Unit 1 - Information Theory.

Measure of Information:

$S = \{S_1, S_2, S_3 \ldots S_q\}$ $P = \{P_1, P_2 \ldots P_q\}$

Amount of information or Self information

$$\boxed{I_k = \log \frac{1}{P_k}}$$

1. Binary symbols 0 & 1 are transmitted with probabilities 1/4 & 3/4 respectively. Find the corresponding self information

→ $I_0 = \log \frac{1}{P_0} = \frac{\log 1/4}{\log 2} = \log_2 4 = 2$ bits

Self information

$$I_1 = \log \frac{1}{P_1} = \frac{\log 3/4}{\log 2} = \log_2 (4/3) = 0.413 \text{ bits}$$

- I cannot be -ve.
- Lowest possible self information is zero. ($\because P_k = 1$)
- I_k & I_l → be two different information

  $I_{kl} = I_k + I_l$

Zero memory source:-

Memoryless. Any source that generates discrete symbols but does not have memory is called memoryless source.

Average Information (Entropy) for a long independent sequence

$S = \{S_1, S_2 \ldots S_q\}$

$P = \{P_1, P_2 \ldots P_q\}$

length of sequence → L symbols

$P_1 L$ no. of messages of type S_1

$P_2 L$ — " — S_2

$\vdots$

$P_q L$ — " — S_q

Self information $I_k = \log \frac{1}{P_k}$

$P_1 L$ no. of message of type S_1 contains $P_1 L \log \frac{1}{P_1}$ bits of information.

$P_2 L$ — " — S_2 $\quad P_2 L \log \frac{1}{P_2}$

$\vdots$

$P_q L$ — " — $\quad P_q L \log \frac{1}{P_q}$

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

$$\boxed{I_{total} = L \sum_{i=1}^{q} P_i \log \frac{1}{P_i}}$$

Average Information $= \frac{I_{total}}{L}$

Entropy $\boxed{H(S) = \sum_{i=1}^{q} P_i \log \frac{1}{P_i}}$ bits/message symbol.

10/2/24

Entropy (ig) uncertainty: Average (~~uncertainty~~) amount of uncertainty / surprises.

1 Let us take source alphabet $S = \{S_1, S_2\}$ with the probability $P = \{1/256, 255/256\}$

→ $H(S) = \sum_{i=1}^{q} P_i \log \frac{1}{P_i}$

$= \frac{1}{256} \log_2(256) + \frac{255}{256} \log_2 \frac{256}{255}$

$= 0.0368$ bits/message symbol.

2 $S: \{S_3, S_4\}$ $\quad P' = \{7/16, 9/16\}$

→ $H(S) = \frac{7}{16} \log_2\left(\frac{16}{7}\right) + \frac{9}{16} \log_2\left(\frac{16}{9}\right)$

$= 0.988$ bits/msg symbol.

3. $S: \{S_5, S_6\}$ $\quad P' = \{1/2, 1/2\}$

$H(S) = 1$ → uncertainty is maximum.

Information rate (R_s):-

Product of avg information content per symbol and message symbol (r_s).

$\boxed{R_s = H(S)\, r_s}$ bits/sec

4. The o/p of info. source consists of 150 symbols, 32 of which occurs with a prob. of 1/64 & the remaining 118 occurs with a prob. of 1/236. The source emits 2000 sym/sec. Assume that symbols are chosen independently. Find the average info rate of the source.

→ $H(S) = 32 \times \frac{1}{64} \log(64) + 118 \times \frac{1}{236} \log(236)$

$= 6.94$ bits/message symbol

$R(S) = H(S) \cdot r(S)$

$= 6.94 \times 2000$

$= 13882$ bits/sec

$= 13.88$ kbits/sec

5 A card is drawn from a deck

(i) You are told it is a spade

(ii) How much info. did you receive if are told the card drawn is an ace.

(iii) If u are told, card drawn is ace of spades. How much info is received.

classmate

Date
Page

(iv) Is the info. obtained in (ii) the sum of info. obtained from (i) & (iii)

→ (i) $P = 1/4$

$\cancel{H(S) = \frac{1}{4} \log_2(4) = 1/2 \text{ bits/symbol}}$

$I = \log_2 \frac{1}{P} = 2$ bits

(ii) $P_k = 4/52 = 1/13$

$I_k = \log_2 13 = 3.700$ bits

(iii) $P = 1/52$

$I = \log_2 \frac{1}{P} = 5.7$ bits

(iv) $I_{k_1} + I_{k_2} = 2 + 3.700$

$= 5.700$ bits

$= I_{k_3}$

6. Discrete source emmits one of the 6 symbols once every ms. The symb probs are $1/2$, $1/4$, $1/8$, $1/16$, $1/32$, $1/32$ respectively. Find source entropy and information rate.

