entropy :-

$$H = \sum_{K=1}^{M} P_K \log_2 \left(\frac{1}{P_K}\right)$$

for example we have two messages with probabilities P and (1-P). Then average information per message interval is

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



A plot of H as a function of P as shown as

As shown H=0 at P=0 and P=1. The ~~the~~ maximum value of H may be located by setting the value of $\frac{dH}{dp} = 0$

$$\because H = P\log_2 \frac{1}{P} + (1-P)\log_2 \frac{1}{1-P} \quad \text{—(a)}$$

Differentiating eqn (a) with respect to P but before this as we know that

$$\log_2 x = \frac{\log_e x}{\log_e 2}$$

after applying this equation (a) look like

$$H = p\frac{\log_e \frac{1}{P}}{\log_e 2} + (1-P)\frac{\log_e\left(\frac{1}{1-P}\right)}{\log_e 2}$$

$$H = \frac{1}{\log_e 2}\left[P\log_e \frac{1}{P} + (1-P)\log_e \frac{1}{1-P}\right] \quad \text{—(b)}$$

~~Now differentiating~~

Scanned by CamScanner

Now differentiating

$$H = \frac{-1}{\log_e 2}\left[P \times \log_e P + (1-P)\log_e(1-P)\right]$$

$$\frac{dH}{dP} = \frac{-1}{\log_e 2}\left[P \times \frac{1}{P} + \log_e P + (1-P) \times \frac{-1}{(1-P)} + \log_e (1-P)(-1)\right]$$

$$\Rightarrow \frac{-1}{\log_e(2)}\left[1 + \log_e P - 1 - \log_e(1-P)\right]$$

$$\Rightarrow \frac{-1}{\log_e(2)}\left[\log_e P - \log_e(1-P)\right]$$

$$\Rightarrow -\log_2 P + \log_2(1-P)$$

$\log \frac{M}{n} \Rightarrow \log M - \log n$

$$\Rightarrow \log_2\left(\frac{1-P}{P}\right)$$

To find out the maximum value of P, equate $\frac{dH}{dP}$ to zero

$$\therefore \log_2\left(\frac{1-P}{P}\right) = 0$$

$$\frac{1-P}{P} = 2^0$$

$$\Rightarrow \frac{1-P}{P} = 1$$

$\Rightarrow 1 - P = P$ is $2P = 1 \Rightarrow P = \frac{1}{2}$

Hence maximum value of H is

$$H_{max} = \frac{1}{2}\log_2 2 + \frac{1}{2}\log_2 2$$

$0 \le H(x) \le \log_2 M$

$H_{max} = \log_2 2 = 1$ bit/message

A source emits an independent sequence of symbol from an alphabet consisting of 5 symbols A, B, C, D, E with Probabilities $\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{3}{16}, \frac{5}{16}$ respectively. Find entropy of source?

$$H = \sum_{i=1}^{4} P_i \log_2 \frac{1}{P_i}$$

$$H = \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8) + \frac{3}{16}\log_2\left(\frac{16}{3}\right) + \frac{5}{16}\log_2\left(\frac{16}{5}\right)$$

$$\Rightarrow \frac{1}{2} + \frac{3}{8} + \frac{3}{8} +$$

$$\Rightarrow .5 + .375 + .375 + .4520 + .52$$

$$H = 2.22 \text{ bit/Symbol}$$

A black & white TV Picture consists of 525 line of Picture information. Assume that each ~~Picture~~ line consist of 525 Picture element & each element have 256 brightness level. Pictures are Repeated at the rate of 30/sec. Calculate the average rate of information content by TV set to viewer.

Soln: Total Picture element = 525×525

Now rate of 525×525 elements are $R = 525 \times 525 \times 30/\text{sec}$

$= 8268750$ elements/sec

Now $H = \sum_{i=1}^{m} \frac{1}{256} \log_2 256$

Now $H = +\sum_{i=1}^{256} \frac{1}{256} \log_2 256$

Because here information are transmitted i.e. element have 256 brightness levels.

i.e total no of message = 256.

equiprobable & each levels are

$\therefore 256P = 1$

$P = \frac{1}{256}$

Sum of all Probability of all messages are equal to 1.

$H = \log_2 256$

Information Rate (R): if a source generates messages at the rate of r message per second, the rate of information R is defined as the average number of bits of information per second.

$R = r\, H(X)$

$$R = \frac{\text{message}}{\text{sec}} \times \frac{\text{bits}}{\text{message}} = \frac{\text{bits}}{\text{sec}}$$

$H = 8$ bit/element

(7)

now information Rate

$R = rH$

$r = \frac{\text{messages}}{\text{sec}}$, $H = \frac{\text{bits}}{\text{message}}$

$R = 8268750 \times 8$

$R = 66.15 \times 10^6$ bit/sec

$R = \frac{\text{bits}}{\text{sec}}$

$R = 6.552 \times 10^6$ bit/sec.

Q. A discrete Source emits one of 5 symbols once every millisecond. The symbol Probabilities are $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}$ respectively. Calculate entropy & rate of information?

$$H = \sum_{i=1}^{q} P_i \log_2 \frac{1}{P_i}$$

$$H = \sum_{i=1}^{5} P_i \log_2 \frac{1}{P_i}$$

$R = rh$

$$H = \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{16}\log_2(16) + \frac{1}{16}\log_2(16)$$

$H = \frac{15}{8}$ bits/symbol or 1.875 bits/sym

~~$R = \frac{bits}{r}$~~ R = Information Rate

$r = \frac{1}{10^{-3}}$ symbol/sec

$R = rh$

$\frac{\text{Symbol}}{\text{sec}} \times \frac{\text{bits}}{\text{Symbol}}$

$R = \frac{15}{8} \times 10^3$ bits/sec.

