

# IC - A2

I.

1. given,  
Variable length codes

$$C(A) \stackrel{\#Tree1}{=} 0$$

$$C(B) = 100$$

$$C(C) = 101$$

$$C(D) = 110$$

$$C(E) = 111$$

#Tree2

$$C(A) = 00$$

$$C(B) = 01$$

$$C(C) = 10$$

$$C(D) = 110$$

$$C(E) = 111$$

$$(a) \underline{01000111101}$$

The codes are prefix free so,

$$\rightarrow A \underline{1000111101}$$

$$\rightarrow A \underline{1000111101}$$

$$\rightarrow A \underline{1000111101} \rightarrow AB \underline{0111101}$$

$$\rightarrow AB \underline{A111101}$$

$$\rightarrow AB \underline{AEC}$$

(b)

$L(C)$  for Tree #1

wkT,

$$P(A) = 0.5 \quad P(B) = P(C) = P(D) = P(E) = 0.125$$

The shortest message averaged over many messages is the Tree which has lower average length  $L(C)$

$l(x)$  is length of codeword

$$L(C)_{\#1} = \sum P(x) \cdot l(x)_{\#1} \quad x \in \{A, B, C, D, E\}$$

$$= 0.5 \times (1) + 0.125(3) + 0.125(3) + 0.125(3) + 0.125(3)$$

$$= 0.5 + 4 \times 3 \times 0.125$$

$$= 0.5 + 4 \times 0.375$$

$$= 0.5 + 1.5$$

$$L(C)_{\#1} = 2 \text{ bits}$$

$$L(C)_{\#2} = \sum P(x) \cdot l(x)_{\#2}$$

$$L(C)_{\#2} = \sum P(x) \cdot l(x)$$

$$= 0.5(2) + 0.125 \times 2 + 0.125 \times 3 + 0.125 \times 2 + 0.125 \times 3$$

$$= 1 + 0.25 + 0.375 + 0.25 + 0.375$$

$$= 2.25 \text{ bits}$$

$\therefore$  Tree #1 has a lower average code length (2.0) bits than Tree #2 (2.25) bits.

$$(c) \quad L_2 = 2.25 \text{ bits/symbol (average)}$$

$\downarrow$  length

for 100 symbol messages

$$L_{200} = 100 \times 2.25 = 225 \text{ bits}$$

For 100 symbol message the expected avg length is 225 bits

2. initial uncertainty for a 3-digit number  
with no clues  $H = \log_2(8) = 3 \text{ bits}$

(a) Alice told number is odd.

the last digit is ~~odd~~ (1)

(b) Bob said it's not multiple of 3  
then, possible values are (0, 3, 6)  
it should be (1 or 5 or 7)

(c) according to Charlie it has  
exactly two 1's then,  
it only '5' because 1 has only ②  
one no. of 1's and 7 has 3 no. of 1's  
where 5 only has 2 no. of 1's

(a) by Alice

Possible numbers are

{001, 011, 101, 111}

$$H_A = \log_2(4) = 2$$

$$I = H - H_A$$

$$I = 3 - 2 = 1 \text{ bit}$$

(b) by Bob possible numbers are

{001, 010, 100, 101, 111}

$$H_b = \log_2(5) = 2.3219$$

$$I = H - H_b = 3 - 2.3219 \approx 0.678 \text{ bits}$$

(c) by Charlie.

Possible values are  $\{011, 101, 110\}$

$$H_c = \log_2^{(3)} \approx 1.585 \text{ bits}$$

$$I = H - H_c = 3 - 1.585$$

$$I \approx 1.415 \text{ bits}$$

(d) from above theory if all these clues are given Deb then,

→ Deb can find exact answer, from (2) is 5 (101)

→ Hence the uncertainty for finding answer for Deb is 0.

$$\text{i.e. } I = H - \log_2^{(1)} = 3 - 0 = 3 \text{ bits}$$

the info gained by deb is 3 bits.

