BLG 381E ADVANCED DATA STRUCTURES
MIDTERM - NOVEMBER 21, 2012, 13:30-15:30 PM (2 hours)

1 (5 pt)	**2 (15 pt)**	**3 (30 pt)**	**4 (30 pt)**	**5 (20 pt)**	**Total (100 pt)**

On my honor, I declare that I neither give nor receive any unauthorized help on this exam.

Student Signature:___________________________

Write your name on each sheet.
Write your answers neatly (in English) in the space provided for them.
You <u>must show</u> all your work for credit.
Books and notes are closed.
Good Luck!

Q1[10 points]: Indicator Random Variables

Use indicator random variables to solve the following problem, which is known as the ***hat-check problem***. Each of *n* customers gives a hat to a hat-check person at a restaurant. The hat-check person gives the hats back to the customers in a random order. What is the expected number of customers that get back their own hat?

Q2[20 points]: Red-Black Trees

a) What are the properties of red-black trees (4 pts)

b) Build a red-black tree with the following numbers (Show red nodes with double circle)(6 pts):

[16 7 19 6 12 20 10 13]

c) What are the two modifying operations in Red-Black Trees to update the tree after INSERT or DELETE operations? (2 pts)

d) Insert "9" to the tree that you build in (b). (8 pts)

Q3[20 points]: Augmenting Data Structures

a) What are the four steps of augmenting a data structure? (4 pts)

b) Perform these steps to augment a data structure for interval trees. Develop a new operation INTERVAL-SEARCH*(T, i)*, which finds a node in tree *T* whose interval overlaps interval *i*. If there is no interval that overlaps *i* in the tree, return a pointer to the sentinel *nil*[*T*]. (16 pts)

Hint:

The interval [t1, t2] represents the set {t ∈ **R** : t1 ≤ t ≤ t2}. We can represent an interval [t1, t2] as an object i , with fields low[i] = t1 (the low endpoint) and high[i] = t2 (the high endpoint).

(This page is intentionally left blank.)

Q4[13 points]:

Q4a)[5pts]

You have a Btree containing 100000 keys and with a minimum degree t=10. What is the maximum number of timesteps it would take to search for an item? Assume that a DISK-READ takes 10 timesteps and all in-memory single step operations (like assignment, comparison) take 1 timestep.

Hint: if minimum degree is t, then there are at least t-1 and at most 2t-1 keys in a node.

ANSWER 4a)

In order to find the maximum number of steps, we need to consider the B-tree height being maximum with the given number of keys ($n=10^5$). B-Tree height will be maximum when every node has the smallest possible number of keys, which is 10-1=9. The number of children will, therefore, be 10. The tree height will be 4 and including the root level there will be 5 levels.

- *depth 0: 1 node 9 keys*
- *depth 1: 10 nodes, 90 keys*
- *depth 2: 100 nodes, 900 keys*
- *depth 3: 1000 nodes, 9000 keys*
- *depth 4: 10000 nodes, 90000 keys*
- *depth 5: 1 nodes, 1 keys ←this is impossible, because a node in level 4 would have <t-1 children. Therefore, there can not be any nodes at this level.*

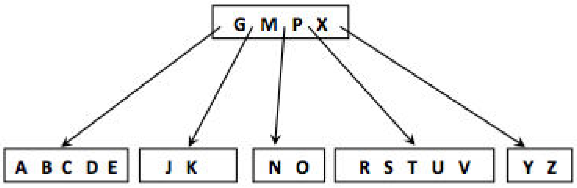
*For each level there will be a DISK-READ operation (5*10), within each level there will be a search for the child pointer, which will take 9 time steps and 4 more additional comparison/assignment type operations. Therefore the maximum number of timesteps will be: 13*5 + 10*5 = 23*5 = 115timesteps.*

Q4b)[8pts]

Insert an element with key "F" into the following B-tree with minimum degree t=3.

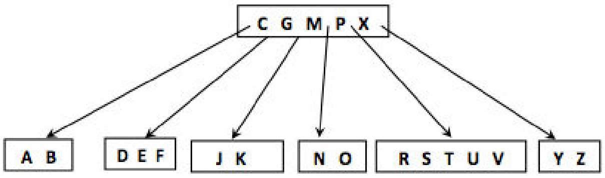
Then delete the element with key "J".