Q. Consider a Discrete Memoryless Source with the alphabet α (S_0, S_1, S_2) with probabilities $P_0 = \frac{1}{4}$, $P_1 = \frac{1}{4}$, $P_2 = \frac{1}{2}$. find out the entropy of orignal source & second order extension entropy?

1st Entropy of orignal source $H(\alpha) = \sum_{i=0}^{2} P_i \log_2 \frac{1}{P_i}$

$$= \frac{1}{4}\log_2(4) + \frac{1}{2}\log_2(2) + \frac{1}{2}\log_2(2)$$

$\Rightarrow \frac{3}{2}$ bits/alphabets symbols or 1.5

$\frac{1}{10^{-3}}$

$\frac{r}{10^{-3}}$

Second order extension entropy:-

Source alphabet $\alpha = (S_0, S_1, S_2)$ consist of three symbols. (8)

$K =$ no of symbols $= 3$

Second order extension source symbols $= K^2 = 3^2 = 9$

These symbols are

$\sigma_0 \; \sigma_1 \; \sigma_2 \; \sigma_3 \; \sigma_4 \; \sigma_5 \; \sigma_6 \; \sigma_7 \; \sigma_8$

$= (S_0S_1, S_0S_0, S_0S_2, S_1S_0, S_1S_1, S_1S_2, S_2S_0, S_2S_1, S_2S_2)$

Probabilities of these are

$P(\sigma_i) = \left(\frac{1}{16}, \frac{1}{16}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{4}\right)$

Entropy of second order extension

$$H(\alpha^2) = \sum_{i=0}^{8} P(\sigma_i) \log_2\left(\frac{1}{P}\right)(\sigma_i)$$

$$H(\alpha^2) = \frac{1}{16}\log_2(16) + \frac{1}{16}\log_2(16) + \frac{1}{8}\log_2(8) + \frac{1}{16}\log_2(16)$$

$$+ \frac{1}{16}\log_2(16) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8)$$

$$+ \frac{1}{4}\log_2(4)$$

$= 3$ bits/sym Ans.

Cross check ~~by~~ the ans by formula

$H(\alpha^n) = n\,H(\alpha)$

$H(\alpha^2) = 2H(\alpha)$

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

Q. The output of an information source consist of 128 syms, 16 of which occurs with a probability of $1/32$ & remaining occur with a prob of $1/224$. The source emits 1000 sym/sec. Assume that symbs are chosen independently find R, $H = \frac{1}{2}(5) + \frac{1}{2}(7.808)$, $H = 6.404$, $R = rH = 6404$ bits/sec.

Q. Consider a DMS with source alphabate $\alpha = (S_0, S_1, S_2, S_3)$ (9)
with Prob $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$. findout Second & third order extension entropy.

$K^2 \Rightarrow \boxed{K \text{ no of sym}}$ $4^2 = 16$

Soln $\alpha^2 = S_0S_0 \; S_0S_1 \; S_0S_2 \; S_0S_3 \; S_1S_0 \; S_1S_1 \; \ldots \; S_3S_0 \; S_3S_1 \; S_3S_2 \; S_3S_3$

total no of symbols $= 4^2 = 16$

Probability of each of 16 blocks of 2 symbols are $\frac{1}{16}$

$$H(\alpha^2) = \sum_{i=1}^{16} P_i \log_2 \frac{1}{P_i}$$

$$H(\alpha^2) = \sum_{i=1}^{16} \frac{1}{16} \log_2(16) = \log_2 16 = 4 \text{ bits/symbol}$$

Now third order extension

α^3 (block of 3 symbols) $= (S_0 S_1 S_2 \,.\, S_0 S_1 S_3 \ldots)$

No of block $= 4^3 = 64$

$$H(\alpha^3) = \sum_{i=1}^{64} \frac{1}{64} \log_2(64) = 6 \text{ bits/symbol}$$

$H(\alpha)$ = first order extension

$$\sum_{i=1}^{4} \frac{1}{4} \log_2(4) = 2 \text{ bits/symbol}$$

we can crosscheck by $H(\alpha^n) = n H(\alpha)$ Ans

Properties of Entropy :-

The entropy $H(\mathcal{L})$ of DMS is bounded as

$$0 \leq H(\mathcal{L}) \leq \log_2 K$$

where K is No of symbols of the alphabet $\mathcal{L}$ of source

1) $H(\mathcal{L}) = 0$, if and only if the Probability $P_K = 1$ for some ~~and~~ K, and the remaing Probabilities in the set are all zero. this is lower bound of ano entropy.

2) $H(\mathcal{L}) = \log_2 K$, if and only if $P_K = 1/K$ for all K (i.e all the symbols in the alphabet $\mathcal{L}$ are equiprobable). This is upper

CODING: ⊗ used to improve the transmission efficiency of ~~a~~ the communication system.

⊗ used for secrecy or minimum probability of error

Source ⟶ channel ⟶ Receiver

without encoder - decoder

Source ⟶ encode ⟶ channel ⟶ Decoder ⟶ Receiver

with encoder - decoder

terminology in coding

1) Letter, character & Symbol → any individual member of the alphabet set

ii) Message or word: A finite sequence of letters of an alphabet.

iii) length of the word:- Number of letters in a message.

(C_1 0 C_2 01 C_3 001)

Suppose $C_1 = 0$, $C_2 = 10$ and $C_3 = 110$

This code is irreducible as an addition of a 0 or a 1 to any of the codewords does not produce other code words.

if we receive 0110101000101101010

$C_1 C_3 C_2 C_2 C_1 C_1 C_2 C_3 C_3 C_2$

it can be said that then ~~for~~ a code is irreducible, it is also uniquely decode.

Coding efficiency:-

Let M be the number of symbols in an encoding alphabet.