~~(A, C, D, B)~~

let,  
if all  $N$  outcomes are equally likely

$$P_i = \frac{1}{N} \quad N = 2^3$$

wkt,

$$H(X) = - \sum_{i=1}^N \left( \frac{1}{N} \log_2 \frac{1}{N} \right) \quad (\infty) \quad \sum_{i=1}^N \frac{1}{N} \log_2 \left( \frac{1}{N} \right)$$

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

$$H(X) = -\log_2 \left( \frac{1}{N} \right) = \log_2(N)$$

$I = H_{\text{before}} - H_{\text{after}}$  and for ~~8~~  $H$  with  
no clues  $H_x = \log_2(N) = \log_2 8 = 3$ .

3. given,

| I     | P(I) | $\log_2(1/P(I))$ | $P(I) \log_2(1/P(I))$ |
|-------|------|------------------|-----------------------|
| A     | 0.22 | 2.18             | 0.48                  |
| E     | 0.34 | 1.55             | 0.53                  |
| I     | 0.17 | 2.57             | 0.43                  |
| O     | 0.19 | 2.40             | 0.46                  |
| U     | 0.08 | 3.64             | 0.29                  |
| Total | 1.00 | 12.34            | 2.19                  |

(a) from above table

$$P(I) = 0.17$$

$$P(U) = 0.08$$

By given data

$$H_1 = 2.19$$

For probability to be  $I \text{ or } U$

is  $P(I) + P(U) = 0.25$

$$P(I|I \text{ or } U) = \frac{0.17}{0.25} = 0.68$$

$$P(U|I \text{ or } U) = \frac{0.08}{0.25} = 0.32$$

~~$$H_2 = 0.68 \log$$~~

$$H_2 = P(I|I \text{ or } U) \log_2 P(I|I \text{ or } U) + P(U|I \text{ or } U) \log_2 P(U|I \text{ or } U)$$

$$H_2 = -0.68 \log_2 0.68 - 0.32 \log_2 0.32$$

$$H_2 = 0.378 + 0.526 \approx 0.904 \text{ bits}$$

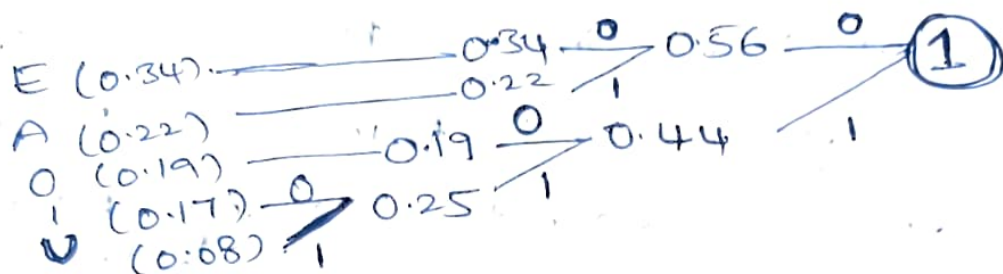
$$I = H_1 - H_2$$

$$I = 2.19 - 0.904$$

$$I = 1.286$$

So, The gain in Information is 1.286 bits.

(b)



$$E = 00$$

$$A = 01$$

$$0 = 100$$

$$1 = 110$$

$$U = 111$$

$$L(C) = P(E)l(E) + P(A)l(A) + P(0)l(0) + P(1)l(1)$$

$$L(C) = 0.34(2) + 0.22(2) + 0.19(2) + 0.08(2)$$

$$L(C) = 2.25 + 0.75 \Rightarrow (1.5 + 0.75)$$

$$L(C) = 2.25$$

(c)

avg for 1 codeword (Vowel) is 2.25

for 100 vowels are

$$L_{100} = 2.25 \times 100 = 225 \text{ bits}$$

d) wkt the Lower bound of Expectation of length  $L(C)$  is entropy

$$H(X) \leq L(C)$$

if bitdiddle transmits 197 bits  
for 100 vowels the 1.97 per vowel  
is below the level of given data  
entropy.

