Name ───────────────   ──────Roll No──────   ─────
Section

National University of Computer and Emerging Sciences, Lahore Campus

Course:	**Information Retrieval**	**Course Code:**	**CS317**
Program:	**BS(Computer Science)**	**Semester:**	**Fall 2018**
Duration:	**25 Minutes**	**Total Marks:**	**14**
Paper Date:	**19-Sept-18**	**Weight**	**3.3%**
Section:	**A**	**Page(s):**	**2**
Exam:	**Quiz 1**	**Roll No:**	

Question 1 [6 marks]

Let V = Vocabulory size,

N= Total number of documents

AveD = Average Document Length

|q| = query length

|posting| = length of posting list of a word

Write time and space complexity of different indexing methods in this table.

	Term Document Matrix	Forward Index	Inverted Index
Time Complexity for relevant document retrieval	|q| * N	|q| * N * AvgD	|q| * |posting|
space Complexity	N * V	N * AvgD	N * AvgD OR V * |posting|

Question 2 (a)

Suppose a company needs to store large number of financial figures. The value of numbers range from 20,000 to 60,000. Which of the following two options will be more space efficient for encoding these numbers. **Why?** [1 Mark]

a) Elias Gamma Encoding
b) 16 bit Fixed Length Encoding

Solution:
Elias Gamma will take more space since it will take 31 bits to store the numbers whereas fixed length can store them in 16 bits.

Question 2 (b)

Suppose a company needs to store large number of financial figures. The value of numbers range from 1 to 7. Which of the following two options will be more space efficient for encoding these numbers. **Why?** [1 Mark]

a) Elias Gamma Encoding
b) 8 bit Fixed Length Encoding

Solution:
8 bit Fixed Length Encoding will take more space since Elias Gamma will take only 5 bits to store the numbers.

Question 3

Decode following into integers using Elias Gamma decoding. [4 Marks]

11100110011001

How many numbers are encoded here?

Solution:

1110011 0 0 11001

1011 = 11
0 = 1
0 = 1
101 = 5

11,1,1,5

4 numbers encoded

Question 4

According to Heap's Law, what is number of unique words of a novel you are expected to see if you have read 25% text of the novel. Total words in novel are N = 100,000. [2 Marks] Assume K = 10, Beta = 0.5.

Solution:

0.25*100,000 = 25000

V = 10 * sqrt (25000) = 1581