There messages $[m_1, m_2, m_3 --- m_N]$ with probabilities $[P(m_1), P(m_2) --- P(m_N)]$.

Let n_i be the number of symbols in the i^{th} message.

The average length of message or avg length per code word is then given by

$$\bar{L} = \sum_{i=1}^{N} n_i P(n_i) \text{ letters/message}$$

$\bar{L}$ ~~Should~~ Should be minimum to have an efficient transmission coding efficiency, then can be defined as $\eta = \frac{\bar{L}_{min}}{\bar{L}}$, $\eta = \frac{H}{N\bar{L}}$

$\eta = \bar{L}_{min}$ $\eta = \frac{H}{...}$

Entropy Coding :-

A) Shannon – fano Coding :-

1 *) List the source symbols in order of decreasing Probability.

2 *) Partition the set into two sets that are close to equiprobable as possible and assign 0 to upper set and one to lower set.

3 *) Continue this Process, each time partitioning the sets with as nearly equal probabilities as possible until further partitioning is not possible.

	P						Code	No of bits
x_1	.30	0	0			2	00	2
x_2	.25	0	1			2	01	2
x_3	.20	1	0			2	10	2
x_4	.12	1	1	0		3	110	3
x_5	.08	1	1	1	0	4	1110	4
x_6	.05	1	1	1	1	4	1111	4

$N = 2(.30) + 2(.25) + 2(.20) + 3(.12) + 4(.08) + 4(0.05)$

$\Rightarrow .60 + .50 + .40 + .36 + .32 + .20$

$\Rightarrow 2.38$ bits / Symbol

$$\boxed{\log_2 x = \frac{\log_{10} x}{\log_{10} 2}}$$

$$H = \sum_{i=0}^{5} P_i \log_2 \frac{1}{P_i}$$

$$= .30 \log_2 \frac{1}{.30} + .25 \log_2 \frac{1}{.25} + .20 \log_2 \frac{1}{.20} + .12 \log_2 \frac{1}{.12} + .08 \log_2 \frac{1}{.08} + .05 \log_2 \frac{1}{.05}$$

$\frac{1}{.30} \Rightarrow \log \Rightarrow / .3010 \Rightarrow \times .30$

$H = 2.36$ bits / symbol

$$\eta = \frac{H(x)}{N \log_2 m} = .99$$

B) Huffman Encoding :-

1 *) List the source symbols in order of decreasing Probability.

2 *) Combine the Probabilities of two symbols (messages) having the lowest probabilities and ~~reordered~~ reorder the resultent probabilities. this step is called reduction 1.

3 *) the same Procedure is repeated until there are two ordered

4) Start encoding with last reduction, which consists of exactly two ordered probabilities. Assign 0 as the first digit in the codewords for all the source symbols associated with the first probability, assign 1 to the second probability.

5) Now go back and assign 0 and 1 to the second digit for the two probabilities that were combined in the previous reduction step retaining all assignment made in step 4.

6) Keep regressing this way until the first column is reached

Q. A message source generates eight message symbols $m_1, m_2 \ldots m_8$ with probabilities 0.25, 0.03, 0.19, 0.16, 0.11, 0.14, 0.08, 0.04 respectively. Give the Huffman codes for these symbols. Calculate the entropy of the source and the average number of bits per symbol.

message symbols	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
Probabilities	0.25	0.03	0.19	0.16	0.11	0.14	0.08	0.04

m_1	0.25_{01}	0.25_{01}	0.25_{01}	$.25_{01}$	$.31_{00}$	$.44_{(1)}$	$.56_{(0)}$
m_3	0.19_{11}	$.19_{11}$	$.19_{11}$	$.25_{10}$	$.25_{01}$	$.31_{(00)}$	$.44_{(1)}$
m_4	0.16_{000}	$.16_{000}$	$.16_{000}$	$.19_{11}$	$.25_{10}$	$.25_{(01)}$	
m_6	0.14_{100}	$.14_{100}$	$.15_{001}$	$.16_{000}$	$.19_{11}$		
m_5	0.11_{101}	$.11_{101}$	$.14_{100}$	$.15_{001}$			
m_7	0.08_{0010}	$.08_{0010}$	$.11_{101}$				
m_8	0.04_{00110}	0.7_{0011}					
m_2	0.03_{00111}						

.25 .19

.25 .50

$\boxed{1\frac{1}{2}, 1\frac{1}{4}}$

.0664
.1171
.1857
.1857
.1857
.1857
.2464
.2461
.3322
.3322
.3322
.4453
.4464

m_1	.25	01	2 ✓
m_2	.03	00111	5
m_3	.19	11	2 ✓
m_4	.16	000	3 ✓
m_5	.11	101	3 ✓
m_6	.14	100	3 ✓
m_7	.08	0010	4 ✓
m_8	.04	00110	5

Source entropy $H = \sum_{i=0}^{7} P_i \log\left(\frac{1}{P_i}\right)$

$\Rightarrow .25 \log\frac{1}{.25} + .03 \log\frac{1}{.03} + .19 \log\frac{1}{.19} + .16 \log\frac{1}{.16} + .11 \log\frac{1}{.11}$
$+ .14 \log\frac{1}{.14} + .08 \log\frac{1}{.08} + .04 \log\frac{1}{.04}$

$\Rightarrow 2.75$ bits / symbols

Avg Number of bits per symbols

$= .25 \times 2 + 0.03 \times 5 + .19 \times 2 + .16 \times 3 + .11 \times 3 + .14 \times 3 + .08 \times 4$
$+ .04 \times 5 =$

$\Rightarrow .50 + .15 + .38 + .48 + .33 + .42 + .32 + .20$

$\Rightarrow 2.78$ bits / symbols

efficiency $\eta = \frac{H}{N}$

$\eta = \frac{2.75}{2.78} = 98.92\ \%$

<u>Discrete Memoryless Source</u> :- Discrete source is termed as memoryless if the successive symbols emitted by the source are statistically independent.

