香 港 中 文 大 學
The Chinese University of Hong Kong



**Course Examinations Second Term, 2020-2021**

**Course Code & Title** : **CSCI2100A/ESTR2102 Data Structures**

**Time allowed** : **1** **hours** **40** **minutes**

**Student I.D. No.** : ..... **Course Code:** .....

---

Please note the following:

1. I will abide with the following rules during the examination time. When I submit my answer script, my consent to the following declaration is automatic. Otherwise, I will not submit my answer script.
   (a) I will not cheat and will not take any unfair advantage of my classmates during the examination.
   (b) Other than what the teacher and/or invigilator(s) have specified, I will not access the Internet or other electronic storage devices to retrieve information related to the course materials during the examination.
   (c) Except with the teacher and/or invigilator(s), I will not discuss any questions with anyone during the examination.
   (d) I confirm I am the person with the identity number shown on the screen, and I am the same person taking this examination.
2. Guidelines on using Zoom
   (a) Make sure your video is on and audio is muted.
   (b) If you want to ask question, please use "private" chat on Zoom so that it is not being broadcasted to everyone.
   (c) Make sure you use your CUHK Student ID number, i.e., "123456789" as your screen name in the participant list so that we can identify who you are.
3. The written midterm is a close-book and open-notes examination. Your notes should fit into one sheet of A4 paper (two sides).
4. No calculator is needed. Turn off your phone.
5. For ESTR2102 students, please put "ESTR2102" after your student ID number on the answer script.

6. For regular students, answer questions only up to 6(h). Your maximum score is 100. You will not receive any additional bonus points for answering questions for ESTR2102 students.
7. For ESTR2102 students, answer all questions, your maximum score is 120.
8. You have 100 minutes on this 100/120 points examination. This means that you should spend about 1 minute for each 1 point.
9. Please note the following about your answer script.
   (a) Write legibly. Anything we cannot decipher will be discarded.
   (b) Put your student ID on top of each answer sheet to identify your work clearly.
   (c) There are three main ways to assemble your answer book.
      i. You can use the pdf file provided and enter your answers directly with software such as Preview (Mac) or Adobe Acrobat (Mac and PC).
      ii. You can use MS Word or other equivalent word processors to enter your answers.
      iii. You can write on clean sheet of papers and then take (or scan) photos of these written pages.
   (d) Submit only one file, e.g., pdf, Word, or zip file, so if you want to submit multiple files, compress them into a single ".zip" file and then submit the ".zip" file.
   (e) The file size should not exceed 25M bytes.
   (f) Submit the one file to **csci2100a.estr2102@gmail.com** within **10 minutes** after the exam is finished.
   (g) **Email Subject:** course code_SID_name_WrittenMidterm, e.g.,
   csci2100a_1155088237_CHEN_Wang_WrittenMidterm;
   estr2102_1155088237_CHEN_Wang_WrittenMidterm;
   (h) **File Name:** follow the same format with the subject, e.g.,
   csci2100a_1155088237_CHEN_Wang_WrittenMidterm.pdf;
   estr2102_1155088237_CHEN_Wang_WrittenMidterm.pdf;
10. My declaration.
   (a) I, with the SID of declare that I will not and have not taken any unfair advantage of my fellow class mates by any illicit behaviors, such as discussing the answer with anyone, cheating by copying others' work, access Internet or electronic materials illegally, etc.
   (b) By signing below, I declare that I have done everything according to the code of conduct and professional practice with utmost integrity.
   (c) If I were to be found violating the policy and guidelines set forth, I shall take the full responsibility, which can mean a fail for the examination, a fail of the course, a demerit, termination of study at CUHK, or any combination of the above.
   (d) Signature: Date:

1. (18) On Program Analysis

   You are given the task to analyze two data structures, $D_1$ and $D_2$, and their related algorithms to solve a problem with $n$ input elements. Now, $D_1$ and $D_2$ have the following time complexity for three different functions, $f_1$, $f_2$, and $f_3$ with their Best, Average, and Worst time complexity respectively as follows:

   Table 1: Time Complexity for $D_1$ and $D_2$

| | $D_1$ | $D_2$ |
|---|---|---|
| $f_1$ | | |
| Best | $4n$ | $3\log n$ |
| Average | $3n^2+2$ | $7n$ |
| Worst | $5n^2\log n$ | $12n(\log n)^2$ |
| $f_2$ | | |
| Best | $3\log n$ | $3n\log 2n$ |
| Average | $4\log n^2+2n+5n^2$ | $4\log n^2+9n$ |
| Worst | $3n!$ | $4n^7+2n^5+5n^2+1$ |
| $f_3$ | | |
| Best | $\log n$ | $350$ |
| Average | $11n$ | $4n^{1.7}$ |
| Worst | $2n\log 4n+n\sqrt{3n}$ | $n^2+3(\log n)^2$ |

   (a) (9) Draw another table of the same size and same format but fill each entry with the corresponding time complexity in the Big-O notation from Table 1.

   (b) (3) Estimate for what $n$ is the interception point for the Average case of $f_2$ for both $D_1$ and $D_2$. In other words, what is the integer $n$ where $f_2^{D_1}(n-1) \leq f_2^{D_2}(n-1)$ and $f_2^{D_1}(n) \geq f_2^{D_2}(n)$ for the Average time complexity.

   (c) (6) Use only the Worst Big-O time complexity to order $f_1$, $f_2$, and $f_3$ for $D_1$ and $D_2$ respectively. (For instance, for $D_1$, if we use the Average Big-O time complexity to order $f_1$, $f_2$, and $f_3$, it will be $f_3 \leq f_1 = f_2$.) You should have one ordering for $D_1$ and another one for $D_2$.