If vowels are independent then there  
definitely some certain error possibility  
for data transmitted using bitdiddle  
obtained code. can be treated as lossy compression.

Thus this ensures data distortion  
during decoding i.e. decoded data can lead  
to different ~~types~~ answers.

So, this ensures data cannot be decoded  
into unique value.

So Ben's code is non-uniquely decodable

4. (a) By given data.

$$I = H - H_i$$

for  $I$  to be small  $H_i$  should  
be larger

$$H_i = -P(x_i) \log \frac{1}{P(x_i)}$$

For to be change in entropy minimum  
(Information)  
 $H_i$  should be maximum probability distribution

⊗ So, The least info department is EECS(VI) ✓

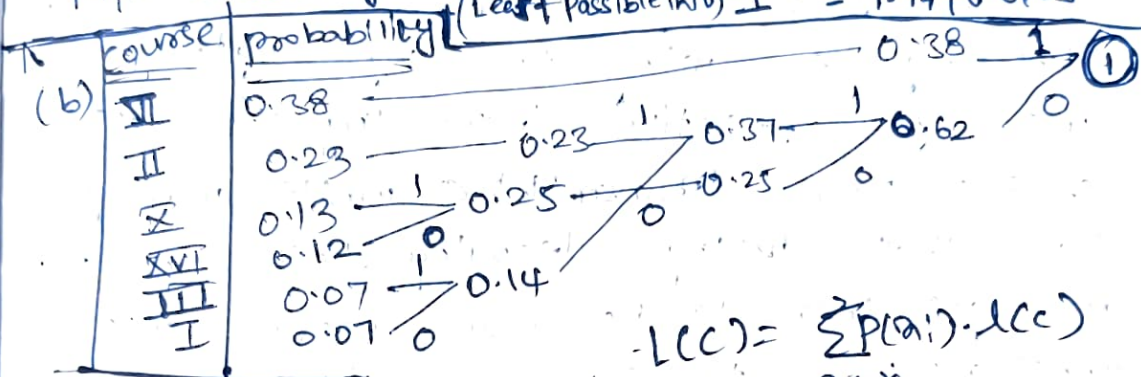
⊗  $H_i$  must be maximum

So if  $P(x) \uparrow$   $H_i \uparrow$   $P(x) \propto H_i \uparrow$

Intuitively, The student department with max probability has very high uncertainty already i.e high info even if involved (provided) it may not change much compare to other.

$$H_i = 0.38 \times \log 0.38 \Rightarrow I = H - H_i = 2.305 - 0.5304$$

$$(Least possible info) I = 1.7746 \text{ bits}$$



$$L(C) = \sum_{n \in X} P(x_i) \cdot L(x_i)$$

course code

I 0100 (4)

III 0101 (4)

II 011 (3)

I 1 (1)

X 001 (3)

XVI 000 (3)

$$H = -\sum P(x_i) \log P(x_i)$$

$$= 0.38 + 0.36 + 0.39 + 0.69 + 0.28 + 0.28$$

$$L(I) = 2.38$$

~~For hundred~~

(c) The average message length for 100 random group students.

For hundred students The total average expected message length =  $2.38 \times 100 = 238 \text{ bits}$  when encoding department for groups of randomly chosen students.

1. (a) Source coding:- codes the messages to be transmitted, it removes redundancy from the encoding process, it assigns less length code words to symbols with more probability of frequency within the message these by decreasing the avg length of the message effectively decreasing the amt of bits to be transmitted while making sure that during decoding these encoded messages there is no error where 2 or more messages have the same encoded message

Ex:- Huffman coding.

(b) Channel coding:- on the other hand is performed on the encoded bit messages obtained after source coding, it adds redundancy to the encoded message such that, even if some error was generated by the channel still the message could mostly be reconstructed, the channel through which the message is sent could introduce errors into the transmitted message for which we introduce redundancy to keep track of the error that crept in and to correct them.