Q. A DMS has 5 equally likely messages. Construct Shannon's Fano code for DMS and calculate efficiency of code? [14]

(a) choose 0.4 versus 0.6 set

(b) choose 0.6 versus 0.4 set

(a) First group's sum of Probability must be equal to .4 and second group's sum of Probability = 0.6 at the starting stage.

message	Prob				No of bits
m_0	.2	0	0		2
m_1	.2	0	1		2
m_2	.2	1	0		2
m_3	.2	1	1	0	3
m_4	.2	1	1	1	3

sum of all message prob = 1

5 equally likely messages.

$\therefore 5p = 1 \quad p = 1/5 = .2$

$N = .2(2) + .2(2) + .2(2) + .2(3) + .2(3)$

$\Rightarrow .4 + .4 + .4 + .6 + .6$

$\Rightarrow 1.2 + 1.2 \Rightarrow 2.4$ bits/message

$$H = \sum_{i=0}^{4} P_i \log \frac{1}{P_i}$$

$$H = \frac{1}{5}\log 5 + \frac{1}{5}\log 5 + \frac{1}{5}\log 5 + \frac{1}{5}\log 5 + \frac{1}{5}\log 5$$

$\Rightarrow \log_2 5 = 2.32$ bits/symbol

efficiency $\Rightarrow \eta = \frac{H}{N} = \frac{2.32}{2.4} = .96$

b)

message	Prob				No of bits
M_0	.2	0	0	0	3
M_1	.2	0	0	0	3
M_2	.2	0	1	0	2
M_3	.2	1	0		2
	.2				

.2	0	0	0	= 3
.2	0	0	1	= 3
.2	0	1		= 2
.2	1	0		2
.2	1	1		2

$N = 2.4$ bits/symbol

$H = 2.32$ bits/Symbol

$e = .96$

Q. Apply the Shannon-fano coding Procedure for the given message.

$[X] = [x_1 \; x_2 \; x_3 \; x_4 \; x_5 \; x_6 \; x_7]$

$[P] = [0.4 \; 0.2 \; 0.12 \; 0.08 \; 0.08 \; 0.08 \; 0.04]$

Assume $M = 2$, & $M = 3$

Solⁿ for the first Partitioning there are two ways :-

[1] $[X_1] = [x_1, x_2]$, $[X_2] = [x_3, x_4, x_5, x_6, x_7]$

[2] $[X_1] = [x_1]$, $[X_2] = [x_2, x_3, x_4, x_5, x_6, x_7]$

2]

Message	Prob	Encoded Message	Code word length
x_1	.4	0	1
x_2	.2	1 0 0	3
x_3	.12	1 0 1	3
x_4	.08	1 1 0 0	4
x_5	.08	1 1 0 1	4
x_6	.08	1 1 1 0	4
x_7	.04	1 1 1 1	4

$\frac{H}{N}$

$N = .4\times1 + .2\times3 + .12\times3 + .08\times4 + .08\times4 + .08\times4 + .04\times4$

$\Rightarrow .4 + .6 + .36 + .32 + .32 + .32 + .16$

$\Rightarrow 2.48.$

[1]

Message	Prob	Encode Message	Code word length
x_1	.4	0 0	2
x_2	.2	0 1	2
x_3	.12	1 0 0	3
x_4	.08	1 0 1	3
x_5	.08	1 1 0	3
x_6	.08	1 1 1 0	4
x_7	.04	1 1 1 1	4

$$\bar{N} = \sum_{k=1}^{7} = (.4\times 2)+(.2\times 2)+(.12\times 3)+(.08\times 3)+(.08\times 3)$$
$$+(.08\times 4)+(.08\times 4)$$

$\bar{N} = 2.52$ Letters / Message.

it prove that Second Method is better as it provides a lower value of $\bar{N}$.

$$\text{Entropy } H(x) = -\sum_{k=1}^{7} P_K \log \frac{1}{P_K}$$

$$\Rightarrow \sum_{k=1}^{7} .4\log\frac{1}{.4} + .2\log\frac{1}{.2} + .12\log\frac{1}{.12} + .08\log\frac{1}{.08} + .08\log\frac{1}{.08}$$
$$+ .08\log\frac{1}{.08} + .04\log\frac{1}{.04}.$$

$H(x) \Rightarrow 2.48$ bits/Message

$$\eta = \cancel{\frac{2.42}{2.48}} \quad \frac{H(x)}{\bar{N}\log M} = \frac{2.42}{2.48\log 2} = 97.6\%$$

-1, 0, 1
0, 1, 2

Case M = 3

Let the encoding alphabet be = -1 0 1

Message	Probability	Encoded Message			No of bits
x_1	.4	-1			1
x_2	.2	0	-1		2
x_3	.12	0	0		2
x_4	.08	1	-1		2
x_5	.08	1	0		2
x_6	.08	1	1	-1	3
x_7	.04	1	1	0	3

$$\bar{N} = \sum_{K=1}^{7} P_K n_K = [(0.4\times1) + (0.2\times2) + (0.12\times2) + (0.08\times2) + (0.08\times2) + (0.08\times3) + (0.04\times3)]$$

$\Rightarrow$ 1.72 letters / message

Hence efficiency $\eta = \frac{H(X)}{\bar{N}\log M} = \frac{2.42}{1.72\log_2 3} = 88.7\%$

SHANNON'S THEOREM / CHANNEL CAPACITY THEOREM:-

This theorem is concerned with rate of transmission of information over a communication channel.

Q: $x_1 -- x_7$

Prob .46, .30, .12, .06, .03, .02, .01

$H(x) = 1.9781$ bits/message

$N = 1.99$ $\eta = .9940$

Information Source:- an information source may be viewed as an object which produces an event.

