Huffman coding →

Steps →

① The source symbols are arranged in order of decreasing probability. Then the two of lowest probability are assigned '0' & '1' bit.

② Then combine last two symbols and move the combined symbol as high as possible.

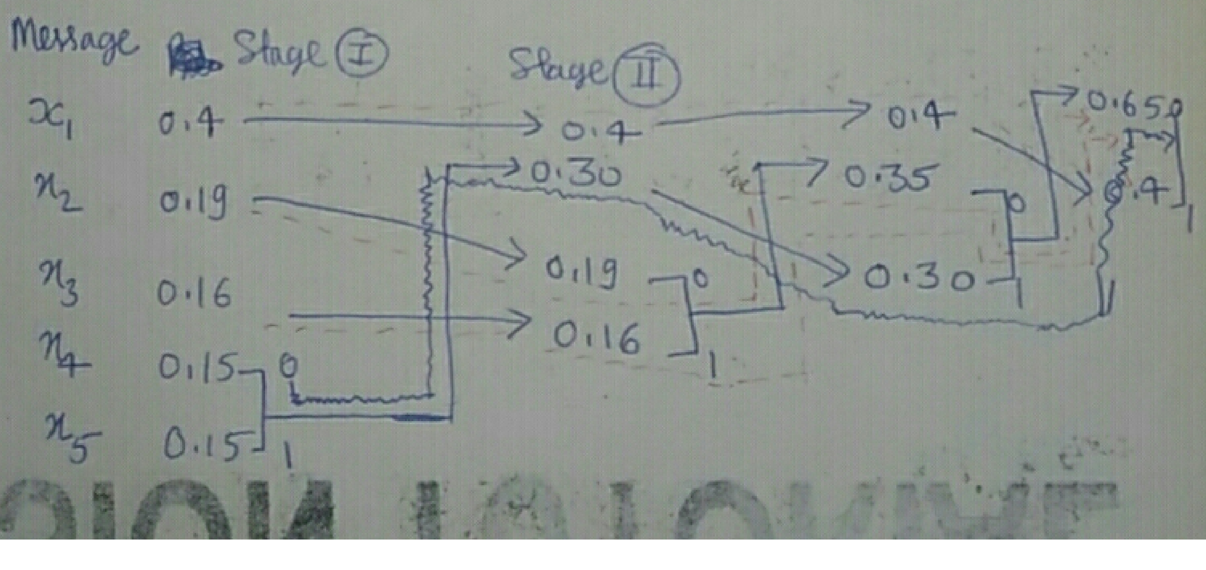
③ Repeat the above steps until end.

④ Code for each symbol is found by moving backward.

ex → A discrete memoryless source has five symbols x_1, x_2, x_3, x_4 and x_5 with probabilities 0.4, 0.19, 0.16, 0.15, 0.15 respectively attached to every symbol

① Construct a Huffman code for the source ② calculate code efficiency.

Sol^n – ①



Message	Probability	Code word	Length
x_1	0.4	1	1
x_2	0.19	000	3
x_3	0.16	001	3
x_4	0.15	010	3
x_5	0.15	011	3

$H = 2.2280$ bits / msg

$\bar{N} = 0.4 \times 1 + 0.19 \times 3 + 0.16 \times 3 + 0.15 \times 3 + 0.15 \times 3$

$\bar{N} = 2.35$

code efficiency $\rightarrow$ $\eta = \frac{H}{\bar{N}}$

$\eta = \frac{2.2280}{2.35}$

$\eta = 0.948$

$\eta = 94.8\%$

Ans

Conditional Entropy →

The conditional entropy $H(X/Y)$ is called equivocation. It is defined as →

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

$H(X/Y)$ → represents uncertainty of X, on average when Y is received.

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

Joint Entropy →

The joint entropy $H(X,Y)$ is given as →

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

Problem → compare the Huffman coding and Shannon-Fano coding algorithms for data compression. For a DMS 'X' with six symbols $x_1, x_2 \ldots\ldots x_6$, find a compact code for every symbol if the probability distribution is as follow →

$P(x_1) = 0.3$, $P(x_2) = 0.25$, $P(x_3) = 0.2$

$P(x_4) = 0.12$, $P(x_5) = 0.08$, $P(x_6) = 0.05$