CSE 202: Fundamentals of Database Systems
Winter 2018
Home Assignment 3
Due Date: 28th March 2018 (Two Week time: No extension will be allowed)

Instructions:
- Write the programs in **Java** only.
- **Only one submission per group.**
- The naming convention for the files: filename_rollnumber.extension like program1_2016001_2016002.java.
- **The assignment can be done individually or in groups of maximum 2 students.**
- **Compress all the input/output files along with the programs as tar.gz. Your submission (.tar.gz) must contain a text file with your group details (name and roll number of student). Your submission must contains files only with no folder hierarchy created.**

In this assignment, you have to implement a B+ index as discussed in the class. Your program must take as input a text datafile and should create an **index file** of B+ tree over a **field name** which is given as an argument. The program must read the datafile from disk and write the **index file** containing the B+ tree to disk. Your program should support addition and deletion of a record in the datafile and in the index. Addition of a record in the datafile would result in appending a data record at the end of the file. Deletion of a record would be a logical deletion that would put a tag '0000' in the beginning of the record to mark that the record is invalid. Any other string of 4 character would denote the record as a valid record. Tag would help to keep the offset unchanged for all the records that appear after the deleted record.

The format of the text data file is as follows:

Validity Tag	Instructor ID	Instructor Name	Department	Salary

The field names: Validity Tag, Instructor ID, are of String type of length 4 bytes, Instructor Name of String type of length 20 bytes, Department of String type of length 10 bytes, and Salary of numeric type of length 10 bytes. You can assume the order of B+ tree as 10. The field Instructor ID and Instructor Name will have unique values only. Two instructors can belong to the same department with the same Salary.

The program must implement the following operations.

- Find (Value V) - Returns the leaf node L and index i such that $L.P_i$ points to the first record with search value V. You will be required to load the index file in memory and then use that to locate the leaf node that contains the record with that key if any.
- PrintAll (Value V) - Print all records with the search value V. You will be required to load the index file first and then perform the printing of records.
- FindRange (L, U) - Find all records with search key values in a specified range (L, U) i.e. between L and U.
- Insert (Record R) - Append the record R at the end of the datafile and update the index file accordingly.
- Delete (Record R) - Delete the record R from the datafile and update the index file accordingly.