[An ~~Practical~~ information source in a communication system is a device which produces messages and it can be either analog or discrete. analog sources can be transformed to discrete sources through the use of sampling and quantization techniques.

A discrete information source is a source which has only a finite set of symbols as possible outputs. The set of source symbols is called the source alphabet and the elements of the set are called symbols or letters.

Classification of Information source:-

information source can be classified as having memory or being memoryless.

A source with memory is one for which a current symbol depends on the previous symbols, a memoryless source is one for which each symbol produced is independent of the previous symbol.

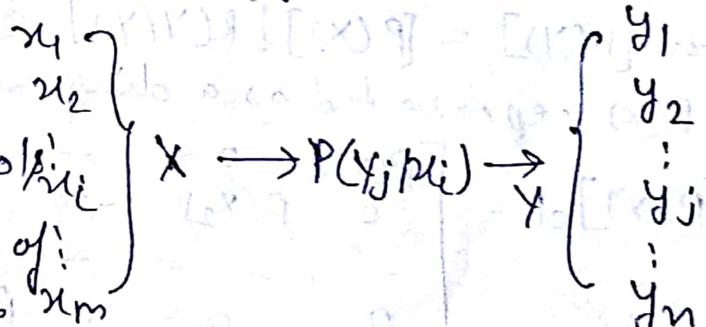
A Discrete memoryless source can be characterized by the list of symbols which are independent of the previous symbols.

DISCRETE MEMORYLESS CHANNELS (DMC):-

(1) **Channel Representation:-** A communication channel may be defined as the path or medium through which the symbols flow to the receiver end.

A DMC is a statistical model with an input X and and output Y. DMC accepts an input symbol from X and in responce it generates an output symbol from Y. the channel is said to be discrete when the alphabets of X and Y are both finite. and it is said to be memoryless when the current output depend on only the current input and not on any of previous inputs.

a diagram of a DMC with m inputs and n outputs. The input X consist of input symbols $x_1, x_2 -- x_m$. The probabilities of these source symbols $P(x_i)$ are assumed to be known.



The output Y consist of output symbols $y_1, y_2 -- y_n$. Each possible input to output path is indicated along with a conditional probability $P(Y_j/X_i)$, when $P(Y_j/X_i)$ is the conditional probability of obtaing output Y_j given that the input is X_i, and is called a channel transition probability.

(b) Channel Matrix:- a channel is completely specified by the complete set of transition Probabilities. Accordingly the channel in fig. is often specified by the matrix of transition Probabilities $[P(Y/X)]$. This matrix is given by

$$[P(Y/X)] = \begin{bmatrix} P(Y_1/X_1) & P(Y_2/X_1) & \cdots & P(Y_n/X_1) \\ P(Y_1/X_2) & P(Y_2/X_2) & \cdots & P(Y_n/X_2) \\ \vdots & & & \vdots \\ P(Y_1/X_m) & P(Y_2/X_m) & \cdots & P(Y_n/X_m) \end{bmatrix}$$

Channel Probability or transition Probability

This matrix $[P(Y/X)]$ is called the channel matrix. Since each input to the channel results some output, each row of channel matrix must sum to unity.

This means that

$$\sum_{j=1}^{n} P(y_j/x_i) = 1 \quad \text{for all } i.$$

if input Probabilities $P(X)$ are represented by row matrix

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

and output Probabilities $P(Y)$ are represented by row matrix as $[P(Y)] = [P(Y_1)\ P(Y_2)\ \cdots\ P(Y_n)]$

then $[P(Y)] = [P(X)][P(Y/X)]$

if $P(X)$ represented as a diagonal Matrix, then we have

$$[P(X)]_d = \begin{bmatrix} P(X_1) & 0 & \cdots & 0 \\ 0 & P(X_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & P(X_m) \end{bmatrix}$$

Joint probability

then $[P(X,Y)] = [P(X)]_d\,[P(Y/X)]$

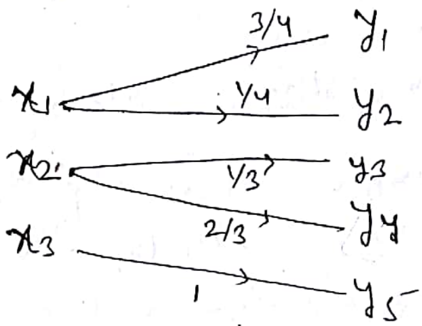
where the (i,j) element of matrix $[P(X,Y)]$ has the form $P(x_i, y_j)$. The matrix $[P(X,Y)]$ is known as the joint Probability matrix. and the element $P(x_i, y_j)$ is the joint Probabilities of transmitting x_i and receiving y_j

20

Type's of channel

Lossless channel:- whose channel matrix has only one non-zero element in each column.

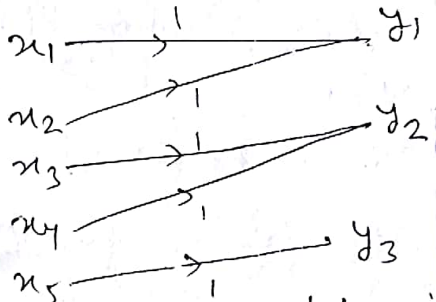
$$[P(Y|X)] = \begin{array}{c|ccccc} & y_1 & y_2 & y_3 & y_4 & y_5 \\ x_1 & \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 \\ x_2 & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ x_3 & 0 & 0 & 0 & 0 & 1 \end{array}$$



no source information is lost in transmission.

