Devansh Shah
1914078
B2

ITC Tutorial 4

Q1) Given the message 'Independence'. Use shannon Fano coding & Hoffman coding to assign unique binary codes to the symbols in the given message. Compare the efficiency of these codes.

Symbols	Probability	stage 1	stage 2	stage 3	stage 4
E	4/12 = 0.333	0	0		
N	3/12 = 0.25	0	1		
D	2/12 = 0.167	1	0		
I	1/12 = 0.083	1	1	0	
P	1/12 = 0.083	1	1	1	0
C	1/12 = 0.083	1	1	1	1

Symbols	Probability	Codeword	length
I	0.083	110	3
N	0.25	01	2
D	0.167	10	2
E	0.333	00	2
P	0.083	1110	4
C	0.083	1111	4

$$\text{Entropy } (H) = \sum P_x \log_2 \frac{1}{P_x}$$

$$= 0.083 \log_2 \frac{1}{0.083} + 0.25 \log_2 \frac{1}{0.25}$$

$$+ 0.167 \log_2 \frac{1}{0.167} + 0.333 \log_2 \frac{1}{0.333}$$

$$+ 0.083 \log_2 \frac{1}{0.083} + 0.083 \log_2 \frac{1}{0.083}$$

$$H = 0.298 \times 3 + 0.5 + 0.43 + 0.52$$

$$\boxed{H(X) = 2.3454 \text{ bits/symbol}}$$

Average Codelength $= \sum P(x_i)(n_i)$

$$= 3 \times 0.083 + 2 \times 0.25 + 2 \times 0.167 + 2 \times 0.33 + 4 \times 0.083 + 4 \times 0.083$$

$$\therefore \boxed{L = 2.407}$$

Code efficiency $= \dfrac{H(X)}{L}$

$$= \frac{2.3454}{2.407}$$

$$= 0.9744$$

$$= 97.44\%$$

Huffman Coding

Symbols	Probability	1st Reduction	2nd Reduction	3rd Redu
E	4/12	4/12	4/12	5/12
N	3/12	3/12	3/12	4/12
D	2/12	2/12	3/12 0	3/12
T	1/12	2/12	2/12 1	
P	1/12 0	1/12		
C	1/12 1			

Symbols	Probability	Codeword	length
I	0.083	011	3
N	0.25	10	2
D	0.167	010	3
E	0.333	00	2
P	0.083	110	3
C	0.083	111	3

$$H = \sum P_{xi} \log_2 \frac{1}{P(x_i)}$$

$$= \left(0.083 \cdot \log_2 \frac{1}{0.083}\right) \times 3 + 0.25 \cdot \log_2 \frac{1}{0.25}$$

$$+ 0.333 \cdot \log_2 \frac{1}{0.333} + 0.167 \log \frac{1}{0.167}$$

$$= 0.298 \times 3 + 0.5 + 0.43 + 0.5214$$

4th Reduction

$$\boxed{H(x) = 2.3454 \text{ bits/symbol}}$$

7/12 0

5/12 1

$$\text{Avg code length} = \sum (P_{xi}\, n_i)$$

$$= 3 \times 0.083 + 2 \times 0.25$$

$$+ 3 \times 0.167 + 2 \times 0.333$$

$$+ 3 \times 0.083 + 3 \times 0.083$$

$$\boxed{L = 2.414}$$

Code efficiency $= \frac{H(x)}{L}$

$= \frac{2.3454}{2.414}$

$= 0.9715$

$= 97.15\%$

Efficiency of shannon fano coding algorithm is higher than hoffman coding

Q2) What is kraft inequality theorem? Show that this theorem holds true for binary codes generated in Q1.

→ Kraft's inequality theorem states that given a list of positive integers $(n_1, n_2 \dots n_r)$, there exists a prefix code with a set of code words $(\sigma_1, \sigma_2 \dots \sigma_r)$ where length of each code word $|\sigma_i| = n_i \ \forall i$ if and only if

$$\boxed{\sum_{i=1}^{n} S^{-n_i} \leq 1}$$

where S is the size of the alphabet 'S'

∴ we are encoding in binary

∴ $S = \{0, 1\}$

∴ $S = 2 \Rightarrow$ $\boxed{\sum_{i=1}^{n} 2^{-n_i} \leq 1}$

codewords through Huffman coding
(011, 10, 010, 00, 110, 111)

$$= \frac{1}{2^3} + \frac{1}{2^3} + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^3}$$

