USN | | | | | | | | | | |

12SSE322

# Third Semester M.Tech. Degree Examination, Dec. 2013 / Jan 2014.
## Information Retrieval

Time: 3 hrs. Max. Marks:100

**Note:** ***Answer any FIVE full questions.***

1 a. Define Information Retrieval. **(02 Marks)**
b. Explain with a neat diagram, the process of retrieving the information. **(08 Marks)**
c. Explain how retrieval of relevant information is directly affected both by the user and the logical view of the documents. **(10 Marks)**

2 a. Describe the extended Boolean model and fuzzy set model in detail. **(12 Marks)**
b. Discuss the main features of structured text retrieval model. **(08 Marks)**

3 a. Discuss any 5 measures used for performance evaluation of information retrieval. **(10 Marks)**
b. Explain the evaluation measures at the TREC reference collection. **(04 Marks)**
c. Consider a set Rq = $\{d_3, d_5, d_9, d_{25}, d_{39}, d_{44}, d_{56}, d_{71}, d_{89}, d_{123}\}$ containing the relevant documents. The ranking of the documents in the answer set is given by

| | | | |
|---|---|---|---|
| 1. $d_{123}$* | 5. $d_8$ | 9. $d_{187}$ | 13. $d_{250}$ |
| 2. $d_{84}$ | 6. $d_9$* | 10. $d_{25}$* | 14. $d_{113}$ |
| 3. $d_{56}$* | 7. $d_{511}$ | 11. $d_{38}$ | 15. $d_3$* |
| 4. $d_6$ | 8. $d_{129}$ | 12. $d_{48}$ | |

i) Find the precision and recall.
ii) Draw precision versus recall graph considering precision at 11 standard recall levels.
iii) Find average precision at seen relevant documents. **(06 Marks)**

4 a. Explain the query expansion through local clustering. **(08 Marks)**
b. Explain the different query protocols. **(05 Marks)**
c. What is relevance feedback? Explain the query expansion and Term Reweighing for the vector model. **(07 Marks)**

5 a. Explain how document preprocessing is divided into text operations. **(05 Marks)**
b. Mention the 2 general approaches to text compression and explain the statistical approach in detail. **(10 Marks)**
c. Discuss the inverted file compression. **(05 Marks)**

6 a. Explain suffix trees and suffix arrays index, with examples. **(12 Marks)**
b. Describe KMP algorithm and Abo – Corasick algorithm used in matching a set of patterns. **(08 Marks)**

7 a. Discuss the MIMD architecture in detail. **(10 Marks)**
b. List the steps required for query processing in a distributed information retrieval system. **(02 Marks)**
c. Explain the Information Action process, with neat diagram. **(08 Marks)**

8 Write short notes on :
a. SGML.
b. Signature files.
c. Hypertext model.
d. Ad – hoc and filtering retrieval. **(20 Marks)**

*****