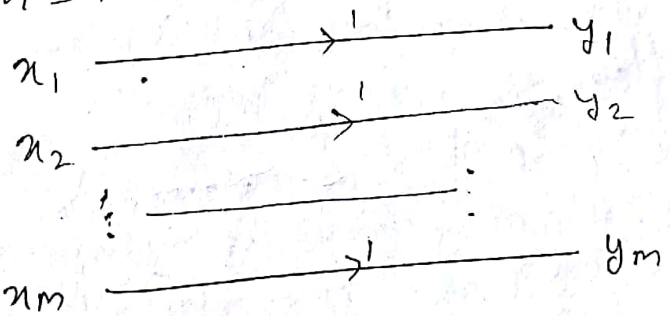
Deterministic channel:- whose channel matrix has one non zero element in each row.

$$[P(Y|X)] = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ x_1 & 1 & 0 & 0 \\ x_2 & 1 & 0 & 0 \\ x_3 & 0 & 1 & 0 \\ x_4 & 0 & 1 & 0 \\ x_5 & 0 & 0 & 1 \end{array}$$



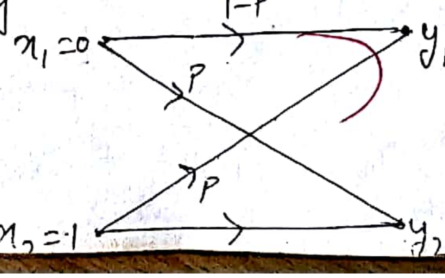
when a given source symbol is sent in the deterministic channel it is clear which output symbol will be received.

Noiseless channel:- A channel is called noiseless if it is both lossless and deterministic. The channel matrix has only one element in each row and in each column and this element is unity. the input and the output are of the same size. that is $m = n$ for noiseless channel.



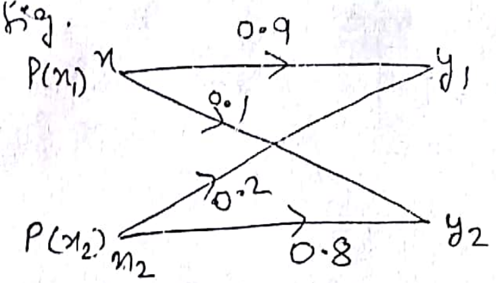
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Binary Symmetric channel (BSC) :- a BSC has two inputs ($x_1 = 0$, $x_2 = 1$) and two outputs ($y_1 = 0$, $y_2 = 1$). This channel is symmetric because the probability of receiving a 1 if a 0 is sent is the same as the probability of receiving a 0 if a 1 is sent. This common transition probability is denoted by P.



$$[P(Y|X)] = \begin{bmatrix} 1-P & P \\ P & 1-P \end{bmatrix}$$

Q. Given a Binary channel shown in fig.



(i) find the channel matrix of the channel

Ans we know that channel matrix is given by

$$[P(Y|X)] = \begin{bmatrix} P(y_1|x_1) & P(y_2|x_1) \\ P(y_1|x_2) & P(y_2|x_2) \end{bmatrix}$$

Substituting all the given values, we get

$$[P(Y|X)] = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

(ii) find $P(y_1)$ and $P(y_2)$ when $P(x_1) = P(x_2) = 0.5$

Ans we know that

$$[P(Y)] = [P(X)]\,[P(Y|X)]$$

Substituting all the values, we get

$$[P(Y)] = [0.5 \quad 0.5] \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$\Rightarrow [P(Y)] = [.55 \quad .45] \Rightarrow [P(y_1) \quad P(y_2)]$$

$P(y_1) = .55$ and $P(y_2) = .45$

(iii) find the joint Probabilities $P(x_1, y_2)$ and $P(x_2, y_1)$ when $P(x_1) = P(x_2) = 0.5$

$$[P(X,Y)] = [P(X)]_d\,[P(Y|X)]$$

$$[P(X,Y)] = \cancel{[P(X)]_d} \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$[P(X,Y)] = \begin{bmatrix} .45 & .05 \\ .1 & .4 \end{bmatrix} \quad \text{——— (1)}$$

standard matrix form

$$[P(X,Y)] = \begin{bmatrix} P(x_1, y_1) & P(x_1, y_2) \\ P(x_2, y_1) & P(x_2, y_2) \end{bmatrix} \quad \text{— (2)}$$

Comparing (1) & (2)

$P(x_1, y_2) = .05$ and $P(x_2, y_1) = .1$ Ans

THE CONDITIONAL AND JOINT ENTROPIES :-

Using the input Probabilities $P(x_i)$, output Probabilities $P(y_j)$, transition probabilities $P(y_j|x_i)$, and the joint probabilities $P(x_i, y_j)$, Let us define the following various entropy functions for a channel with m inputs and n outputs

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

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

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

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

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

H(X) is the average uncertainty of the channel input and H(Y) is the average uncertainty of the channel output.

The conditional entropy H(X|Y) is a measure of the average uncertainty remaining about the channel input after the channel output has been observed.

H(X|Y) is sometimes called the equivocation of X with respect to Y. The conditional entropy H(Y|X) is the average uncertainty of the channel output given that X was transmitted.

The joint entropy H(X,Y) is the average uncertainty of the communication channel as a whole. two useful relationships among the above various entropies are as under:

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

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

MUTUAL INFORMATION:-

The mutual information denoted by $I(X;Y)$ of a channel is defined by

$$I(X;Y) = H(X) - H(X|Y) \quad b/\text{Symbol}$$

Since $H(X)$ represents the uncertainty about the channel input before the channel output is observed and $H(X|Y)$ represents the uncertainty about the channel input after the channel output is observed, the mutual information $I(X;Y)$ represents the uncertainty about the channel input that is resolved by observing the channel output.