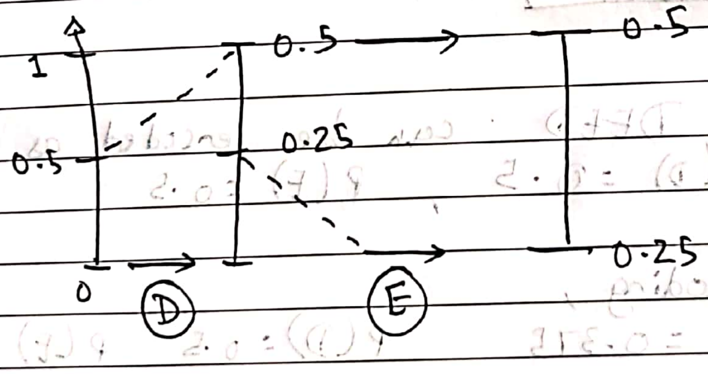
$$= \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = \underline{\underline{1}}$$

∴ It is a prefix code satisfying kraft inequality problem.

Q3) Show how arithmetic coding and decoding is done using message 'DEED'.

→ Message 'DEED'

Symbol	Probability	CDF
D	2/4 = 0.5	0.5
E	2/4 = 0.5	1



(i) For symbol D,

Range of symbol = lowerlimit + (diff × Prob of symbol)

$R(D) = ll + diff \times P(D)$
$= 0 + (0.5 - 0)\, 0.5$
$= \underline{0.25}$

$$R(E) = ll + diff \times P(d)$$
$$= 0.25 + 0.5 \times 0.5$$
$$= \underline{0.5}$$

(ii) For symbol E,

$$R(D) = ll + diff \times P(d)$$
$$= 0.25 + (0.25 \times 0.5)$$
$$= 0.375$$

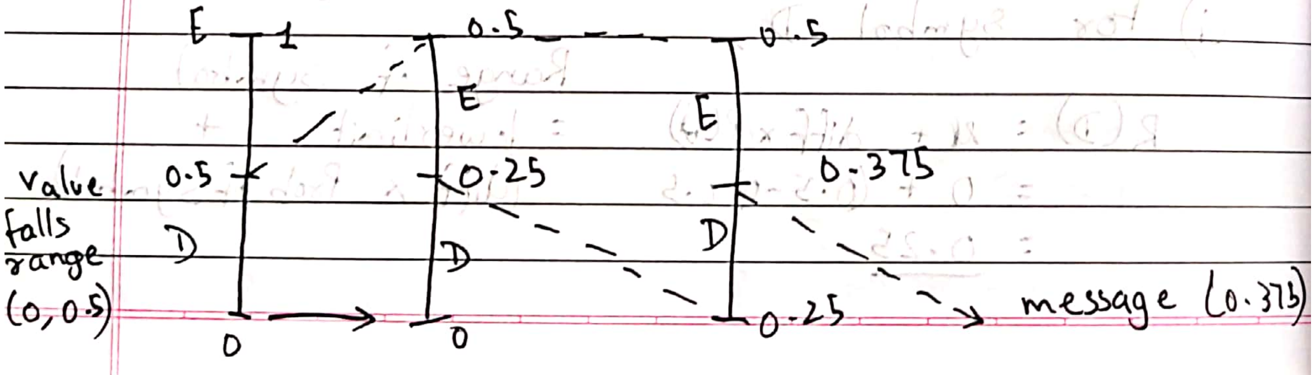
$$R(E) = l.l. + diff \times P(d)$$
$$= 0.375 + 0.5 \times 0.5$$
$$= 0.375 + 0.125$$
$$= 0.5$$

$$Tag = \frac{\text{Upper limit} + \text{lower limit}}{2}$$
$$= \frac{0.5 + 0.25}{2}$$

$$\boxed{Tag = 0.375}$$

Message DEED can be encoded as 0.375 when $P(D) = 0.5$, $P(E) = 0.5$

Now decoding,
Message $= 0.375$ $P(D) = 0.5$ $P(E) = 0.5$



$R(D) = 0 + 0.5 * 0.5$
$= 0.25$
$R(E) = 0.25 + 0.5 \times 0.5$
$= 0.5$

Now,
$R(D) = 0.25 + 0.25 \times 0.5$
$= 0.375$
$R(E) = 0.375 + (0.25 \times 1)$
$= 0.5$

Initially value falls in range $(0, 0.5) \rightarrow D$
Then falls in range $(0.25, 0.5) \rightarrow E$
then falls in range $(0.375, 0.5) \rightarrow E$
value at $D \rightarrow 0.375 \longrightarrow D$

$\Rightarrow$ Message is DEED