Total No. of Questions : 10]

SEAT No. :

P2072

[Total No. of Pages : 3

[5059] - 678

B.E. (Information Technology)

INFORMATION STORAGE AND RETRIEVAL

(2012 Pattern) (Elective - III)

Time : 2½ Hours] *[Max. Marks :70*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Figures to the right indicate full marks.***
4) ***Assume Suitable data, if necessary.***

Q1) Why single pass algorithm is better than Rocchio's Algorithm? Form the document cluster of following document term matrix using single pass clustering algorithm. Consider **[10]**

Membership Function: Sum of product

Centroid calculation Function: Average

Threshold = 11

	Dl	D2	D3	D4	D5
T1	1	1	0	1	1
T2	2	1	2	3	0
T3	3	0	1	0	1
T4	2	2	0	3	0
T5	2	2	1	2	1

OR

Q2) a) Explain working of suffix tree. Construct suffix tree for following example. **[6]**

"This is a text. A text has many words. Words are made from letters."

b) Write a short note on matching coefficients. **[4]**

P.T.O.

Q3) a) Write a note on "Ontology languages for semantic web". **[5]**

b) Write a note on "cluster based retrieval". **[5]**

OR

Q4) Consider a reference collection and its set of example information request. If q is the information request and a set **[10]**

Rq=(d3, d5, d9, d25, d39, d44, d50, d70, d80, dl20). Now consider new retrieval algorithm has been designed and has been evaluated for information request q returns, ranking of the documents in the answer set as.

1.	d120	6.	d9	11.	d38
2.	d84	7.	d58	12.	d48
3.	d50	8.	d129	13.	d230
4.	d6	9.	d143	14.	d113
5.	d8	10.	d25	15.	d3

The documents that are relevant to the query q are underlined. Calculate precision and recall for the documents that are relevant to the query q

Q5) a) Describe the query processing in Distributed IR. **[8]**

b) Explain Data models used for Multimedia IR. **[8]**

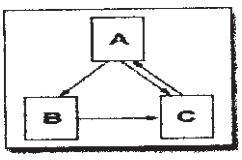
OR

Q6) a) What do you mean by collection partitioning & source selection in Distributed IR? **[8]**

b) Explain the generic multimedia Indexing Approach. **[8]**

Q7) a) What is Metasearcher? **[6]**

b) What is importance of page ranking? Calculate page rank of following web pages. **[12]**



OR

Q8) a) What are the challenges while searching the web? **[12]**

b) What is the role of crawler in web searching? Explain the strategies used by web crawler. **[6]**

Q9) a) Define Recommender System? Explain in brief collaborative filtering. **[8]**

b) Explain the method of extracting data from text. **[8]**

OR

Q10)a) Explain the concept of semantic web .How it is useful in web searching? **[8]**

b) Explain in detail content based recommendation of documents. **[8]**