Property of Mutual Information $I(X;Y)$

i) $I(X;Y) = I(Y;X)$

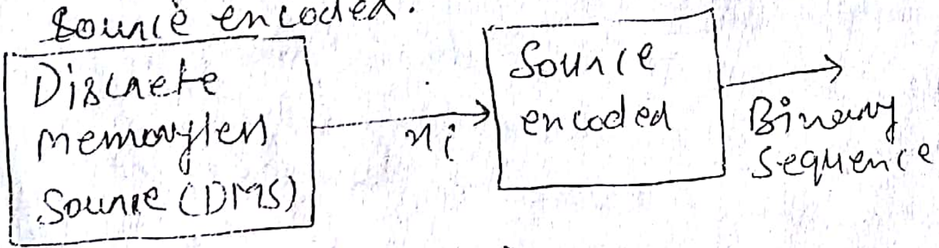
ii) $I(X;Y) \geq 0$

iii) $I(X;Y) = H(Y) - H(Y|X)$

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

THE SOURCE CODING:-

(1) Definition :- A conversion of the output of a discrete memoryless source into a sequence of binary symbols (i.e binary code-word) is called source coding. The device that perform this conversion is called source encoder.



$X = (x_1, x_2, x_3 \ldots x_i, \ldots x_m)$

(2) Objective of Source Coding :- ~~an~~ an objective of source coding is to minimize the average bit rate required for representation of the source by reducing the redundancy of the information source.

Few Terms related to Source Coding Process :-

(i) Codeword length :- Let X be a DMS with finite entropy H(X) and an alphabet $\{x_1, \ldots x_m\}$ with corresponding Probabilities of occurrence $P(x_i)$ ($i = 1, \ldots m$). Let the binary codeword assigned to symbol x_i by the encoder have length n_i, measured in bits. The length of a codeword is the number of binary digit in the codeword.

(ii) Average codeword length :- The average codeword length L, per source symbol is given by

$$L = \sum_{i=1}^{m} P(x_i)\, n_i$$

the Parameter L represents the average number of bits Per source symbol used in the source coding Process.

(iii) Code efficiency :- code efficiency η is defined as

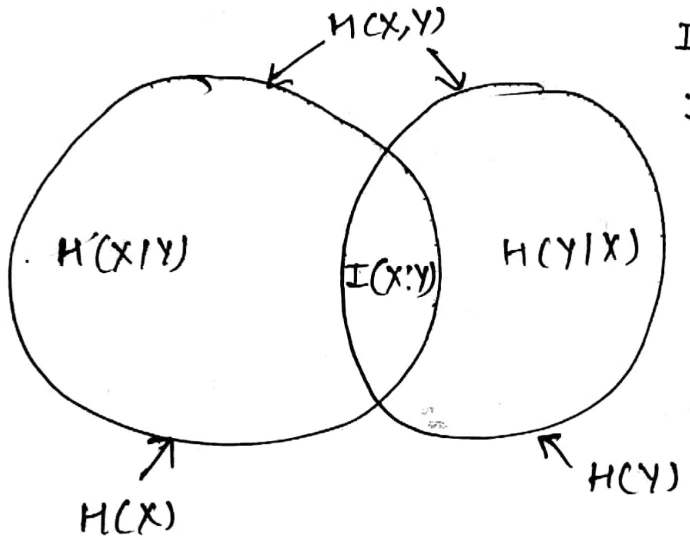
$$\eta = \frac{L_{min}}{L}$$

where L_{min} is the minimum Possible value of L. when η approaches unity the code is said to be efficient.

(iv) Code Redundancy :- Code Redundancy γ is defined as $\gamma = 1 - \eta$

Source coding theorem :- The source coding theorem

Mutual info & entropy! -



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

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

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

$I(X;Y) = I(Y;X)$

Scanned by CamScanner

Source Coding theorem :- the source coding theorem states that for a DMS X, with entropy H(X), the average codeword length L per symbol is bounded as

$$L \geq H(X)$$

and further, L can be made as close to H(X) as desired for some suitably chosen code. thus, with $L_{min} = H(X)$, the code efficiency can be rewritten as

$$\eta = \frac{H(X)}{L}$$

Classification of code :-

x_i	Code 1	Code 2	Code 3	Code 4	Code 5	Code 6
x_1	00	00	0	0	0	1
x_2	01	01	1	10	01	01
x_3	00	10	00	110	011	001
x_4	11	11	11	111	0111	0001

fixed length code :- A fixed length code is one whose codeword length is fixed. Code 1 and code 2 are fixed length codeword with length 2.

Variable length codes :- a variable length code is one whose codeword length is not fixed. all codes of table except codes 1 and 2 are variable length codes.

Distinct Codes :- if Prefix free then uniquely decodable
if uniquely decodable not necessary prefix code.

A code is distinct if each codeword is distinguishable from other code words. All codes of table except code 1 are distinct codes - notice the codes for x_1 and x_3

Prefix - free codes :- A code in which no codeword can be formed by adding code symbols to another codeword is called prefix-free code. Thus, in a prefix-free code no codeword is a prefix of another. Code 2, 4, 6 are prefix free code

Uniquely decodable codes:-

~~if the~~ A distinct code uniquely decodable if the original source sequence can be reconstructed perfectly from the encoded binary sequence.

Code 3 of the table is not a uniquely decodable code. for example the binary sequence 1001 may correspond ~~is~~ to the source sequence $x_2, x_3 x_2$ or $x_2 x_1 x_1 x_2$.

A sufficient condition to ensure that a code is uniquely decodable is that no code word is a prefix of another. Thus, the prefix codes 2, 4 and 6 are uniquely decodable codes.

Note that the prefix free condition is not a necessary condition for unique decodability. for example code 5 of table does not satisfy the condition ~~for unique decodability~~ and yet it is uniquely decodable since the bit 0 indicates the beginning of each codeword of the code.