Show the resulting B-tree after each operation.

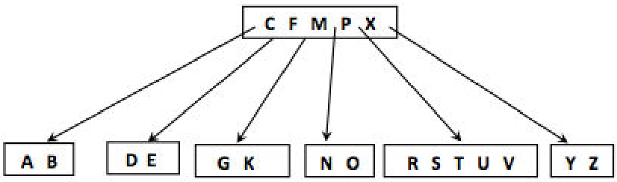


ANSWER4B:

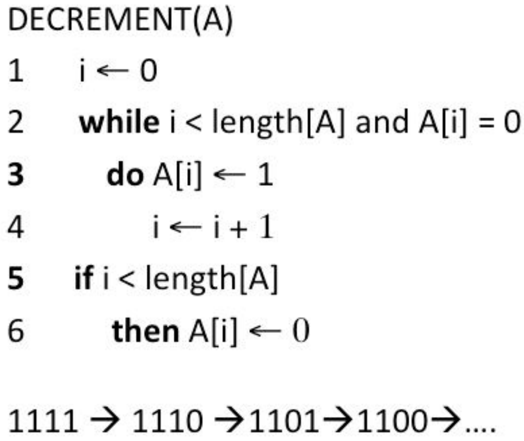
Insert F:



Delete J:



Q5[17 points] (Amortized Analysis)
You have a 4 bit down counter that you will decrement from 1111.
Q5a) [8pts] Write down the pseudocode for the decrement operation.
Q5b) [9pts] What is the amortized cost of one decrement operation among a sequence of k decrement operations?
ANSWER5a)



```
DECREMENT(A)
i ← 0
while i < length[A] and A[i] = 0
   do A[i] ← 1
      i ← i + 1
if i < length[A]
   then A[i] ← 0
```

1111 → 1110 →1101→1100→....

ANSWER5b)

In DECREMENT there is only one 1→0 bit flip and there are some 0→1 bit flips.

For every 1→0 bit flip, we can pay 2TL, we can use 1TL for the 1→0 bit flip and the remaining 1TL when the bit needs to be flipped back to 0. Therefore for k decrement operations we pay 2k TL. Amortized cost of one decrement operation is 2k/k = 2.

Q7) [8pts] Medians and Order Statistics

Write down the fastest algorithm that you can write to compute the minimum (i.e. the 1st order statistics) in an array of size n (note: your algorithm must be faster than O(n)).

ANSWER7:

Although RANDOMIZED-SELECT may be the first algorithm that comes to mind, finding the minimum has expected time complexity of Θ(n). However, if the array is kept in a heap, then minimum can be found in O(logn) time.

Q6) [6pts] Hashing

Q6a)[6pts] Insert the following elements into a hash table of size 11 using open addressing with double hashing. A = [10, 22, 34, 5]. Write down your hash function clearly and show the collisions you got for each insertion.

ANSWER7:

*You can use h(k,i)=(h'(k)+i*h''(k))mod 11 where h'(k)=k mod a and h''(k)=k mod b, and a, b and m are relatively prime. We could choose, for example, b=13, a=19*

h(10,0)=((10 mod 19)+ 0) mod 11 = 10

h(22,0)=((22 mod 19)+ 0) mod 11 = 3

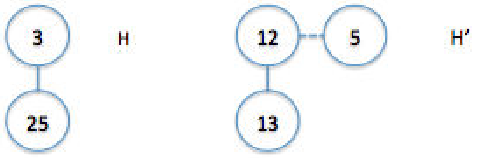
h(34,0)=((34 mod 19)+ 0) mod 11 = 15

h(5,0)=((5 mod 19)+ 0) mod 11 = 5

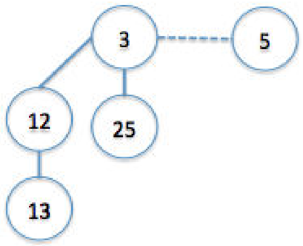
There are no collisions.

Q8) [6pts] Binomial Heap

Two binomial heaps H and H' contain 2 and 3 elements respectively. Show the binomial heap which is the union of H and H'.



ANSWER8:



(This page is intentionally left blank.)