Ph. D. 2\textsuperscript{nd} Semester Examination
Jan.2014
Subject – 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.

\textbf{Note:} Question No. 1 is Compulsory and attempt any two questions from each section. All questions carry equal marks.

1. \(5\times4=20\)
   (a) Describe the positional index with example. Also List few advantages of positional indexes.
   (b) Give some characteristics of Web and explain it in brief.
   (c) Define term frequency and inverse document frequency (tf-idf). Why is the idf of a term always finite?
   (d) What is ranking? Explain about relevance score.

\textbf{SECTION - A}

2(a) What were reasons for origination of information retrieval system?
   What reasons force to do research in information retrieval system? \hspace{1cm} (10)
   (b) Discuss the extended Boolean model for information retrieval. How it is different than ranked retrieval method. \hspace{1cm} (10)

3(a) Differentiate between parametric and zone indexes. \hspace{1cm} (10)
   (b) Explain the efficient cosine scoring scheme. \hspace{1cm} (10)

4. Explain the concepts of term, document and query as vector? Describe the Vector space model in brief. \hspace{1cm} (20)

\textbf{SECTION - B}

5(a) What is Cluster pruning? Explain it with the help of suitable diagram. \hspace{1cm} (10)
   (b) Discuss the components of an Information retrieval system. \hspace{1cm} (10)

6(a) Describe the Bowtie structure of a web graph. \hspace{1cm} (10)
   (b) Write short notes on Markov chain. \hspace{1cm} (10)

7(a) What is crawling? Explain the crawler architecture with suitable diagram. Explain its working. \hspace{1cm} (10)
   (b) Describe the language model for Information retrieval system and their need. \hspace{1cm} (10)