

Code No: **R42059**

R10

Set No. 1

IV B.Tech II Semester Supplementary Examinations, July/Aug - 2015

INFORMATION RETRIEVAL SYSTEMS

(Computer Science and Engineering)

Time: 3 hours

Max. Marks: 75

**Answer any FIVE Questions
All Questions carry equal marks**

- 1 a) What is Information Retrieval System? Distinguish between IRS and DBMS. [8]
b) What problems need to address when using a DBMS as part of an Information retrieval System? [7]
- 2 Explain the search capabilities of Information Retrieval System. [15]
- 3 a) Explain about N-Gram Data Structures and PAT Data Structure. [8]
b) What is Hypertext data structure? Where this data structure is extensively used in? [7]
- 4 a) Show the Data Flow in Information Processing System using Automatic Indexing. [8]
b) What are the benefits of a weighted index system over a Binary index system? [7]
- 5 a) Discuss the three basic methods for generation of a thesaurus. [5]
b) Explain about Automatic Term Clustering. [10]
- 6 a) How the Use of Hidden Markov Models for searching textual corpora has introduced a new paradigm for search? [8]
b) Explain about Ranking Algorithms. [7]
- 7 Explain Information Visualization Technologies. [15]
- 8 Use the Boyer-Moore text search algorithm to search for the term FANCY in the text string FANCIFUL FANNY FRUIT FILLED MY FANCY.
a) Show all of the steps and explain each of the required character shifts.
b) How many character comparisons are required to obtain a match?
c) Compare this to what it would take using the Knuth-Pratt-Morris algorithm [15]

Code No: **R42059**

R10

Set No. 2

IV B.Tech II Semester Supplementary Examinations, July/Aug - 2015

INFORMATION RETRIEVAL SYSTEMS

(Computer Science and Engineering)

Time: 3 hours

Max. Marks: 75

**Answer any FIVE Questions
All Questions carry equal marks**

- 1 a) What is Information Retrieval System and what are their Objectives? [8]
b) What are the two major measures commonly associated with information systems and also explain the Effects of Search on Total Document Space. [7]
- 2 a) Explain about Miscellaneous Capabilities. [8]
b) Explain Automatic indexing in detail. [7]
- 3 Discuss in detail about Stemming Algorithms. [15]
- 4 a) Explain about Probabilistic Weighting approach. What are the different areas in which the probabilistic approach may be applied? [8]
b) Explain about concept indexing with an example. [7]
- 5 a) Explain the concept of clustering and the process of clustering. What are the guidelines on the characteristics of the classes? [9]
b) Explain the concept Clustering Using Existing Clusters. [6]
- 6 a) What is Search Statements and Binding? [8]
b) What is Relevance Feedback? Is the use of positive feedback always better than using negative feedback to improve a query? [7]
- 7 a) Describe the need for information visualization. Under what circumstances is information visualization not useful. [10]
b) Describe how other senses could be used in displaying results from searches. [5]
- 8 Construct finite state automata for each of the following set of terms:
a) BIT, FIT, HIT, MIT, PIT, SIT
b) CAN, CAR, CARPET, CASE, CASK, CAKE
c) HE, SHE, HER, HERE, THERE, SHEAR
Be sure to define the three sets I, S, and P along with providing the state drawing [15]

Code No: **R42059**

R10

Set No. 3

IV B.Tech II Semester Supplementary Examinations, July/Aug - 2015

INFORMATION RETRIEVAL SYSTEMS

(Computer Science and Engineering)

Time: 3 hours

Max. Marks: 75

Answer any FIVE Questions

All Questions carry equal marks

- 1 a) Explain about Digital Libraries and Data Warehouses. [8]
b) What is the difference between the concept of a “Digital Library” and an Information Retrieval System? What new areas of information retrieval research may be important to support a Digital Library? [7]
- 2 a) Explain about Browse Capabilities. [7]
b) What is Indexing? Explain the objectives of indexing and also discuss about Automatic indexing. [8]
- 3 a) What are the two major data structures in any information system? [7]
b) Describe the similarities and differences between term stemming algorithms and n-grams. [8]
- 4 a) Explain in detail about Vector Weighting. What are the general problems with the Vector Model? [8]
b) Explain about Natural Language Processing. Describe how use of Natural Language Processing will assist in the disambiguation process [7]
- 5 What is clustering? If clustering has been completed on two different domains. Discuss the impact of merging the domains into a single cluster for both term clustering and item clustering. What factors will affect the amount of work that will be required to merge the clusters together? [15]
- 6 a) Explain Similarity Measures and Ranking. [8]
b) What are the two major approaches to generating queries? Explain in detail. [7]
- 7 a) Explain in detail about Information Visualization. [6]
b) What are the Aspects of the Visualization Process? [9]
- 8 a) What are the three classical text retrieval techniques have been defined for organizing items in a textual database? Explain Text Streaming Architecture. [5]
b) Explain about Software Text Search Systems approach. [10]

Code No: **R42059**

R10

Set No. 4

IV B.Tech II Semester Supplementary Examinations, July/Aug - 2015

INFORMATION RETRIEVAL SYSTEMS

(Computer Science and Engineering)

Time: 3 hours

Max. Marks: 75

Answer any FIVE Questions

All Questions carry equal marks

- 1 What are the four major functional processes that Information Storage and Retrieval System is composed? Explain in detail with neat figures. [15]
- 2 a) What are the similarities and differences between use of fuzzy searches and term masking? What are the potentials for each to introduce errors? [8]
b) What is the relationship between vocabulary browse and thesauri/concept Classes? [7]
- 3 a) What is the most common data structure used in both database management and Information Retrieval Systems? Explain in detail. [8]
b) Explain about Hidden Markov Models with an example. [7]
- 4 a) Explain the following algorithms
i) Simple Term Frequency Algorithm
ii) Inverse Document Frequency
iii) Signal Weighting
iv) Discrimination Value [12]
b) Under what conditions would the Bayesian and the Vector approach be the same? [3]
- 5 Discuss about Hierarchical clustering in Information Retrieval focuses on the area of hierarchical agglomerative clustering methods. [15]
- 6 a) How would you define an item on the Internet with respect to a search statement and similarity function? [7]
b) What are the six key characteristics of intelligent agents? [8]
- 7 Explain in detail about Cognition and Perception. [15]
- 8 a) Explain about Hardware Text Search Systems approach. [10]
b) Trade off the use of hardware versus software text search algorithms citing advantages and disadvantages of each in comparison to the other. [5]

