M.Tech (CSE) 2nd Semester Examination

Jun.2013

Concepts of Information Retrieval

Subject Code: CSL 518

Time Allowed: 03 hours.   Maximum Marks: 100

Before answering the question paper the candidate should ensure that they have been supplied the correct question paper. Complaints in this regard, if any, shall not be entertained after the examination.

Note: Attempt any five questions and all questions carry equal marks.

SECTION - A

1(a) What is the difference between Data retrieval and Information retrieval. Also With the help of block diagram explain typical information retrieval system. ?

(b) Explain the following terms:

- Positional index
- Tired index
- Bi word indexes
- Inverted index
- Forward index

2(a) What is the importance of indexing in information retrieval and what are the various hardware basics of indexing construction? Difference between dynamic and distributed indexing?

(b) Explain various posting list compression techniques also explain Zipf’s law and heap’s law for posting compression?

3(a) Explain algorithm for query processing using OR,AND,NOT logic?

(b) Explain SPIMI and BSBI with algorithms? Explain its pros and cons?

SECTION-B

4(a) Explain the process of cluster pruning in details?

(b) Draw and explain the architecture of web crawler in details also explain the principle of robot exclusion?

5(a) What do you mean by web as a graph? Explain its bowtie structure?

(b) Explain the concept of cloaking AND spamming?

6. (a) What is Reputation of a page? How it is different from page rank? Derive the formula for computation of page rank of a page?

(b) What is the need of language models in information retrieval. Explain the concept of query likelihood model?