ITC Tutorial 3

Devansh Shah
19I4078
B2

Q1)

Symbol	Probability $P(X_i)$	Code	Code length $L(X_i)$	$P(X_i) \times L(X_i)$
A	0.3	0	1	0.30
B	0.25	10	2	0.50
C	0.2	110	3	0.60
D	0.15	1110	4	0.60
E	0.1	11110	5	0.50

Avg Code length $= \Sigma P(X_i) \times L(X_i)$

$= 0.3 + 0.5 + 0.6 + 0.5 + 0.6$

$= 2.5$ bits

Entropy $= \Sigma P(X_i) \log\left(\frac{1}{P(X_i)}\right)$

$= 0.3 \log_2 \frac{1}{0.3} + 0.25 \log_2 \frac{1}{0.25} + 0.2 \log_2 \frac{1}{0.2} + 0.15 \log_2 \frac{1}{0.15} + 0.1 \log \frac{1}{0.1}$

$= 2.228$

Code Efficiency $= \frac{\text{Entropy}}{\text{Avg Code length}}$

$= \frac{2.228}{2.5}$

$= \underline{0.8912}$

Code efficiency % $= 0.8912 \times 100$

$= 89.12\%$

Code redundancy = 1 - code efficiency
= 1 - 0.8912
= 0.1088

Code redundancy % = 0.1088 × 100
= 10.88 %

Q2) Difference between lossy and lossless Compression

Lossy Compression	Lossless Compression
1) Lossy compression is the method which eliminate that data which is not noticable.	While lossless does not eliminate that data is not noticeable
2) In this a file does not restore or rebuilt in its original form	In this file can be restored in it's original form.
3) Data quality's compromised	Does not compromise the data's quality.
4) This has more data holding capacity	This has less data holding capacity comparatively

When average code length approaches the entropy of DMS it is said to be entropy encoding. It is a loseless data compressing scheme.
Examples:
1) Shannon Fano coding
(i) List the symbols in order of decreasing probability
(ii) Split the 2 sets that are as close to equiprobable as possible. Upper set - 0, lower set 1

iii) Continue this process each time positioning them into sets until further splitting is not possible

(2) Huffmann coding

i) List the source symbols in decreasing possibility order

ii) Combine probabilities of 2 symbols with resultant probabilities. This is a reduction Continue this step until possibilities are remaining

iii) Start encoding the last reduction, first digit - 0 last digit - 1

iv) Now assign 0 & 1 to 2nd digit in the previous step and keep doing till first column is reached.