IR Spring 2019 (CS) **Topic**: Crawler **QUIZ** 1 **Roll** 

**No----- Date**: 7/2/2019

## Question1:

Find out the simHash of following sentence, show all working. *one fish two fish* 

The binary value for each word is given blow

word	Binary value
one	1100
fish	1010
two	0001

## Question 2:

Consider a URL frontier, with 3 front queues and 4 back queues. As shown.

Move all possible URLS from front queues to back queues.

Front Queue 1 Front Queue 2

Assume than whenever a back queue is empty a URL of higher quality available will be added to it.

Front Queue 3

		_					1			1
	Comsat.p	k/news		Geon	ews.c	om/p				
	nu.edu.pk/1				k/faculty		Stanf		ord.com/	
Back Queu	e 1	Back	(Que	eue 2	ı [	Back Q	(ueue	3	Back Que	eue 4
						nu.edu	.pk/1	6		
Comsat.pk/	/home					nu.edu	.pk/1	4		