→ $H(S) = \sum_{i=1}^{6} P \log(1/P)$

$= 1/2 \log_2(2) + 1/4 \log_2(4) + 1/8 \log_2(8) + 1/16 \log_2(16) + \frac{2}{32} \log_2(32)$

$H(S) = 1.9375$

Find information rate:

$R = \cancel{\log_2(2) + \log_2(4) + \log_2(8) + \log_2 16 + 2\log_2(32)}$

$R_s = r_s \times H(S)$

$= 1000 \times 1.9375 \times 10^3$

$= \cancel{1.9375 \text{ Kbps}}$ 1.9375 Kbps

classmate

Date
Page

7. Find relation b/w Hartley, nats, bits.

$I = \log_{10} \frac{1}{P}$ Hartley

$I = \log_e \frac{1}{P}$ nats

$I = \log_2 \frac{1}{P}$ bits

H to b

$\log_2 10 = 3.321$

$1 H = 3.321$ bits

→ $1 \text{ Hartley} = \frac{I}{\log_{10} 1/P}$

$= \frac{\log_e 1/P}{\log_{10} 1/P} \Rightarrow \frac{-\log_e P}{-\log_{10} P}$

$= \log_e 10$

$= \frac{\log_e 10}{\log_e e}$

$\boxed{\because \log_a b = \frac{1}{\log_b a}}$

$\boxed{1 \text{ Hartley} = \log_e 10 = 2.303 \text{ nats}}$

|||ly

$1 \text{ nats} = \log_2 e = 1.443$ bits

$1 \text{ Hartley} = \log_2 10 = 3.322$ bits.

8. A code is composed of dots & dashes. Assuming dash is 3 times as long as dots & has $1/3^{rd}$ the probality of occurrence. Calculate
(i) The information in a dot & dash
(ii) Entropy of · – code
(iii) The average rate of information if a dot lasts for 10 ms & this time is allowed b/w symbols

→ $I_{dot} = \log_2(\cancel{3/4})\ 4/3 = \cancel{0.584}\ 0.415$

$I_{dash} = \log_2(4) = \cancel{1.584}\ 2$

$I = I_{dot} + I_{dash} = \cancel{2.169}\ 2.415$

$P_{dot} + P_{dash} = 1$

$P_{dot} + \frac{1}{3} P_{dot} = 1$

$P_{dot} = 3/4$

$P_{dash} = 1/4$

(ii) $H(S) = \sum_{i=1}^{2} p \log_2 (1/p)$

$= \frac{3}{4} \log_2 \left(\frac{4}{3}\right) + \frac{1}{4} \log_2 (4)$

$= 0.811$ bits/sym

(iii) $R_s = H(S)\, r_s$

$R_{dot} = 0.811 \times \frac{1}{10 \times 10^{-3}}$

$= 81.1$ bps

$R_{dash} = 0.811 \times \frac{1}{30 \times 10^{-3}}$

$= 27.03$ bps

$R_{s(avg)} = \frac{R_{dot} + R_{dash}}{2} = 54.065$ X

(iii) Prob of dot $= 3/4$ ∴ In 4 symbols, 3 will be dots & 1 will be dash

Prob of dash $= 1/4$.

$R(S) = H(S) \cdot r_s$

$= 0.811 \times 40$

$= 32.44$ bits/sec.

$\therefore \underset{\text{dash}}{30} + \underbrace{10 + 10 + 10}_{\text{dots}} + \underbrace{10 + 10 + 10 + 10}_{\text{delays in b/w symbols}}$

$\Rightarrow 100$ ms

$r_s = \frac{4}{100\text{ms}} = 40$ sym/sec.

9. Find the entropy of a source in nats per symbol of a source that emits 1 out 4 symbols a, b, c & d in a statistical independent sequence with probs 1/2, 1/4, 1/8, 1/8

→ $H(S) = \sum_{i=1}^{4} p \log_e 1/p$

$= \frac{1}{2} \log_e 2 + \frac{1}{4} \log_e 4 + \frac{2}{8} \log_e 8 = 1.213$ bi nats/sym

$\log_2 \rightarrow 1.75$ b/sym

1.75

1 Hartley $= 2.303$ nats

? ← 1.213

$\frac{1.213}{2.303} = 0.526$ (bits) Hartley

1 bit $= 0.693$ nats.

$1.75 = ?$ $= 1.213$ nats

1 bit $= 0.301$ Hart

$1.75 = ?$ $= 0.526$ Hartley

Properties of Entropy :-

1. Probability is b/w 0 & 1 (exclusive)
2. Entropy is symmetric function of its argument

$H[P_k, (1-P_k)] = H[(1-P_k), P_k]$

S_A, S_B, S_C

$P_A = \{P_1, P_2, P_3\}$

$P_B = \{P_2, P_1, P_3\}$ ∴ $H(S_A) = H(S_B) = H(S_C)$

$P_C = \{P_3, P_1, P_2\}$

3. Extremal property:

$H(S)_{max} = \log q$ bits/sym

where q = No. of symbols

4. Properties of Additivity:

$H'(S) \geq H(S)$

$S = \{S_1, S_2, S_3\}$

$P = \{P_1, P_2, P_3\}$

Sub group $\Rightarrow H'(S)$

where $H'(S)$ should be greater than =

5. Source efficiency:

$\boxed{\eta_s = \frac{H(S)}{H(S)_{max}}}$

Source Redundancy

$\boxed{R_{\eta s} = 1 - \eta_s}$

10. A discrete message source is S emmits two independant symbol if the prob is 0.55 & 0.45 respectively. Calculate the efficiency of the source & redundancy.

