Assignment I, Database Systems

Due on 30th January 2014, in class by 2:10PM.

ONLY HANDWRITTEN ANSWERS WILL BE GRADED.

1. Consider a hard disk Megatron 777, assume it has two heads diametrically opposite of each other. The number of tracks traversed by each disk head is half the total number of tracks (i.e., number of tracks/number of heads). Each head has its own disk controller. Compute average time taken to access a disk block. What are the advantages or disadvantages of having two heads? Suppose k is the number of equally spaced disk heads, what is the ***optimal*** number of disk heads you would want to minimize disk access time? Make and state your justifiable assumptions. Formulate the problem and solve it.

Consider your roll number to be made of digits ***abcdefghi*** (for example, for roll number 201569214, *a* is 2, *b* is 0, …, *g* is 2, *h* is 1, and *i* is 4, *ab* is number 20). The correctness in mapping your roll number to above keys, and correctness in solving the problem is required to get marks.

2. Show the step by step building of dynamic hashing structure, when following keys made of digits ***bi, af, hd, def, cie, bac, hig, d, g, aihd, fed, gab*** are inserted in this order (duplicate values are allowed).

3. Show step by step building of the linear hash structure with the same keys (duplicate values are allowed), ***bi, af, hd, def, cie, bac, hig, d, g, aihd, fed, gab*** inserted into a linear hashing structure in this order where $h_0 = K \pmod{2^0}$.

4. Consider record length as ***agh***, key length as ***ag***, attribute A's length is ***ah***, length of block pointer is **16** bytes, length of record pointer is **32** bytes. Assume disk block or page size to be ***maximum(aghi, agh)***. The number of records is ***efghic***. *50% of marks are allotted to correctness of the values computed.*

   a. Calculate the blocking factor ***bfr***, and number of file blocks, assuming unspanned organization.

   b. How many blocks need to be accessed on an average to retrieve a record using linear scan.

   c. How many blocks need to be accessed on an average to retrieve a record for a key value assuming file is ordered on Key.

   d. How many disk block accesses on an average to access a record for a particular key value using Primary Index.

   e. Assuming each attribute A value has at least 5 rows, and at most 150 rows, what is the minimum and maximum number of disk block accesses needed to retrieve all rows for a particular attribute value, using (i) Clustering index, and (ii) Secondary index on A.

   f. What is number of levels for B+ Tree on the Key attribute? Where each block (both leaf and non-leaf) is kept at most 66% full.