

GUJARAT TECHNOLOGICAL UNIVERSITY**BE - SEMESTER-VI (NEW) - EXAMINATION – SUMMER 2016****Subject Code:2161603****Date:11/05/2016****Subject Name: Data Compression and data Retrival****Time: 10:30 AM to 01:00 PM****Total Marks: 70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. * **h** need to be considered as blank space

- Q.1** (a) Define: - Compression Ratio, Entropy, Distortion, Data Retrieval, Query Optimization, HINT and Run Length Coding. **07**
- (b) Explain modeling and coding. Explain how this will help to reduce entropy for following data. **07**
9,11,11,11,14,13,15,17,16,17,20,21
- Q.2** (a) Explain Huffman Coding in detail with example. Define minimum variance Huffman codes. **07**
- (b) Explain Scalar Quantization in detail. **07**
- OR**
- (b) Explain Vector Quantization in detail. **07**
- Q.3** (a) Encode “aacdeaab” using Adaptive Huffman code. Derive Output string, Codes and final tree. **07**
- (b) Generate GOLOMB code for m=9 and n=8 to 13. **07**
- OR**
- Q.3** (a) Define Arithmetic Coding. Encode and Decode “BACBA” with arithmetic coding. ($P(A)=0.5, P(B)=0.3, P(C)=0.2$) **07**
- (b) Write procedure to generate TUNSTALL code. Generate TUNSTALL code with probability of $P(A)=0.6, P(B)=0.3, P(C)=0.1$ and n=3 bits. **07**
- Q.4** (a) Given an initial dictionary consisting of the letters a b r y **h**, encode the following message using the LZW algorithm: **h**bar**h**array**h**by**h**bararray**h**bay. **07**
- (b) Encode the sequence etab**h**cetab**h**and**h**betab**h**ceta using Burrows-Wheeler transform and move to front coding.* **07**
- OR**
- Q.4** (a) Encode the following sequence using the LZ77 and LZ78 algorithm: **h**bararray**h**bar**h**by**h**bararray**h**ba **07**
Assume you have a window size of 30 with a look-ahead buffer of size 15. Furthermore assume that $C(a)=1, C(b)=2, C(\mathbf{h})=3, C(r)=4$, and $C(y)=5$. *
- (b) Write a short note on Old JPEG standard and JPEG-LS. **07**
- Q.5** (a) Explain and compare Incident matrix and Inverted index with example. **07**
- (b) Explain Lemmatization and Stemming in detail. **07**
- OR**
- Q.5** (a) Explain skip pointers and Phrase queries with example **07**
- (b) Explain challenges in XML information retrieval. **07**