→

$$H(S) = \sum_{i=1}^{q} P \log \frac{1}{P}$$

$$= 0.55 \log_2 \frac{1}{0.55} + 0.45 \log_2 \frac{1}{0.45}$$

$$= 0.9927 \text{ bits/msg sym}$$

$$\eta_s = \frac{H(S)}{H(S)_{max}} = \frac{0.9927}{1} \times 100 = \underline{\underline{99.27\%}}$$

$$R_{\eta s} = 1 - \eta_s = 1 - 0.992 = 0.008 = \underline{0.8\%}$$

11. A pair of dice are tossed simultaneously. The outcome of the first die is recorded as x_1 & that of second die as x_2. Two events are defined as follows
$A = \{(x_1, x_2)$ such that $x_1 + x_2 \leq 7\}$
$B = \{(x_1, x_2)$ such that $x_1 > x_2\}$ which event has more info.
Support your answer by numerical computation.

→

S { (1,1) (1,2) (1,3) (1,4) (1,5) (1,6)
(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)
(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)
(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)
(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)
(6,1) (6,2) (6,3) (6,4) (6,5) (6,6) }

$P_A = 21/36$ $\qquad$ $P_B = 15/36$

$I_A = \log 1/P$ $\qquad$ $I_B = \log(1/P)$

$= 0.777$ $\qquad$ $= 1.26$

Event B conveys more information

12. Shortly before a horse race, book maker believes that several horses entered the race at have the following probability of winning $\boxed{A: 0.04 \quad B: 0.42 \quad C: 0.31 \quad D: 0.12 \quad E: 0.11}$
He then receives a message that owing to a minor injury one of the horses is not participating in the race. Explain how would you assess from a information theory point of view, the information value of this message
(i) If the horse in question is known
(ii) If it is not known

→ (i) $I_A = 4.64$

$I_B = 1.25$

$I_C = 1.689$

$I_D = 3.05$

$I_E = 3.18$

(ii) $I_{total} = \frac{\Sigma I}{5} = 13.809$

$H(S) = 0.04 \times 4.64 + 0.42 \times 1.25 + 0.31 \times 1.689 + 0.12 \times 3.05 + 0.11 \times 3.18$

$= 1.85$ bits/msg sym

13. Black & white TV picture consists of 525 lines of picture information. Assume that each line consists of 525 picture elements & that each element can have 256 brightness levels. Picture are repeated at the rate of 30 frames/second. Calculate the avg rate of information conveyed by a TV set to a viewer.

→ 525 lines

525 pixels

256 brightness levels

$q = (256)^{525 \times 525}$

$H(max) = \log q$

$=$

Extension of zero-memory source:-

$S = \{S_1, S_2\}$

$P = \{P_1, P_2\}$ where $P_1 + P_2 = 1$

2^{nd} extension :- extension

[no. of source symbol]

$[2]^2 = 4$

S_1S_1 occurs with prob $P_1P_1 = P_1^2$

S_1S_2 — " — P_1P_2

S_2S_1 — " — $P_2P_1 = P_2P_1$

S_2S_2 — " — $P_2P_2 = P_2^2$

$$P_1^2 + 2P_1P_2 + P_2^2 = 1$$

$H(S) = \sum_{i=1}^{2} P_i \log 1/P_i$

$= P_1 \log_2 1/P_1 + P_2 \log_2 1/P_2$

$H(S^2) = \sum_{i=1}^{4} P_i \log_2 1/P_i$

$P_1^2 \log 1/P_1^2 + P_1P_2 \log_2 1/P_1P_2 + P_2P_1 \log_2 1/P_2P_1 + P_2^2 \log_2 1/P_2^2$

$= 2P_1^2 \log_2 1/P_1 + P_1P_2 \log_2 1/P_1P_2 + P_1P_2 \log_2 1/P_1P_2 + 2P_2^2 \log_2 1/P_2$

$= 2P_1^2 \log_2 1/P_1 + 2P_1P_2 \log_2 \frac{1}{P_1} + 2P_1P_2 \log_2 1/P_2 + 2P_2^2 \log_2 1/P_2$ add

$= 2P_1(P_1+P_2) \log_2 1/P_1 + 2P_2(P_1+P_2) \log_2 1/P_2$ $(\because P_1+P_2=1)$

