RollNo: _____ Name: _____

National University of Computer and Emerging Sciences, Lahore Campus

Course:	Advanced Database Concepts	Course Code:	CS451
Program:	BS(Computer Science)	Semester:	Spring 2017
Duration:	60 Minutes	Total Marks:	30
Paper Date:	07-Apr-2017	Weight	12.5%
Section:	CS	Page(s):	5
Exam:	Midterm-2		

Instruction/Notes: Scratch sheet can be used for rough work however, all the questions and steps are to be shown on this question paper. No extra/rough sheets should be submitted with question paper.
You will not get any credit if you do not show proper working, reasoning and steps as asked in question statements. Calculators are allowed.

Q1. *(10 points)* Assume that you have just built a dense B^+-tree index on a heap (unordered) file containing 50,000 records. The key field for this B^+-tree index is a 40-byte string, and it is a candidate key. Pointers (Record/block) are 10-byte values. The size of one disk page is 500 bytes. The index was built in using the bulk-loading algorithm, and the nodes at each level were filled up as much as possible.
How many levels does the resulting tree have? For each level of the tree, how many nodes are at that level? Show each step.

Ans:
order p = **10**; (p * 10) + ((p - 1) * 40) <= 500 [p <= (500+40)/50]
order p_{leaf} = **9**; (p_{leaf} * (40+10)) + 10 <= 500 [p_{leaf} <= (500-10)/50]
b1= ceiling(50,000/9) = 5556
fo = 10
b2= ceiling(5556/10)= 556
b3= ceiling(556/10)= 56
b4= ceiling(56/10)= 6
b5= 1

a) x=5; OR x = ceiling($\log_{fo}$ (b1)) + 1 = ceiling($\log_{10}$ 5556) + 1 = 4 + 1 = 5 levels
b) Lev1(b1)=5556, Lev2(b2)=556, Lev3(b3)=56, Lev4(b4)=6, Lev5(b5)=1

RollNo: _____ Name: _____

Q2. *(10 points)* Consider a relation R (a, b, c) with 20,000 records, 2,000 blocks (10 records fit on each block), and where *a* is a non-negative integer primary key. How many blocks will be read from disk to answer the selection query $\sigma_{a \text{ IN } (15, 35, 45, 65, 85)}(R)$ in each of the following scenarios? Justify your answer.

a) Relation R is stored in an unordered (heap) file. There also exists a B^+-tree index with search key a. Assume height of index is four and one node of the B^+-tree is stored in one block on the disk. **None** of the index blocks are in memory.

b) Relation R is stored in an unordered (heap) file and there is a B^+-tree index with search key a. All index blocks are already in memory.

c) Relation R is stored in an ordered (sequential) file sorted on a and there is a B^+-tree index with search key a. **None** of the index blocks are in memory.

d) Relation R is stored in an ordered (sequential) file sorted on a and there is a B^+-tree index with search key a. All index blocks are already in memory.

e) Relation R is stored in an unordered (heap) file. There also exists a hash-based index with search key a. 80 index entries fit on each block and **none** of the index blocks are in memory.

Ans:

a) 5 * (index cost + base table cost) = 5 * (X+1) = 5 * (4+1) = 25

b) Base table cost only = 5; (we will potentially make one block IO for each tuple for a cost of 5 block IOs)

c) 5 * (index cost + base table cost) = 5 * (X+1) = 5 * (4+1) = 25 (i.e. Same as part(a))

d) Base table cost only = 5; (we will potentially make one block IO for each tuple for a cost of 5 block IOs) (i.e. Same as part(b))

e) Index cost + base table cost = 5 + 5 = 10; (Since the index is a hash-based index and the predicate is an equality predicate, the best that we can do is scan the entire index. The total cost will be the index access cost i.e. 5 (one I/O for each value) plus 5 block IOs to actually fetch the data. Total cost = 10 IOs)

RollNo: _____ Name: _____

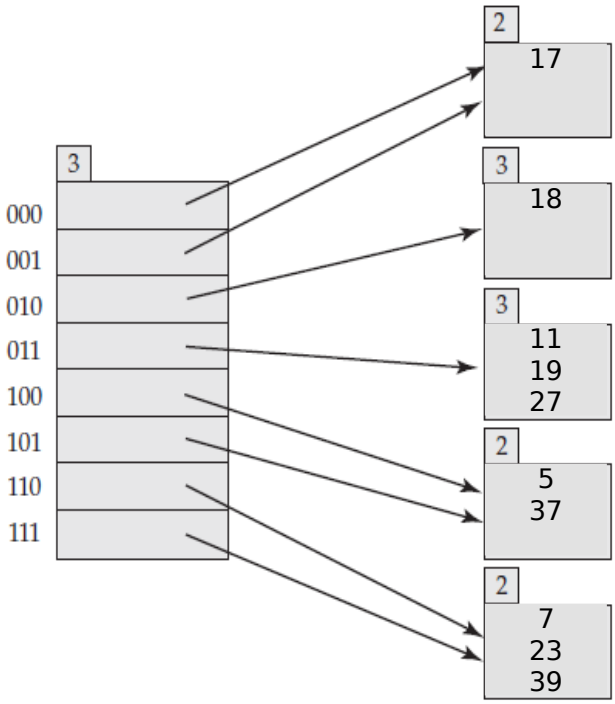
Q3. *(3 points)* Let R and S be relations with no indices, and assume that the relations are not sorted and r < s. Assuming infinite memory, what is the lowest amount of memory required (in terms of I/O operations) to compute R $\bowtie$ S? Assume tuples and blocks in R and S are r, s , b_r and b_s respectively.

Ans:
We can store the entire smaller relation in memory, read the larger relation block by block and perform nested loop join using the larger one as the outer relation. The number of I/O operations is equal to br + bs , and memory requirement is min(br , bs) + 2 pages.

RollNo: _____ Name: _____

Q4. *(7 points)* Suppose that we are using extendable hashing on a file that contains records with the following search-key values: 5, 7, 11, 17, 18, 19, 23, 27, 37, 39
Show the extendable hash structure for this file if the hash function is $h(k) = k \bmod 8$ and buckets can hold three records. Show your working.

Ans:
Extendable hash structure:



RollNo: _____ Name: _____