Ex:- Hamming Code

1(b) for a  $C(n, k)$  linear block code

$$u^T H = 0$$

$$\dim(V) + \dim(V^\perp) = n$$

$k$

$(n-k)$

$\downarrow$

$\dim(G)$

$\dim(H)$

$G$  is dual  $H$  and vice versa

$G$  &  $H$  are subspace of  $\mathbb{R}^n$

$$H^T G = 0$$

$$C H^T = 0$$

$$H C^T = 0$$

Linear block code

Generator matrix ( $G$ ):

The generator matrix is a matrix of  $k \times n$  dimensions that generates  $n$  length encoded messages from  $k$  length message codewords.

Codewords of  $C(n, k)$  generated from  $G$  are linear combination of rows of  $G$ .

Set of Codewords are subspace  $\mathbb{F}_2^n$

$G$  & its rows are linearly independent

Parity check matrix ( $H$ ):

The parity check matrix is used to detect error in received codewords/data. It is an  $(n-k) \times n$  dim. matrix that when multiplied with a valid codeword, generates a Null matrix.

$H$  also can be found by Tanner graph representation

\* These are limits for error detection and error corrections based on no. of redundant / parity bits.

2) (5,1) REP code  $\rightarrow$  REP-5

wkt,

$$G = [1 \ 1 \ 1 \ 1 \ 1] \Rightarrow [I \mid \underset{5-1}{1's}] \Rightarrow [1 \ 1 \ 1 \ 1 \ 1]_{1 \times 5}$$

$$C = I \cdot G$$

$I$  can be  $[0]$  or  $[1]$

Parity check matrix  $H$

$$H = [P^T \mid I_{5-1}]$$

$$H = \left[ \begin{array}{c|c} 1 & I_4 \end{array} \right]$$

$$H_{\text{REP-5}} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

4x5

$H \cdot C^T = 0$  or  $C \cdot H^T = 0$  for above  $H$  &  $C$  ✓

\*  $G$  of REP- $n$  is dual of  $H$  of SPC- $n$  similarly  $H$

$H$  of REP- $n$  is dual of  $G$  of SPC- $n$

3) (6,5) single parity check code

$C$  = linear combination of any 5 digit numbers which belongs to  $\mathbb{F}_2^5$

$\dim(G) = 5$ ,  $\mathbb{F}_2^5 \subset \mathbb{F}_2^6$ ,  $\mathbb{F}_2^5$  is span{each row in  $G$ }

$$G = \left[ \begin{array}{c|c} I_5 & P \end{array} \right] = [I_5 \mid P]$$

wkt,

$$H = [P^T | I_{n-k}] = [P^T | I_{6-5}]$$

$$H = [11111 | 1] = [111111]_{1 \times 6}$$

as we can observe it also could be

→ It is also systematic G & H.

→ H for (6,5) single parity check matrix is  
G for Rep-6 code

→ Similarly G for (6,5) single parity check  
matrix is H for REP-6 code

4. given  $m$  (no. of parity bits) = 4  
 $k = 11$  (no. of message bits).

acc to hamming code

$$k \leq 2^m - m - 1 \quad (\text{or}) \quad 2^m \geq k + m + 1$$

$$11 \leq 2^4 - 4 - 1 = 16 - 5 = 11 \quad \checkmark$$

So, we can proceed with 4 parity bits  
in each code word.

→ data bits  $k = 11$

so, we can write ~~the~~ parity check matrix H as  
const of all non-zero binary column vectors length  $m=4$   
(11) we can derive parity eqn from

$$H = [1 - 15]$$



By labelling  
P's inside boxes

$$P_1 = D_1 \oplus D_2 \oplus D_4 \oplus D_5 \oplus D_7 \oplus D_9 \oplus D_{11}$$

$$P_2 = D_1 \oplus D_3 \oplus D_4 \oplus D_6 \oplus D_7 \oplus D_{10} \oplus D_{11}$$

$$P_3 = D_2 \oplus D_3 \oplus D_4 \oplus D_8 \oplus D_9 \oplus D_{10} \oplus D_{11}$$

$$P_4 = D_5 \oplus D_6 \oplus D_7 \oplus D_8 \oplus D_9 \oplus D_{10} \oplus D_{11}$$