$= 2[P_1 \log_2 1/P_1 + P_2 \log_2 1/P_2]$

$\boxed{H(S^2) = 2H(S)}$

Third extension :-

$[2^3] = 8$

$S_1S_1S_1$ occurs with prob $P_1P_1P_1 = P_1^3$

$S_1S_1S_2$ — " — $P_1P_1P_2 = P_1^2P_2$

$S_1S_2S_1$ — " — $P_1P_2P_1 = P_1^2P_2$

$S_1S_2S_2$ — " — $P_1P_2P_2 = P_1P_2^2$

$S_2S_1S_1$ — " — $P_2P_1P_1 = P_1^2P_2$

$S_2S_1S_2$ — " — $P_2P_1P_2 = P_1P_2^2$

$S_2S_2S_1$ — " — $P_2P_2P_1 = P_1P_2^2$

$S_2S_2S_2$ — " — $P_2P_2P_2 = P_2^3$

$$P_1^3 + 3P_1^2P_2 + 3P_1P_2^2 + P_2^3 = 1$$

$H(S^3) = \sum_{i=1}^{8} P_i \log_2 1/P_i$

Finally, on simplification

$H(S^3) = 3H(S)$

$\therefore$ for n^{th} extension,

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

14. A zero memory source has source alphabets $S = (S_1, S_2, S_3)$ with prob. $P = \{1/2, 1/4, 1/4\}$. Find the entropy of the source. Also determine the entropy of second extension & also verify that $H(S^2) = 2H(S)$

→ $H(S) = \sum_{i=1}^{3} P_i \log_2 1/P_i$

$= P_1 \log_2 1/P_1 + P_2 \log_2 1/P_2 + P_3 \log_2 1/P_3$

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

$H(S) = 1.5$

$S = 3$ II extension $= 3^2 = 9$

S_1S_1 occurs with prob P_1^2

S_1S_2 — " — P_1P_2

S_1S_3 — " — P_1P_3

S_2S_1 — " — P_1P_2

S_2S_2 " — P_2^2

S_2S_3 — " — P_2P_3

S_3S_1 — " — P_1P_3

S_3S_2 — " — P_2P_3

S_3S_3 — " — P_3^2

$$P_1^2 + 2P_1P_2 + 2P_1P_3 + 2P_2P_3 + P_2^2 + P_3^2 = 1$$

$P_1 = 0.5$

$P_2 = 0.25$

$P_3 = 0.25$

$$H(S^2) = \sum_{i=1}^{9} P_i \log \frac{1}{P_i}$$

$$= P_1^2 \log_2 \frac{1}{P_1^2} + 2P_1P_2 \log_2 \frac{1}{P_1P_2} + 2P_1P_3 \log_2 \frac{1}{P_1P_3} + 2P_2P_3 \log_2 \frac{1}{P_2P_3} + P_2^2 \log_2 \frac{1}{P_2^2} + P_3^2 \log_2 \frac{1}{P_3^2}$$

$= \cancel{2.54}$

$= 3$

15 A source emits one of the 4 probable messages m_1, m_2, m_3, m_4 with prob $\{7/16, 5/16, 1/8, 1/8\}$ respectively. Find the entropy of the source. List all the elements of the second extension of the source. Hence show that $H(S^2) = 2H(S)$

→ $H(S) = \sum_{i=1}^{4} P_i \log \frac{1}{P_i}$

$= 1.796$ bits/msg sym

$m_1m_1 - P_1^2$ $m_2m_2 - P_2^2$ $m_3m_3 - P_3^2$ $m_4m_4 - P_4^2$

$m_1m_2 - P_1P_2$ $m_2m_3 - P_2P_3$ $m_3m_4 - P_3P_4$

$m_1m_3 - P_1P_3$ $m_2m_4 - P_2P_4$ $m_4m_1 - P_4P_1$

$m_1m_4 - P_1P_4$ $m_3m_1 - P_3P_1$ $m_4m_2 - P_4P_2$

$m_2m_1 - P_2P_1$ $m_3m_2 - P_3P_2$ $m_4m_3 - P_4P_3$

$$H(S^2) = P_1^2 \log_2 \frac{1}{P_1^2} + 2P_1P_2 \log_2 \frac{1}{P_1P_2} + 2P_1P_3 \log_2 \frac{1}{P_1P_3} + 2P_1P_4 \log_2 \frac{1}{P_1P_4} + P_2^2 \log_2 \frac{1}{P_2^2} + 2P_2P_3 \log_2 \frac{1}{P_2P_3} + 2P_2P_4 \log_2 \frac{1}{P_2P_4} + 2P_3P_4 \log_2 \frac{1}{P_3P_4} + P_3^2 \log_2 \frac{1}{P_3^2} + P_4^2 \log_2 \frac{1}{P_4^2}$$

$= 3.59$ bits/msg sym

$2H(S) = 2 \times 1.796$

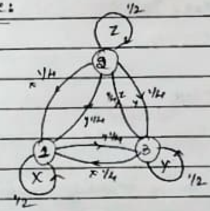
$= 3.592$

$= H(S^2)$

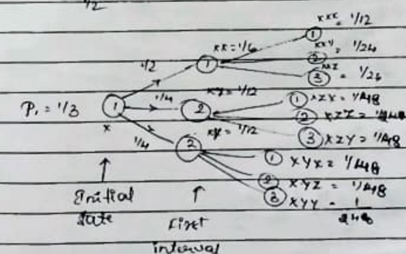
Entropy Average Information content of a symbol in a long dependent sequences:

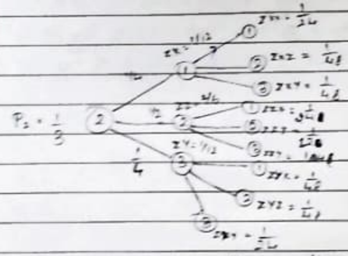
Markoff Statistical model for information source.

