1. [Relational Algebra, SQL, TRC, DRC]. Consider the following relations containing airline flight information:

Flights( flno, from, to, distance, departs, arrives )  
Aircraft( aid, aname, cruiserange )  
Certified( eid, aid )  
Employees( eid, ename, salary )

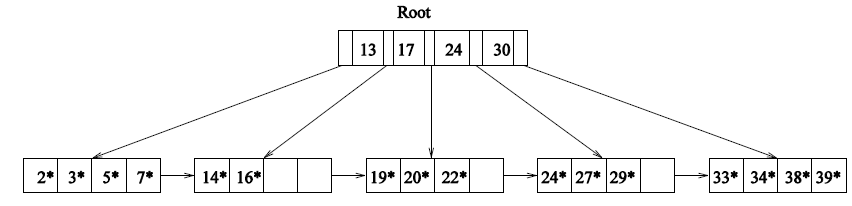
Note that Employees relation describes pilots and other types of employees as well. Every pilot is certified to fly some aircraft, and only pilots are certified to fly aircrafts. Write the following query in relational algebra, SQL, tuple relational calculus and domain relational calculus.

Query: Find names of all pilots certified for some Boeing aircraft.

2. [Functional Dependencies]. Suppose you are given a relation R with four atomic attributes ABCD. For the set of functional dependencies F: {AB --> C, AB--> D, C--> A, D--> B} that hold on R, do the following:

(a) Identify the candidate key(s) for R.  
(b) Identify the highest normal form that R satisfies (1NF, 2NF, 3NF, or BCNF).  
(c) If R is not in BCNF, decompose it into a set of BCNF relations that preserve the dependencies

3.[Indexing]. Consider the B+ tree shown below:



i) Identify a list of five data entries such that:

(a) Inserting entries in the order shown and then deleting them in the opposite order (e.g., insert a, insert b, delete b, delete a) results in the original tree. Explain your answer.

(b) Inserting entries in the order shown and then deleting them in the opposite order results in a different tree.

ii) What is the minimum number of insertions of data entries with distinct keys that will cause the height of the original tree to change from its current value of 1 to 3? Explain.

4.[Query Processing]. i. Consider joining two relations R(x,y) and S(x,z) on their common attribute x. The size of relation R is 150 blocks and the size of relation S is 100 blocks. Attribute x has 50 different values and is evenly distributed in both R and S. Suppose that both relations are not sorted by attribute x.

(a) Suppose the memory buffer has 15 blocks, compute the cost of join using a block-nested loop join.  
(b) Suppose the memory buffer has 15 blocks, compute the cost of join using a sort-merge join.  
(c) Can you estimate the size of the output relation by joining R and S on x. If no, explain why.

ii. Consider a relation R(a, b, c, d) that has a clustering index on a and non-clustering indexes on each of the other attributes. The relevant parameters are:

B(R) = 1000, T(R) = 5000, V (R, a) = 20, V (R, b) = 1000, V (R, c) = 5000, V (R, d) = 500

Give the best query plan and the disk I/O cost for each of the following selection queries:

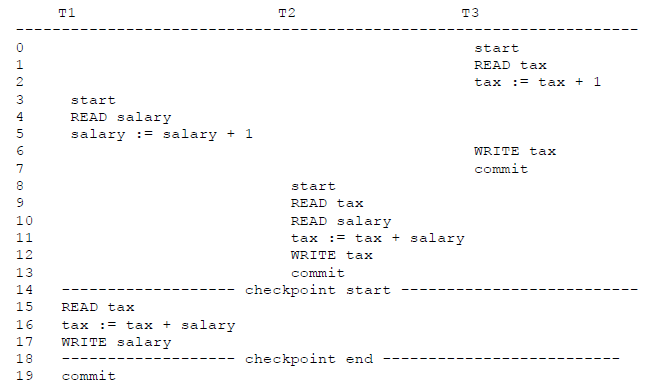
(a) http://202.88.229.62/qeee/coursepack/filedirectory/2323/634/1506432813954.PNG  
(b) http://202.88.229.62/qeee/coursepack/filedirectory/2323/634/1506432815481.PNG

5.[Concurrency Control]. Consider the schedules below:

i. T1:R(X), T2:W(X), T1:W(X), T2:Commit, T1:Commit  
ii. T1:R(X), T2:W(X), T1:W(X), T3:R(X), T1:Commit, T2:Commit, T3:Commit

(a) State which of the following classes it belongs to: conflict serializable, view serializable, recoverable, avoids-cascading-rollbacks.

6.[Recovery]. Consider the schedule given below. There are three transactions, T1, T2, and T3. Initially, the salary = 1 and the tax = 2. The assignments happen within the local memory space of the transactions and the effects of these assignments are not reflected in the database until the WRITE operation.



(a) Show the undo/redo log file entries that would be generated by this execution. For each log entry, indicate what line above generates it.