EL9343 Midterm Exam (2021 Spring)

Name: **ID:**

March 28, 2021

Write all answers on your own blank answer sheets, scan and upload answer sheets to newclasses at the end of exam, keep your answer sheets until the grading is finished.

Multiple choice questions may have multiple correct answers. You will get partial credits if you only select a subset of correct answers, and will get zero point if you select one or more wrong answers.

1. (**20 points**) **True or False (2 points each)**

   (a) **T or F:** Universal hashing functions can eliminate the worst-case scenario where all keys are hashed to the same slot;

   (b) **T or F:** Heap-sort is not stable;

   (c) **T or F:** Inorder tree walk of a binary search tree will print out all keys in sorted order;

   (d) **T or F:** Height of a binary search tree with n elements is $\Theta(\log n)$;

   (e) **T or F:** $3^n = \omega(2^n)$;

   (f) **T or F:** In Randomized QuickSort, if the array has n elements, the probability for any pair of elements to be compared during the first partition call is $\frac{1}{n}$.

   (g) **T or F:** It is impossible to develop a comparison sorting algorithm with $o(n \log n)$ complexity, where $o(\cdot)$ stands for little-o.

   (h) **T or F:** In a rooted tree, all nodes at the same level have the same height;

   (i) **T or F:** With simple uniform hashing, to store n keys in a hash table with m slots, the average number of keys hashed to any slot is $\frac{n}{m}$;

   (j) **T or F:** Merge-sort is in-place;

++++++ continue on next page ++++++

2. (**4 points**) Given a max-heap with height 8, if all the elements are distinct, the 3-rd largest element can be at height of:

(a) 1; (b) 5; (c) 6; (d) 7; (e) 8.

3. (**4 points**) If $f(n) = \Omega(g(n))$, and $z(n) = \omega(g(n))$, which of the following can be true?

(a) $f(n) = O(g(n))$;

(b) $f(n) = \Theta(z(n))$;

(c) $g(n) = \Theta(z(n))$;

(d) $f(n) = o(g(n))$;

(e) $f(n) = \omega(z(n))$;

4. (**4 points**) Which of the following about binary search tree is true?

(a) when deleting a node z with two children, we can replace z by its predecessor;

(b) in an AVL tree, the left sub-tree and right sub-tree of any node have the same height;

(c) v's successor can be the right child of v's left child;

(d) a new node is always inserted as a leaf node;

(e) the worst-case search time in an AVL tree with n nodes is $O(\log n)$;

5. (**4 points**) When handling collisions using chaining, which of the following statements about a hash table of m slots and stored n elements are true?

(a) The linked list at each slot is sorted;

(b) Worst-case time to search is $\Theta(n)$;

(c) Worst-case insertion complexity is always $\Theta(1)$;

(d) To have an average-case search time of $O(1)$, n must be $O(m)$;

(e) In universal hashing, a new hash function is picked randomly each time when inserting;

6. (**4 points**) Which of the following function pairs satisfy $f(n) = \Omega(g(n))$?

(a) $f(n) = 2n$ and $g(n) = \log n$;

(b) $f(n) = n^k$ and $g(n) = k^n$, where $k > 1$ is a constant;

(c) $f(n) = n$ and $g(n) = \sqrt{n}$;

(d) $f(n) = n \log n$ and $g(n) = 3n + \log n$;

(e) None of the above

++++++ **continue on next page** ++++++

7. (**12 points**) Solve the following recurrences:

   (a) (**4 points**) Use the iteration method to solve $T(n) = T(\frac{n}{3}) + T(\frac{n}{4}) + 5n$;

   (b) (**4 points**) Use the substitution method to verify your solution for Question 7a.

   (c) (**4 points**) Solve the recurrence $T(n) = 3T(\sqrt[3]{n}) + \log n$.

8. (**8 points**) When we build an AVL tree from scratch, keys are inserted in the order of:

$$40 \to 30 \to 20 \to 29 \to 24 \to 27,$$

   Please plot the AVL tree after each key is inserted, and mark the type of rotation taken, if any, at each step.

9. (**8 points**) If heap-sort is applied to an array of $[2, 7, 26, 25, 19, 17, 1, 90, 3, 36, 15]$,

   (a) (**4 points**) plot the initial max-heap built from the array;

   (b) (**4 points**) plot the max-heap after the third-maximum is extracted.

10. (**10 points**) The pseudocode of SELECTION-SORT is shown below.

```
Alg.: SELECTION-SORT(A)
 n <-- length[A]
 for j <-- 1 to n - 1
     do smallest <-- j
          for i <-- j + 1 to n
              do if A[i] < A[smallest]
                 then smallest <-- i
     exchange A[j] <-- A[smallest]
```

   (a) (**4 points**) How many key comparisons does SELECTION-SORT do? What arrangement of keys incurs the largest number of key exchanges for SELECTION-SORT? What arrangement of keys incurs the least number of key exchanges?

   (b) (**6 points**) Prove the correctness of SELECTION-SORT by applying Loop-Invariant to the outer loop only.

11. (**10 points**) Given an unsorted array $A[1 \cdots n]$,

   (a) (**4 points**) Construct an $\Theta(n)$ algorithm to find the pair of indexes (i^*, j^*), $1 \le i^* \le j^* \le n$, such that $|A[i^*] - A[j^*]|$ is the maximum among all the possible (i, j) pairs with $1 \le i \le j \le n$, where $|\cdot|$ stands for the absolute value. Write down pseudocode, analyze its complexity.

   (b) (**6 points**) Construct an $O(n \log n)$ divide-and-conquer algorithm to find the pair of indexes (i^*, j^*), $1 \le i^* \le j^* \le n$, such that $A[i^*] - A[j^*]$ is the maximum among all possible (i, j) pairs with $1 \le i \le j \le n$. Write down pseudocode, analyze its complexity.

++++++ continue on next page ++++++

12. (**12 points**) We showed that there is a worst-case linear time SELECT algorithm that selects the i-th order statistic from n numbers with $T(n) = \Theta(n)$ comparisons. Now suppose there are n min-heaps, each min-heap contains m numbers ($1 \leq m \leq n$), develop an algorithm that selects the i-th ($1 \leq i \leq n$) order statistic from the combined nm numbers from all min-heaps with the complexity of $\Theta(n + i(\log i + \log m))$.

(a) (**8 points**) describe the main steps of your algorithm (no need to write down the detailed pseudo code);

(b) (**4 points**) analyze the complexity of your algorithm.