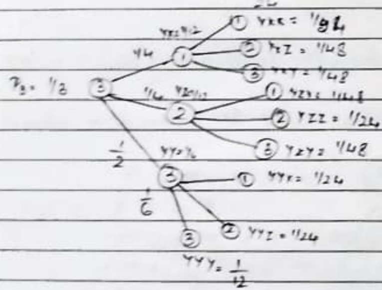
Example:



$P_1 = P_2 = P_3 = 1/3$







Entropy and information source using Markoff source

Entropy of each state $H_i = \sum_{j=1}^{n} P_{ij} \log_2 1/P_{ij}$

Average entropy $H = \sum_{i=1}^{n} P_i H_i$

$$H = \sum_{i=1}^{n} P_i \left[\sum_{j=1}^{n} P_{ij} \log_2 1/P_{ij} \right]$$

Information Rate

$R_s = r_s H$ bits/sec

Theorem:

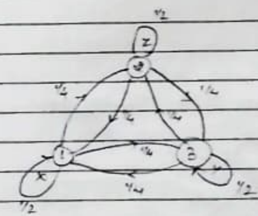
$$G_N = \frac{1}{N} \sum_i P(m_i) \log 1/P(m_i)$$

$$G_N = 1/N \cdot H(\bar{S}) \qquad H(\bar{S}) \to \text{adjoint entropy}$$

Information rate

$\lim_{N\to\infty} G_N = H$ bps

16. For the Markov source shown in fig (previous example)
(i) Entropy of each state
(ii) Entropy of source
(iii) G_1, G_2, G_3 then s.t. $G_1 > G_2 > G_3 > H$



$$H_i = \sum_{j=1}^{3} P_{ij} \log_2 1/P_{ij}$$

$$H_1 = \sum_{j=1}^{3} P_{1j} \log_2 1/P_{1j}$$

$$= P_{11} \log_2 \frac{1}{P_{11}} + P_{12} \log_2 \frac{1}{P_{12}} + P_{13} \log_2 \frac{1}{P_{13}}$$

$$= 1/2 \log_2(2) + 1/4 \log_2(4) + 1/4 \log_2(4)$$

$$= 3/2$$

$= 1.5$ bits/msg sym

$$H_2 = P_{21} \log_2 \frac{1}{P_{21}} + P_{22} \log_2 \frac{1}{P_{22}} + P_{23} \log_2 \frac{1}{P_{23}}$$

$$= 1/4 \log_2(4) + 1/2 \log_2(2) + 1/4 \log_2(4)$$

$$= 3/2$$

$= 1.5$ bits/msg sym

$$H_3 = P_{31} \log_2 \frac{1}{P_{31}} + P_{32} \log_2 \frac{1}{P_{32}} + P_{33} \log_2 \frac{1}{P_{33}}$$

$$= 3/2$$

$= 1.5$ bits/msg sym

$$H = \sum_{i=1}^{3} P_i H_i$$

$$= P_1 H_1 + P_2 H_2 + P_3 H_3$$

$$= \frac{1}{3}\cdot\frac{3}{2} + \frac{1}{3}\cdot\frac{3}{2} + \frac{1}{3}\cdot\frac{3}{2} = \underline{\underline{3/2}} = 1.5 \text{ bit}$$

$$H(\bar{S}) = \sum_{i=1}^{3} P(m_i) \log \frac{1}{P(m_i)} = P_1 \log \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2} + P_3 \log_2 \frac{1}{P_3}$$

$$= \frac{1}{3}\log 3 + \frac{1}{3}\log 3 + \frac{1}{3}\log 3$$

$G_1 = \frac{1}{1} H(\bar{S}) = 1.58$ bits/msg sym.

$G_2 = \frac{1}{2} H(\bar{S}^2) = \frac{1}{2} \cdot 3 \times \left[\frac{1}{6}\log_2 6 + \frac{1}{12}\log_2 12 + \frac{1}{12}\log_2 12\right]$

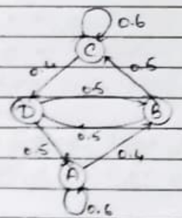
$= 1.54$ bits/msg sym

$G_3 = \frac{1}{3} H(\bar{S}^3) = \frac{1}{3}\, 3 \times \left[\frac{1}{12}\log_2(12) + \frac{4}{24}\log_2 24 + \frac{4}{48}\log_2 48\right]$