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

(or) of finding H  
[Both ways can be done]

4x15

$$H_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

4x15

$$H = \left[ I_4 \mid P \right]_{4 \times 15}$$

wkt,

$$G = \left[ P^T \mid I_{n-k} \right]$$

$$G = \left[ P^T \mid I_{11} \right]_{11 \times 15}$$

after writing  $I_{11}$  on right of  $11 \times 15$  matrix  $P$  can be written correspondingly by previously obtained equations or else  
Just write  $-P^T$  of  $H_2$  and  $I_{11}$

$$G = \left[ P_{H_2}^T \mid I_{11} \right]_{11 \times 15} \text{ (or) } \left[ I_{11} \mid P^T \right]_{11 \times 15}$$

$$G(15,4) = G =$$

$$\left[ \begin{array}{cccccccccccccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{array} \right]_{11 \times 15}$$

11x15

⑤ given  $G = \begin{bmatrix} I_5 & 11 \end{bmatrix}$

(a) rate =  $\frac{\text{no. of info bits}}{\text{total bits}}$

by given data  $G$  is  $5 \times 6$  matrix

→ by this redundancy is 1 & Blocklength (SPC-6) is 6.

→  $C$  is  $(6,5)$  & looks like a single parity

→ No. of Information bits  $(k) = 5$  (rows of  $I_5$ )  
 check code  
 rate =  $\frac{5}{6} = 0.83$

(b)  $[w \ x \ y \ z \ u] \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = [p] = [C]_{1 \times 6}$

$5 \times 6$

Set of all code words are  $[w \ x \ y \ z \ u \ p]$

where  $p$  is  $F_2$  sum (XOR) of  $w, x, y, z$

$p = w \oplus x \oplus y \oplus z$  (or)  $p = w + x + y + z + u$

$\boxed{+ \in F_2}$

$$C = [m_1, m_2, m_3, m_4, m_5, p]$$

$$P = \sum_{i=1}^5 m_i (\text{mod } 2) \quad \text{or} \quad + \rightarrow \mathbb{F}_2 (+; 0, 1)$$

$$\text{code word for } (10010) = [1, 0, 0, 1, 0, (1+0+0+1+0)]$$

$$= [1, 0, 0, 1, 0, 0]$$

$$= [100100]$$

$$\text{wkt, } G = [I_5 | 1] \text{ since we noticed}$$

1 acts as a parity in parity ~~check~~ ~~matrix~~ <sup>column</sup>

for  $C(6,5)$

The parity check equation

$$G = [I_5 | P]$$

let  $[c_1, c_2, c_3, c_4, c_5, c_6]$  be <sup>form of</sup> code word

$$\text{wkt, for above parity check eqn } c_1 + c_2 + c_3 + c_4 + c_5 + c_6 = 0 \pmod{2}$$

$$H = [P^T | I_{6-5}] = [11111 | 1]$$

$$H = [111111] \quad [c * [111111]^T \Rightarrow P_e] \text{ where } P_e \text{ should } P_e = 0 \text{ for no error transmission}$$

$$\text{For any valid code word: } H \cdot c^T = 0 \pmod{2} \quad \text{or} \quad c \cdot H^T = 0 \pmod{2}$$

b) given

$$G = [I_3 | P]$$

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

$$H = [P^T | I_{n-k}]$$

by the we can say

$$\text{than } n = 7$$

$$k = I_n = 3$$

$$P_n = 4$$

$$H = [P^T | I_{7-4}] = [P^T | I_3]_{4 \times 7}$$

$$H = \left[ \begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right]_{4 \times 7}$$

$$H = \left[ \begin{array}{ccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

$$C \cdot H^T = 0 \quad [P + P = 0]$$

$$4 \times 7 \times 4 \checkmark$$

G & H are Systematic