

V3V3V3V3V3V3V3V3 V3V3V3V3V3V3V3V3V3V3V3V3V3V3V3V3V3

National University of Computer & Emerging Sciences

FAST-Karachi Campus

CS4051- Information Retrieval

Quiz#1

Dated: February 15, 2023

Marks: 20

Time: 20 min.

Std-ID: \_\_\_\_\_ Sol \_\_\_\_\_

**Question No. 1**

Give a list of merits for Boolean Retrieval Model for Information Retrieval. [5]

- Very simple, intuitive and mathematical support available.
- Easy to implement idea and support exact matching with query.
- It is very general and can be extended to other than text data.
- Users can easily verify the results

**Question No.2**

Consider a Boolean query given below:

$Q = T1 \text{ AND } T2 \text{ AND } T3 \text{ AND } T4$

If we know the following facts:

Frequency of  $T1 >$  Frequency of  $T2$ ; and Frequency of  $T3 >$  Frequency of  $T4$ ;

**What will be an efficient query order in the form of query work? Justify your answer.**

As per the given information we have ( $T1 > T2$ ) and ( $T3 > T4$ ) but no information about  $T1/T2$  and  $T3/T4$  pairs are given, from the set of possible execution the place of we can have two execution sequences  $[(T1 \text{ AND } T2) \text{ AND } (T3 \text{ AND } T4)]$  or  $[(T3 \text{ AND } T4) \text{ AND } (T1 \text{ AND } T2)]$

### Question No.3

Suppose a query has a total of 6 relevant documents in the collection. System A and System B have each retrieved 10 documents, and the relevance status of the ranked lists is shown below:

System A: [- + - - + - - - -]

System B: [+ + + - - - - -]

where the leftmost entry corresponds to the highest ranked document, and the rightmost entry corresponds to the lowest ranked document. A “+” indicates a relevant document and a “-” corresponds to a non-relevant one. For example, the top ranked document retrieved by System A is non-relevant, whereas the top ranked document retrieved by B is relevant. [5]

Answer the following queries:

<b>Precision of System A</b>	2/10
<b>Precision of System B</b>	3/10
<b>Recall of System A</b>	2/6
<b>Recall of System B</b>	3/6

### Question No.4

Illustrate the differences between following pair of terms.

<b>Dictionary</b>	<b>Thesaurus</b>
A dictionary contains an alphabetical list of words that includes the meaning, etymology and pronunciation. Organization of words in dictionary in lexicographic order. Dictionary is used to see the meaning, type and pronunciation of word. Dictionary may show use of the word in a sentence.	A thesaurus is a book that contains relationships between words like: synonyms and antonyms. Organization of words in thesaurus in generally in thematic order (conceptual order). Thesaurus is used to see the similarity and differences between pair of words or groups. Thesaurus may show the right usage or different context or sense of words.
<b>Bi-word Index</b>	<b>Extended Bi-word Index</b>
Bi-word index contains consecutive words in ordering of their appearance in the document. Example: “Stanford university palo alto” may be index as three bi-words like “stanford university” “university palo” and “palo alto”.	Extended Bi-word index may contain non-consecutive words generally in order of appearance in the document. Example: “Coin in the pocket” may be index as “coin pocket”