$= 1.528$ bits/msg sym

$\therefore G_1 > G_2 > G_3 > H$

17 Consider a state diagram of the markoff source as shown in fig.

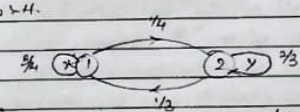


(i) Compute state probabilities
(ii) Find entropy of each state
(iii) Find entropy of source

→ $P(A) = 0.6\,P(A) + 0.5\,P(D) \rightarrow ①$
$P(B) = 0.4\,P(A) + 0.5\,P(D) \rightarrow ②$
$P(C) = 0.5\,P(B) + 0.6\,P(C) \rightarrow ③$
$P(D) = 0.4\,P(C) + 0.5\,P(B) \rightarrow ④$

$P(A) + P(B) + P(C) + P(D) = 1$
$0.4\,P(A) - 0.5\,P(D) = 0$
$0.4\,P(C) - 0.5\,P(B) = 0$

18. The given state diagram for a markov source shown in fig. (i) find entropy H of the source, (ii) Find G_1, G_2 hence show that $G_1 > G_2 > H$.

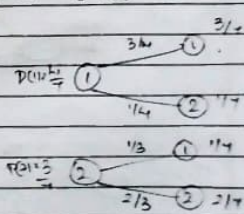


$H_1 = \sum_{j=1}^{2} P_{ij} \log 1/P_{ij} = 3/4 \log_2 4/3 + 1/4 \log_2 4 = 0.811$ bits/msg.

$H_2 = \sum_{j=1}^{2} P_{ij} \log_2 1/P_{ij} = 1/3 \log_2 3 + 2/3 \log_2 3/2 = 0.918$ bits/msg

$P(1)$: $3/4\,P(1) + 1/3\,P(2) \longrightarrow 1/4\,P(1) - 1/3\,P(2) = 0$
$P(2) = 1/4\,P(1) + \frac{2}{3}P(2) \longrightarrow 1/4\,P(1) - 1/3\,P(2) = 0$

$P(1) = 4/7 \qquad P(2) = 3/7$



$H = P(1) H_1 + P(2) H_2$
$= 4/7 \times 0.811 + 3/7 \times 0.918$
$= 0.856$ bits/msg sym

$H(\bar{S}^2) =$

$\frac{1}{N} H(S^N)$

$$G_1 = \frac{1}{1}\left[\frac{4}{7}\log_2 \frac{7}{4} + \frac{3}{7}\log_2 \frac{7}{3}\right]$$

$= 0.985$ bits / msg sym

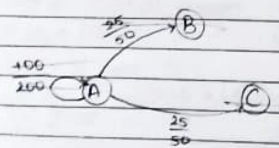
$\frac{1}{N} H(S^N)$

$$G_2 = \frac{1}{2}\left[\frac{3}{7}\log_2 \frac{7}{3} + \frac{2}{7}\log_2 7 + \frac{2}{7}\log_2 \frac{7}{2}\right]$$

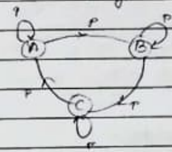
$= 0.9211$ bits / msg sym

$\therefore\ G_1 > G_2 > H$

19 Design information system which gives the information every year for about 200 students passing out with BE - ECE degree from the uni. The students can get into one of the fields as given below. (i) Go abroad for higher studies - A (ii) Join MBA or Civil service - B (iii) Join industry in India - C. Based on the data given below, construct the model for the source and find the source entropy.
On the average 100 students are going abroad. Out of 100 going abroad this year 50 were expected going abroad next year, while 25 each went to MBA & Civil services or joined industries in India. (iii) Out of 100 remaining in India this year 50 continued to do so, while 50 went abroad next year.



20. For the 2 order Markoff source, source alphabet is {A, B, C} shown in fig. (i) Compute the probs of states (ii) Find $H(S)$ & $H(S^2)$



$P(A) = pP(A) + pP(C)$
$P(B) = pP(A) + pP(B)$
$P(C) = pP(B) + pP(C)$

19.→ $H_A = \sum_{j=1}^{3} P_{ij} \log_2 \frac{1}{P_{ij}}$

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

$= 1.5$

$H_B = \sum_{j=1}^{3} P_{ij} \log_2 \frac{1}{P_{ij}}$

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

$= 1$

$H_C = \sum_{j=1}^{3} P_{ij} \log_2 \frac{1}{P_{ij}}$

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

$= 1$

$P_A = 0.5P(A) + 0.5P(B) + 0.5P(C) \rightarrow 0.5P(A) - 0.5P(B) - 0.5P(C) = 0$

$P_B = 0.5P(B) + 0.25P(A) \qquad 0.5P(B) - 0.25P(A) = 0$

$P_C = 0.5P(C) + 0.25P(A) \qquad 0.5P(C) - 0.25P(A) = 0$

On solving, $P(A) = 0.5$
$P(B) = 0.25$
$P(C) = 0.25$

$H = \sum_{i=A}^{C} P_i H_i$

$= P_A H_A + P_B H_B + P_C H_C$

$= 0.5 \times 1.5 + 0.25 \times 1 + 0.25 \times 1$

$\boxed{H = 1.25}$

$P(A) = pP(A) + pP(C) \rightarrow ①$
$P(B) = pP(A) + pP(B) \rightarrow ②$
$P(C) = pP(B) + pP(C) \rightarrow ③$

