Written re-exam

MED8

Algorithms, Data Structures and Software Engineering for Media Technology

Tuesday 23 October 2018

Name:_____

Cpr.no.: _____

Study no.: _____

Aalborg University | MSc in Medialogy | Med8

Algorithms, Data Structures and Software Engineering for Media Technology

Extra-ordinary Re-Examination

23 October 2018

Instructions

- You have 3 hours to complete this examination.
- Neither electronic devices nor written material are allowed in the examination room.
- This examination consists of 10 questions. Each question is worth 10 marks. You must obtain at least 50 marks to pass.
- Do not write any answers on this question paper—answers written on the question paper will be ignored by the examiner. Write all your answers on the writing paper provided.
- Do not write your answers in pencil and do not use a pen with red or green ink. Use a pen with blue or black ink.
- Hand in no more than one answer to each question.
- Do not turn over until you are told to do so by the invigilator.

Question 1

For each of the following equations, state whether it is true or false.

a) $5n^4 + 3n\log_3(n^8) + 100 = \Theta(n^4)$

b) $2n^3\log_2(n^5) = \Theta(n^4)$

c) $n^3\log_2 n = O(n^4)$

d) $3^n = O(n^3)$

e) $9n^3 = \Omega(n^3)$

f) $n^5 + 12n = o(n^5)$

g) $n\log_2(12n^4) = o(n^2\log_2 n)$

h) $n^3 + 19n^2 + n + 4n\log_2 n + 13 = \omega(n^2\log_2 n)$

i) $\left(\sqrt{n}\right)^5\log_2 n = \omega(n^2)$

j) $2n = \Omega\left(\sqrt{n}\right)$

[1 mark for each correct part]

Question 2

The Master Theorem is stated as follows:

Theorem 4.1 (Master theorem)

Let $a \geq 1$ and $b > 1$ be constants, let $f(n)$ be a function, and let $T(n)$ be defined on the nonnegative integers by the recurrence

$$T(n) = aT(n/b) + f(n),$$

where we interpret n/b to mean either $\lfloor n/b \rfloor$ or $\lceil n/b \rceil$. Then $T(n)$ has the following asymptotic bounds:

1. If $f(n) = O(n^{\log_b a - \epsilon})$ for some constant $\epsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$.
2. If $f(n) = \Theta(n^{\log_b a})$, then $T(n) = \Theta(n^{\log_b a}\lg n)$.
3. If $f(n) = \Omega(n^{\log_b a + \epsilon})$ for some constant $\epsilon > 0$, and if $af(n/b) \leq cf(n)$ for some constant $c < 1$ and all sufficiently large n, then $T(n) = \Theta(f(n))$. ■

Given the Master Theorem, as stated above, write down the order of growth in terms of Θ notation for each of the following recurrences.

a) $T(n) = 25T(n/5) + n\log_2 n$

b) $T(n) = 8T(n/2) + 4n^3$

c) $T(n) = 4T(n/2) + \left(\sqrt{n}\right)^5$

d) $T(n) = 3T(n/9) + \sqrt{n}$

e) $T(n) = 9T\left(n/\sqrt{3}\right) + 3n^3$

[2 marks for each correct part]

Question 3

a) Which **one** of the following statements is true?
   A. The POP operation on a stack removes and returns the last element saved in the stack.
   B. The PUSH operation on a stack can lead to stack underflow.
   C. The ENQUEUE operation on a queue can lead to queue underflow.
   D. The DEQUEUE operation on a queue removes and returns the element at the tail of the queue.

b) Which **one** of the following statements is true?
   A. It takes constant time to search for a given key in a doubly-linked list.
   B. It takes constant time to find the minimum key in an unsorted doubly-linked list.
   C. It takes constant time to insert a new element at the head of a doubly-linked list.
   D. Given a pointer to a specific element in a doubly-linked list, it then takes linear time to delete that element from the list.

c) Which **one** of the following statements is true?
   A. It takes $\Theta(n \log_2 \mathrm{n})$ time to carry out an inorder tree walk on a binary search tree.
   B. It takes $O(h)$ time to search for a given value in a binary search tree of height h.
   C. It takes $\Theta(n)$ time on average to find the minimum value in a binary search tree.
   D. It takes constant time to insert an element into a binary search tree.

d) Which **one** of the following statements is true?
   A. The worst-case running time of heapsort and merge sort is $\Theta(n)$.
   B. The minimum of an unsorted array of n elements can be found in constant time.
   C. The ith order statistic of an unsorted array of n elements can be found in $\Theta(1)$ time using a randomized algorithm.
   D. The ith order statistic of an unsorted array of n elements can be found in $O(n \log_2 n)$ time by using a comparison sort and then indexing the ith element in the output array.

e) Which **one** of the following statements is true?
   A. A binary heap is always a complete binary tree.
   B. In a heap, the index of the parent of a node whose index is i can be found by shifting the binary representation of i by one bit to the right.
   C. The max-heap property is that the key at each node is greater than or equal to the key of its parent.
   D. The optimal worst-case running time of building a max heap from an unordered input array is $O(n \log_2 n)$ time.

