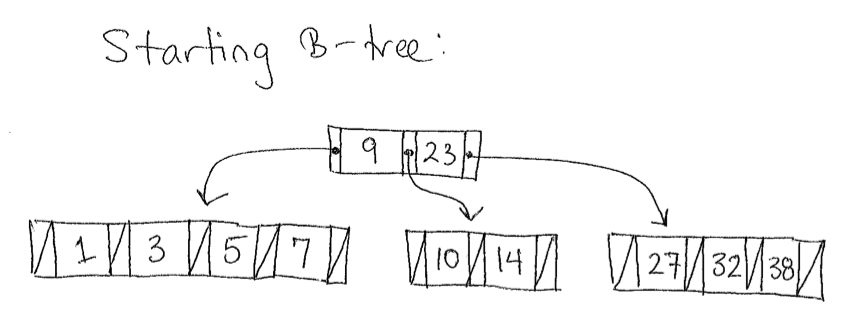
Databases – Spring 2022

Homework 5 – Use for practice & studying

1. Assume you have a B-tree of degree 2 (d=2) as shown in the figure below.   
     
     
   You want to execute the five operations listed below (a through e). Show the B-tree at after each step is completed (you will have one tree per operation). The operations are not independent; they should be done in the order given. **Use only the insertion and deletion algorithms discussed in class**. There are handouts available on the class website for how to do insert & delete.  
   1. Insert 28
   2. Insert 4
   3. Delete 38
   4. Delete 14
   5. Delete 9
2. Consider the relation R(A, B, C, D) with the following dependencies (notice the first two are functional and the last one is multivalued):  
     
   AB → C  
   C → D  
   D →→ B  
     
   Is this relation in 4NF? If it is not in 4NF, decompose it into a set of 4NF relations. Show your work and make sure your decomposition is lossless.
3. Recall the nested loop join algorithm used to join relations R and S on some condition C:  
     
   (1) For each tuple r in R do:  
   (2) For each tuple s in S do:  
   (3) If r and s satisfy C, then   
   (4) Add new tuple (r, s) to the output relation  
     
     
     
   Consider the following three relations:

|  |  |
| --- | --- |
| **Students** | |
| **Rnum** | **Name** |
| 1 | Alice |
| 2 | Bob |
| 3 | Carla |
| 4 | David |

|  |  |
| --- | --- |
| **Courses** | |
| **Cnum** | **Title** |
| 141 | CS I |
| 142 | CS II |
| 172 | Discrete |
| 231 | Comp Org |
| 241 | CS III |

|  |  |
| --- | --- |
| **Enroll** | |
| **Rnum** | **Cnum** |
| 1 | 141 |
| 2 | 142 |
| 2 | 172 |
| 2 | 231 |
| 3 | 172 |
| 3 | 241 |

Consider the query  
  
SELECT Title   
FROM Students NATURAL JOIN Enroll NATURAL JOIN Courses;  
  
Assume none of the relations have indexes, so we can’t use anything other than the nested loop join algorithm.

1. There are two possible join orders for these three tables, assuming we keep the left-to-right sequence of tables the same as the query specifies. What are these two join orders?
2. Assume in the nested loop join algorithm, the “if” test (line 3) takes one unit of time, and the “add new tuple” command also takes one unit of time. For each of the two possible join orders, determine the total amount of time each join order would take to execute the query. Do not consider time for projection, just the joins. Is one query faster than the other?
3. Suppose we want to add a selection operation to this query:  
     
   SELECT Title   
   FROM Students NATURAL JOIN Enroll NATURAL JOIN Courses   
   WHERE Rnum = 1;  
     
   Assume a simple selection algorithm to select all tuples in R that satisfy condition C:  
     
   (1) For each tuple r in R do:  
   (2) If r satisfies C, then   
   (3) Add r to the output relation  
     
   Lines 2 and 3 each take one unit of time. Describe how to optimize the new query to minimize total time, and state how much total time this new query will take. In other words, give a precise ordering of the selection and join operations along with calculating the total time this ordering takes. Show your work as much as possible.