20 $P(A) = pP(A) + pP(C)$

$P(A)+P(B)+P(C) = 1$

$2pP(A) + 2pP(B) + 2pP(C) = 1$

$2p\,[P(A)+P(B)+P(C)] = 1$

$2p = 1$

$p = 1/2$

$① \Rightarrow P(A) = 1/2\,P(A) + 1/2\,P(C) \;=\; P(A) = P(C)$

$② \rightarrow P(B) = 1/2\,P(B) + 1/2\,P(A) \Rightarrow P(B) = P(A)$

$\therefore\; P(A) = P(B) = P(C) = 1/3$

$H(A) = \sum_{j=1}^{2} P_{ij} \log_2 1/P_{ij}$
$= 1/2 \log_2 2 + 1/2 \log_2 2 = 1$ bits/msg sym.

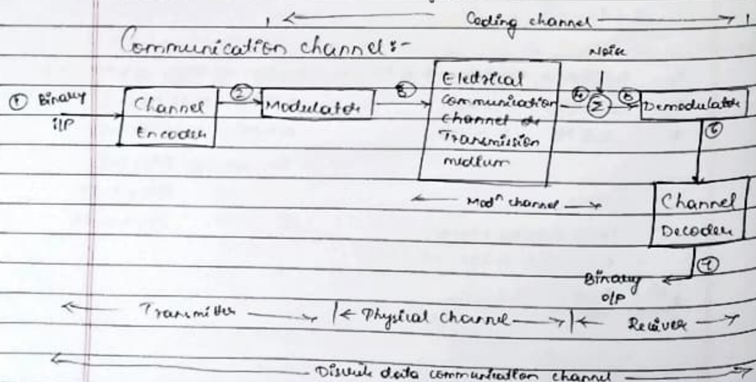
$H_B = 1$ bits/msg sym

$H_C = 1$ bit/msg sym.

$H(S) = \sum_{i=1}^{3} P_i H_i = 1/3 + 1/3 + 1/3 = 1$ bit/sym

$H(S^2) = 2 \cdot H(S) = 2$ bits/msg sym

Communication channel :-



Discrete communication channel :-

Representation of channel :-

Input 'A' with 'r' symbols

Output 'B' with 's' symbols

$A = \{a_1, a_2, a_3 \dots a_r\}$

$B = \{b_1, b_2, b_3 \dots b_s\}$

Conditional probability

$P(b_j/a_i)$

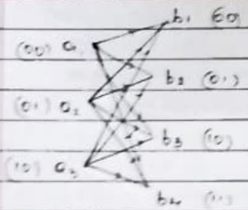
$$A \left\{ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_r \end{matrix} \quad P(b_j/a_i) \quad \begin{matrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_s \end{matrix} \right\} B$$

Channel matrix $(r \times s)$ or Noise matrix :-

$$P(B/A) = P(b_j/a_i) = \begin{bmatrix} P(b_1/a_1) & P(b_2/a_1) & P(b_3/a_1) & \dots & P(b_s/a_1) \\ P(b_1/a_2) & P(b_2/a_2) & P(b_3/a_2) & \dots & P(b_s/a_2) \\ \vdots & \vdots & \vdots & & \\ P(b_1/a_r) & P(b_2/a_r) & P(b_3/a_r) & \dots & P(b_s/a_r) \end{bmatrix}_{r \times s}$$

$A = \{a_1, a_2, a_3\} = \{00, 01, 10\}$

$B = \{b_1, b_2, b_3\} = \{00, 01, 10, 11\}$



$P_{11} = P(b_1/a_1)$

$P_{12} = P(b_2/a_1)$

$P_{13} = P(b_3/a_1)$

$P_{11} + P_{12} + P_{13} + P_{14} = 1$

$P(b_1/a_1) + P(b_2/a_1) + P(b_3/a_1) + P(b_4/a_1) = 1$

$\sum_{j=1}^{s} P(b_j/a_i) = 1$

Similarly

$\sum_{j=1}^{s} P(b_j/a_i) = 1$

Sum of any row in channel matrix is 1.

$P(a_1) + P(a_2) + \dots + P(a_r) = 1$

$\sum_{i=1}^{r} P(a_i) = 1$

$P(b_1) = P(b_1/a_1)P(a_1) + P(b_1/a_2)P(a_2) + \dots + P(b_1/a_r)P(a_r)$

$P(b_2) = P(b_2/a_1)P(a_1) + P(b_2/a_2)P(a_2) + \dots + P(b_2/a_r)P(a_r)$

$\vdots$

$P(b_s) = P(b_s/a_1)P(a_1) + P(b_s/a_2)P(a_2) + \dots + P(b_s/a_r)P(a_r)$

This is called Theorem of total probability.