[2 marks for each correct part]

Question 4

Study the following program and answer the questions that follow it.

```
1  #include <stdio.h>
2  #include <stdlib.h>
3  #define N 3
4
5  __global__ void addVectors(float* dev_a, float* dev_b, float* dev_c) {
6      int i = blockIdx.x;
7      if (i < N) {
8          dev_c[i] = dev_a[i] + dev_b[i];
9      }
10 }
11
12 int main(void) {
13     float a[N] = {1,2,3}, b[N] = {4,5,6}, c[N];
14     float *dev_a, *dev_b, *dev_c;
15
16     cudaMalloc((void**)&dev_a, N*sizeof(float));
17     cudaMalloc((void**)&dev_b, N*sizeof(float));
18     cudaMalloc((void**)&dev_c, N*sizeof(float));
19
20     cudaMemcpy(dev_a, a, N*sizeof(float),cudaMemcpyHostToDevice);
21     cudaMemcpy(dev_b, b, N*sizeof(float),cudaMemcpyHostToDevice);
22
23     addVectors<<<N,1>>>(dev_a, dev_b, dev_c);
24
25     cudaMemcpy(c, dev_c, N*sizeof(float), cudaMemcpyDeviceToHost);
26
27     cudaFree(dev_a);
28     cudaFree(dev_b);
29     cudaFree(dev_c);
30
31     for(int i = 0; i < N; i++)
32         printf("%.0f ", c[i]);
33
34     return EXIT_SUCCESS;
35 }
```

a) What does this program print to the console?
b) In this program, how many thread blocks are there in each grid and how many threads are there in each block?
c) Is the array, a, (defined in line 13) stored in GPU memory or CPU memory?
d) Does the pointer dev_a point at a region of device memory or host memory?
e) In which lines is the GPU memory deallocated?

[2 marks for each correct part]

Question 5

a) What is the worst-case running time of quicksort and under what conditions does this worst case occur?
b) What is the expected running time of quicksort?
c) What is the best-case running time of insertion sort and under what conditions does this best case occur?
d) What is the worst-case running time of insertion sort and under what conditions does this worst case occur?
e) What is the average-case running time of bucket sort, assuming the input is drawn from a uniform distribution?

[2 marks for each correct part]

Question 6

In the context of machine learning, provide a brief definition of each of the following terms:

a) supervised learning model
b) *k*-fold cross validation
c) ensemble method
d) bagging
e) boosting

[2 marks for each correct part]

Question 7

a) Explain why a single perceptron cannot implement an XOR logical gate. [3 marks]
b) Briefly explain the backpropagation learning algorithm. [3 marks]
c) In the context of evolutionary algorithms, explain what is meant by the terms *mutation* and *crossover*. [4 marks]

Question 8

a) For each of the following components, state whether it is the user or the service-provider that has responsibility for providing the component in a PaaS model:
    i. Application
    ii. Data
    iii. Operating system
    iv. Storage
b) Explain the meaning of the term *multi-tenancy* in the context of cloud computing.
c) Explain the meaning of the terms *vertical scaling* and *horizontal scaling* in the context of cloud applications.
d) What is a Content Delivery Network (CDN) and what are the benefits of using one?
e) Explain the meaning of the term *polyglot persistence* in the context of cloud computing.

[2 marks for each part, a-e]

Question 9

a) What is the output printed to the console by the following Python script:

```
a = [[1,2,3,4],[5,6,7,8],[9,10,11,12]]
print(a[0][2])
print(a[0][2:])
print(a[:1][:])
print(a[-1][-2:])
```

b) Give two examples of distance measures that can be used to measure similarity in a *k*-nearest neighbour classifier.

c) Explain the meaning of the term *collision* in the context of a hash table.

d) Explain how chaining can be used to handle collisions in a hash table.

e) What is the expected (average) running time for the dictionary operations in a hash table?

[2 marks for each part, a-e]

Question 10

Write a program in Python that takes a list of numbers as input and outputs the maximum and minimum values in the list. The program should use fewer than $2n$ comparisons.

[10 marks]

END OF EXAMINATION