

# DB-126

December-2017

## 5<sup>th</sup> Year M.Sc., (CA & IT) Integrated (Data Compression)

Time : 3 Hours]

[Max. Marks : 100

1. Attempt any **four** : 20
- (a) Answer the Questions :
- (i) What is Data Compression ? Give few examples of compressed data formats.
  - (ii) Define lossless and lossy compression.
- (b) Define :
- (i) Compression ratio
  - (ii) Compression rate
  - (iii) Distortion
  - (iv) Modelling
  - (v) Coding
- (c) Define Entropy. Generate entropy for alphabet  $A = \{a_1, a_2, a_3, a_4, a_5\}$  with probabilities  $P(a_1) = 0.5, P(a_3) = 0.05, P(a_5) = 0.25, P(a_2) = P(a_4) = 0.10$ .
- (d) Explain Unique Decodability. Check for Unique Decodability  $\{00, 001, 10, 1001, 10011\}$ .
- (e) Prove mathematically - Information obtained from the occurrence of two independent events is the sum of the information obtained from the occurrence of the individual events.
2. Attempt any **four** : 20
- (a) Find Golomb code for integers  $n = 0, 13$  &  $1$  which is parameterized by  $m = 3$ .
- (b) Find Tunstall Code of 3 bits for  $A = \{Y, O, U\}$  where  $P(Y) = 0.7, P(O) = 0.1, P(U) = 0.2$ .
- (c) Generate Minimum Variance Huffman coding for alphabet  $A = \{a_1, a_2, a_3, a_4, a_5\}$  with probabilities  $P(a_1) = 0.5, P(a_3) = 0.05, P(a_5) = 0.25, P(a_2) = P(a_4) = 0.10$ .

- (d) Given the alphabet  $\{A, C, N\}$  with model  $F_x(C) = 0.25$ ,  $F_x(A) = 0.501$ ,  $F_x(N) = 1$  where  $F_x$  is cumulative frequency of the symbol. Encode sequence "CAN" using Arithmetic coding with scaling and generate Tag (Binary conversion of Tag is not required).
- (e) Give at least 3 differences
- Arithmetic Coding - Huffman coding
  - Huffman Coding - Tunstall Coding
3. Attempt the following : 20
- Generate count arrays for context order  $-1, 0, 1$  and  $2$  for "r a t a t a t a t" using PPMB. Initial Dictionary : 1-a, 2-b, 3-d, 4-n, Decode the sequence with LZW  
1 5 2 6 9 5 11 9 8 4 3
  - Decode using LZ78 with  $c(c) = 1$ ,  $c(d) = 2$ ,  $c(e) = 3$ ,  $c(n) = 4$ ,  $c(o) = 5$   $\langle 0, 3 \rangle$   $\langle 0, 4 \rangle$   $\langle 0, 1 \rangle$   $\langle 0, 5 \rangle$   $\langle 0, 2 \rangle$   $\langle 1, 2 \rangle$   $\langle 1, 1 \rangle$   $\langle 4, 2 \rangle$   $\langle 7, 5 \rangle$   $\langle 5, 3 \rangle$   $\langle 1, 4 \rangle$   $\langle 10, 1 \rangle$   $\langle 8, 3 \rangle$
  - Give difference between LZ78 and LZW.
  - Encode following sequence using LZ77 algorithm "PATRAPASATRAR" with window size as 13. Assume six as size of the look-ahead buffer.
4. Attempt any **Four** : 20
- Define Quantization, Uniform Quantization, Non-Uniform Quantization and Noise with example.
  - Explain MSE, SNR, PSNR.
  - Explain Gaussian Distribution, Laplacian Distribution & Gamma Distribution.
  - Quantize the sequence 4.3, 9, 3.5, 6.7, 2.9, 8.8 into a 4 level uniform quantizer.
  - Quantize the sequence 1.3, 6, 3.45, 6.7, 2.9, 9.9 into a 3 level uniform quantizer.
5. Attempt any **Four** : 20
- Explain in detail Vector Quantization.
  - Explain in detail Scalar Quantization.
  - Explain differential coding techniques in detail.
  - Find Extended Huffman code for every two symbols taken from the source with probability model  $P(a_1) = 0.837$ ,  $P(a_2) = 0.163$ .
  - Find Extended Huffman code for every two symbols taken from the source with probability model  $P(a_1) = 0.837$ ,  $P(a_2) = 0.163$ .