i/p probability $P(a_i)$

o/p — " — $P(b_j)$

Conditional probability $P(b_j/a_i)$

o/p — " — $P(a_i/b_j)$

$P(a_i/b_j)P(b_j) = P(b_j/a_i)P(a_i)$

$P(a_i/b_j) = \dfrac{P(b_j/a_i)P(a_i)}{P(b_j)}$

Joint Probability:-

$P(a_i;b_j)$ $(P(a_i \cap b_j)) = P(b_j/a_i)P(a_i) = P(a_i/b_j)P(b_j)$

$$P(b_j/a_i)P(a_i) = \begin{bmatrix} P(b_1/a_1)P(a_1) & P(b_2/a_1)P(a_1) & \dots & P(b_s/a_1)P(a_1) \\ P(b_1/a_2)P(a_2) & P(b_2/a_2)P(a_2) & \dots & P(b_s/a_2)P(a_2) \\ \vdots & & & \\ P(b_1/a_r)P(a_r) & P(b_2/a_r)P(a_r) & \dots & P(b_s/a_r)P(a_r) \end{bmatrix}$$

$$P(a_i, b_j) = \begin{array}{c|cccc} & b_1 & b_2 & & b_s \\ \hline a_1 & P(a_1,b_1) & P(a_1,b_2) & \dots & P(a_1,b_s) \\ a_2 & P(a_2,b_1) & P(a_2,b_2) & \dots & P(a_2,b_s) \\ & \vdots & & & \vdots \\ a_r & P(a_r,b_1) & P(a_r,b_2) & \dots & P(a_r,b_s) \end{array}$$

Property 1:-

$P(b_1) = P(b_1/a_1)P(a_1) + P(b_1/a_2)P(a_2) + \dots + P(b_1/a_r)P(a_r)$

$P(b_1) = P(a_1,b_1) + P(a_2,b_1) + \dots + P(a_r,b_1)$

$P(b_2) = P(a_1,b_2) + P(a_2,b_2) + \dots + P(a_r,b_2)$

$\vdots$

$P(b_s) = P(a_1,b_s) + P(a_2,b_s) + \dots + P(a_r,b_s)$

$\sum_{i=1}^{r} P(a_i;b_j) = P(b_j)$

Property 2:

$P(a_1) = P(b_1/a_1).P(a_1) + P(b_2/a_1).P(a_1) + \dots + P(b_s/a_1).P(a_1)$

$\rightarrow$ $P(a_1) = P(a_1,b_1) + P(a_1,b_2) + \dots + P(a_1,b_s)$

$P(a_2) = P(a_2,b_1) + P(a_2,b_2) + \dots + P(a_2,b_s)$

$\vdots$

$P(a_r) = P(a_r,b_1) + P(a_r,b_2) + \dots + P(a_r,b_s)$

$\sum_{j=1}^{s} P(a_i,b_j) = P(a_i)$

Property 3:-

$P(a_1) + P(a_2) + \dots + P(a_r) = 1$

$P(a_1, b_1) + P(a_1, b_2) + \dots + P(a_2, b_s)$

$+ P(a_2, b_1) + P(a_2, b_2) + \dots + P(a_2 b_s) \dots$

$+ P(a_3, b_1) + P(a_3, b_2) + \dots + P(a_3 b_s) = 1$

In a comm. system tr has 3 i/p syms i.e $A = \{a_1, a_2, a_3\}$ & receiver also has 3 o/p symbols $B = \{b_1, b_2, b_3\}$. The mat given below shows JPM in some marginal prob.

(i) Find missing probabilities in the table

(ii) Find $P(b_3/a_1)$ & $P(a_1/b_3)$

(iii) Are the events a_1 & b_1 statistically independent?

→

$a_i \backslash b_j$	b_1	b_2	b_3
a_1	1/12	*	5/36
a_2	5/36	1/9	5/36
a_3	*	1/6	*

$P(b_j) =$	1/3	14/36	*

→ $P(b_1) + P(b_2) + P(b_3) = 1$

$1/3 + 14/36 + P(b_3) = 1$

$P(b_3) = 5/18$

$P(a_1, b_1) + P(a_2, b_1) + P(a_3, b_1) = 1/3$

$P(a_3, b_1) = 1/9$

$P(a_1 b_2) + P(a_2, b_2) + P(a_3, b_2) = 14/36$

$P(a_1, b_2) = 1/9$

$P(a_1, b_3) + P(a_2, b_3) + P(a_3, b_3) = 5/18$.

$P(a_3, b_3) = 0$

$P(a_1, b_3) = P(b_3/a_1) \cdot P(a_1)$

$P(b_3/a_1) = \frac{P(a_1, b_3)}{P(a_1)} = \frac{5/36}{1/3} = \frac{5}{12}$

$P(a_1, b_3) = P(a_1/b_3) \cdot P(b_3)$

$P(a_1/b_3) = \frac{P(a_1, b_3)}{P(b_3)} = \frac{5/36}{5/18} = 1/2$

$P(a_1 \wedge b_1) = P(a_1, b_1) = P(a_1) P(b_1)$

$1/12 \neq 1/3 \times 1/3$