Instantaneous codes:-

A uniquely decodable code is called an instantaneous code if the end of any codeword is recognizable without examining subsequent code symbols.

in instantaneous code no codeword is a prefix of another codeword. for this reason prefix free codes are sometimes known as instantaneous code.

Optimal code:-

A code is said to be optimal if it is instantaneous and has minimum average L for given source with a given probability assignment for the source symbols.

if prefix free $\xrightarrow{\text{then}}$ uniquely decodable
if uniquely decod. $\xrightarrow{\text{not necess.}}$ prefix free

Kraft Inequality:-

Let X be a DMS with alphabet $\{x_i\}$ $(i = 1, 2, \ldots M)$. Assume that the length of the assigned binary codeword corresponding to x_i is n_i.

a necessary and sufficient condition for the existence of an instantaneous binary code is

$$K = \sum_{i=1}^{M} 2^{-n_i} \leq 1$$

which is known as the Kraft inequality.

Q. Consider a DMS X with symbols $x_i, i = 1,2,3,4$. Table lists four possible binary codes.

x_i	Code A	Code B	Code C	Code D
x_1	00	0	0	0
x_2	01	10	11	100
x_3	10	11	100	110
x_4	11	110	110	111

i) Show that all the codes except code B satisfy the Kraft inequality.

ii) Show that code A & D are uniquely decodable but code B and C are not uniquely decodable.

Soln i) for code A $\quad n_1 = n_2 = n_3 = n_4 = 2$

therefore $K = \sum_{i=1}^{4} 2^{-n_i} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$

for code B $\quad n_1 = 1, n_2 = n_3 = 2, n_4 = 3$

therefore $K = \sum_{i=1}^{4} 2^{-n_i} = \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} = 1\frac{1}{8} > 1$

for code C $\quad n_1 = 1, n_2 = 2, n_3 = n_4 = 3$

$K = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = 1$

for code D $\quad n_1 = 1, n_2 = n_3 = n_4 = 3$

$K = \frac{1}{2} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{7}{8} < 1$

So all codes except B satisfy the Kraft inequality.

Soln ii)

Lempel Ziv Coding! - The Lempel - Ziv coding is used for Lossless data compression. The logic behind Lempel Ziv universal coding is ~~as follows~~. The compression of an arbitrary sequence of bits is possible by coding a series of 0's and 1's as some previous such string (the prefix string) plus one new bit. Then the new string formed by adding the new bit to the previously used prefix string becomes a potential prefix string for future string. These variable length blocks are called phrases. These phrases are listed in dictionary which stores the existing phrases and their locations.

Dictionary coding!-

Dictionary coding techniques rely upon the observation that there are correlations between parts of data (recurring patterns). The basic idea is to Replace those repetitions by shorter references to a "dictionary" containing the orignal.

Encode the Sequence of Symbols [a b aab a bb bbbb b abbbbba]

Encoded Packet	<0,a>	<0,b>	<1,a>	<2,a>	<2,b>	<5,b>	<5,a>	<6,b> <4>
address	1	2	3	4	5	6	7	8
Content	a	b	aa	ba	bb	bbb	bba	Bbbb

<dictionary address, next data character>

Encode the sequence of Symbols [a b aab a bbb bbbb abbbbba]

Encoded Packet	<0, 0, a>	<0,0,b>	<2,1, a>	<3,2, b>	<6,5,a>	<1, 7, a>
Address	a	b	aa	bab	bbbbba	bbbbbbba
Running text	a	ab	abaa	abaabab	wholeseq.	abaababbbbbbba

Channel Capacity :-

Shannon introduced the concept of channel capacity, which is the limit at which data can be transmitted through a medium. The errors in the transmission medium depends on the energy of the signal, the energy of the noise and the bandwidth of the chennel. if the bandwidth is high, we can ~~pump~~ transmitt more data in the channel. if the signal energy is high the effect of noise is reduced.

According to shannon the bandwidth of the channel and signal energy and noise energy are related by formula

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

where C is channel capacity in bits/sec.

B is bandwidth of channel in Hz. (Hertz)

$\frac{S}{N}$ is signal to noise power ratio (SNR). SNR generally measured in dB using the formula

$$\left(\frac{S}{N}\right)dB = 10 \log\left(\frac{\text{Signal Power}}{\text{Noise Power}}\right)$$

The value of the channel capacity obtained using this formula is the theoretical maximum. as an example noiseless channel has infinite capacity. if there is no noise in the channel then $N=0$, Hence $\frac{S}{N} = \infty$. Such channel is called noiseless channel. Then capacity of such channel will be

$$C = B \log_2(1+\infty) = \infty$$

Q. A gaussian channel has 1 MHz bandwidth. Calculate the channel capacity if the signal power to noise spectrum density ratio is 10^5 Hz. ~~also find the maximum information rates.~~

Soln

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

$$= B \log\left(1+\frac{S}{\eta B}\right)$$

N = noise power

η = noise spectrum density

S = Signal Power

$\because \eta B$ for a Gaussian channel

$$= 10^6 \log\left(1+\frac{10^5}{10^6}\right) \quad \left[\because \frac{S}{\eta} = 10^5,\ B = 10^6 Hz\right]$$

$\simeq 137503.52$

$\simeq 138000$ bits/sec

The maximum information rate is

$$R_{max} = \frac{1.44 \frac{S}{\eta N}}{}$$

$= 1.44 \times 10^5 = 144000$ bits/sec.

Q. Calculate the minimum bandwidth required to transmit Picture Signal with following data.

a) ~~The~~

Q. A Television Transmission requires 30 frames of 3,00,000 Pictures elements each to be transmitted per second. Estimate the Theoritical bandwidth of the AWGN channel if the SNR at the receiver is required to be at least 50dB. each of the elements can assume 10 brightness levels with equal probability.

Soln Information Per picture frame = 9.96×10^5 bits