

Code No: 117DX

**R13****JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****B. Tech IV Year I Semester Examinations, November/December - 2016****INFORMATION RETRIEVAL SYSTEMS****(Information Technology)****Time: 3 Hours****Max. Marks: 75****Note:** This question paper contains two parts A and B.

Part A is compulsory which carries 25 marks. Answer all questions in Part A.

Part B consists of 5 Units. Answer any one full question from each unit.

Each question carries 10 marks and may have a, b, c as sub questions.

**Part- A****(25 Marks)**

- 1.a) What is a Retrieval strategy? [2]
- b) What is the use of probabilistic retrieval strategy? [3]
- c) What are retrieval utilities? List them. [2]
- d) What is Regression analysis? [3]
- e) What is stemming? [2]
- f) What is CLIR and what are the resources needed to implement CLIR? [3]
- g) What is an Index and posting list? [2]
- h) What is the method to improve both efficiency and effectiveness of an Information retrieval? [3]
- i) What is a Non-First Normal form? [2]
- j) Draw the diagram that integrates structured data and text. [3]

**Part-B****(50 Marks)**

2. What are language models? Explain with example. [10]
- OR**
3. Explain briefly about
    - a) Simple term weights
    - b) Non-binary independence model
 [5+5]
  - 4.a) Explain relevance feedback in vector space model.
    - b) Explain briefly how clustering is implemented without a precomputed matrix? [5+5]
- OR**
- 5.a) Explain about N-grams.
    - b) Explain how document clustering is used to generate Thesaurus? [5+5]
  6. Discuss about distance measures in semantic networks. [10]
- OR**
7. Explain about simple phrases and complex phrases of parsing. [10]
  8. Explain about building and compressing inverted index. [10]
- OR**
9. Explain about
    - a) The construction of signature files
    - b) Scanning to remove false positives
 [5+5]
  10. Discuss about semi-structured search using a relational schema. [10]
- OR**
- 11.a) Explain with diagram about Distributed document retrieval.
    - b) Write short notes on web search. [5+5]