


National University of Computer and Emerging Sciences, Lahore Campus

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

Question 1 [6 marks]

Let V = Vocabulary size,

N = Total number of documents

AveD = Average Document Length

$|q|$ = query length

$|\text{posting}|$ = 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 * \text{AveD}$	$ q * \text{posting} $
space Complexity	$N * V$	$N * \text{AveD}$	$N * \text{AveD}$ OR $V * \text{posting} $

Question 2

What is advantage of using Elias Gamma encoding as compared to fixed length encoding? [2 Marks]

Ans: Fixed Length encoding takes more space. For example if we encode small numbers using 32 bits then most of the bits are wasted. Elias Gamma Encoding takes less bits for small numbers and more bits for large numbers. Elias Gamma Encoding is more space efficient.

Name _____
Section _____

Roll No _____

Question 3

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

1100110001110111

How many numbers are encoded here?

Ans: 4 number are encoded

11001 100 0 1110111

5, 2, 1, 15

Question 4

According to Zipf's law, what fraction of text is occupied by 5th most frequent word? . [2 Marks]

$F = c/\text{rank}$

$F = 0.1/5 = 0.02 = 2\%$