2. (15) On Stacks and Queues

   (a) (6) Given the initial sequence of $3, 6, 5, 2, 1, 7, 4$, your target is to judge this sequence can be changed to another final sequence using a stack or a queue. Note: at each time you can either (1) push an element from the initial sequence into the stack/queue, or (2) directly output an element from the initial sequence, or (3) pop an element from the stack/queue and output it. Once the element is output, it cannot be back again. Fill the blanks in the following table with **Stack**, **Queue**, **Both** or **None**. **Both** means either a stack or a queue can be used to obtain the final sequence. **None** means neither a stack nor a queue can be used to obtain the final sequence.

   (b) (4) Given the initial sequence in (a), is there any permutation(s) of it which cannot be generated by either a stack or a queue? If yes, please give an example; otherwise, justify your answer.

Table 2: Final Sequences

| | *Final Sequence* | *Data Structure* |
|---|---|---|
| (1) | $6, 7, 4, 1, 2, 5, 3$ | |
| (2) | $3, 5, 6, 1, 2, 4, 7$ | |
| (3) | $2, 3, 6, 5, 4, 1, 7$ | |

(c) (5) Now you are asked to implement a queue using two stacks, please specify your steps to implement `enqueue(x)` and `dequeue()` functions for a queue using pseudocode. You are only allowed to use the standard operations of a stack, i.e., `push(x)` and `pop()`.

3. (15) On Trees

(a) (4) Draw a complete and full binary tree with height = 2. Using pre-order traversal, label all the nodes in the order of traversal with unique integers, 1, 2, $\cdots$ (so the traversal will result in the unique integer sequence).

(b) (5) Using the labels in (a), what are the leaf nodes in the above tree? What are the output sequences of in-order and post-order traversal?

(c) (3) In (a), if we change the height to $n$ and use in-order traversal, what is the label of the root of the complete binary tree.

(d) (3) Given a binary tree with integer as labels, if the pre-order and in-order output are 1, 3, 2, 5, 4 and 3, 5, 2, 1, 4 respectively, please draw the tree.

4. (20) On AVL Trees

(a) (10) Show the result of inserting $12, 14, 11, 15, 19, 16, 13, 17$ into an empty AVL tree. Show all intermediate steps and indicate whether no rotation, a single rotation, or a double rotation has been performed in each step.

(b) (3) What is the height of the resulting AVL tree?

(c) (3) What is the depth of the node with the label "11" in the above resulting AVL tree?

(d) (4) What is the upper bound and lower bound for the number of nodes in an AVL tree with height $n$?

5. (12) On Heaps

(a) (8) Show the result of inserting $20, 15, 22, 12, 11, 9, 17, 16, 18, 13$ one at a time, into an initially empty binary max heap (larger number has higher priority). Show all intermediate steps and the final configuration of the heap using the array data structure.

(b) (4) Show the result of performing two `Delete_Max` operations in the above heap. Show the heap configuration after each `Delete_Max` operation clearly.

6. (20/28) Short Answers (Please give only concise and short answers!)

(a) (4) Give a closed-form solution for the following recurrence relation:

$$T(n) = 3T(n-1) + 2, T(1) = 1. \quad (1)$$

(b) (2) Explain in a few sentences what is "divide and conquer" method.

(c) (4) Analyze the following `for` loop statements and give (1) the precise $f(n)$ with $n$ being the input ($n \geq 1$) and (2) the Big-O notation of $f(n)$. (You may interpret $f(n)$ as the number of times that statement "`x = x + 2`" is executed.)

```
int f(int n){
    int x = 0;
    for i = 1 to n
        for j = 1 to i
            for k = j to n
                x = x + 2;
    return x;
    }
```

(d) (2) Rewrite the following expression in the postfix order to the prefix order: A B * C D * E F * / +.

(e) (2) What is the minimum and maximum number of nodes in a heap of height 7?

(f) (2) What is the major difference between binary search trees and min-heaps?

(g) (2) True or false: Every binary heap is a complete binary tree.

(h) (2) Is a heap always an AVL tree? Justify your answer.

(i) (**For ESTR2102**) (3) There are $n$ people currently in a room, give an expression for the probability of having two people of the same birthday (not necessary to be the same birth year). Assume we have 365 days in a year.

(j) (**For ESTR2102**) (5) Prove by induction on $n \geq 1$ that $\sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6$.

7. (**For ESTR2102**) (4) Given a series of union operations, please fill in the final array after applying quick-union method. The initial array (the root of each node is itself) is the following:

Table 3: Initial and Final Arrays

| id[] | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| final id[] | | | | | | | | | | |

The series of union operations (Note: Union($p$, $q$) means to set the parent of $p$'s root to $q$'s root, no path compression here.): Union(8, 9); Union(7, 5); Union(1, 2); Union(0, 1); Union(2, 7); Union(9, 0); Union(3, 4); Union(6, 5); Union(3, 8).

8. (**For ESTR2102**) (8) Using Deterministic Finite Automata to find all the occurrences of "abbaab" in "aaabbaabbabaaba".

(a) (4) Construct an automaton for "abbaab" and write down the transition function table. The alphabet is $\{a, b\}$.

(b) (4) Fill in the following Table 4.

**-End-**

Table 4: States

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $T[i]$ | $a$ | $a$ | $a$ | $b$ | $b$ | $a$ | $a$ | $b$ | $b$ | $a$ | $b$ | $a$ | $a$ | $b$ | $a$ |
| State | | | | | | | | | | | | | | | |