Ordinary Written exam

MED8

Algorithms, Data Structures and Software Engineering for Media Technology

Wednesday June 13th 2018

Name:_____

Cpr.no.: _____

Study no.: _____

Aalborg University | MSc in Medialogy | Med8

Algorithms, Data Structures and Software Engineering for Media Technology

Ordinary Examination

13 June 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. The maximum number of marks that can be obtained is 50. You must obtain at least 25 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) $3n^2 + 2n + 100 = \Theta(n^2)$

b) $4n \log(n^2) = \Theta(n \log n)$

c) $4n^2 = O(n^3)$

d) $2^n = O(n^2)$

e) $100^{\log n} = \Omega(2^n)$

f) $0.0001n^3 = \Omega(n^3)$

g) $4n^2 = o(n^2)$

h) $3n \log n = o(n^2 \log n)$

i) $9n^3 + 3n^2 + 2n + 5n \log n + 1000000 = \omega(n^3)$

j) $4n^2 \log n = \omega(n^2)$

[1/2 mark for each correct part]

Question 2

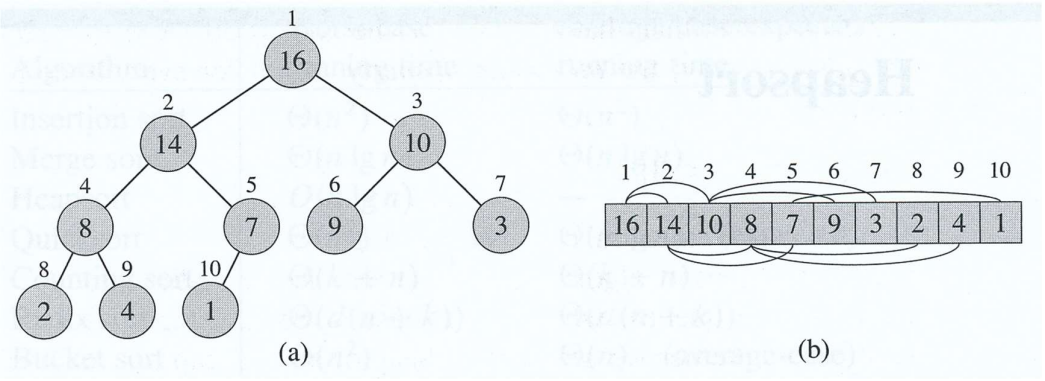

Figure (a) above shows a binary tree and figure (b) shows a representation of this tree as an array (1-based indexing is used).

a) State the max-heap property. [1 mark]

b) Is the tree in Figure (a) above a max-heap? [1 mark]

c) Suppose *n* is a node in a binary tree and let *i* be the index of node *n* in an array representation of the tree, when the tree is represented in the same manner as that in which the array in Figure (b) above represents the binary tree in Figure (a). Write down algorithms that take *i* as input and efficiently compute the index in the array for (1) the parent of *n*, (2) the left child of *n* and (3) the right child of *n*. For each of the three algorithms, explain how it can be implemented efficiently using bit-shifting. [3 marks]

Question 3

a) Write down the operations that a data structure must support if it is to function as a *dictionary*. [1 mark]

b) What is the worst-case running time for searching for an element in a hash table? Under what circumstances does this worst case occur. [2 marks]

c) What is the expected running time for searching for an element in a hash table? Under what conditions does this expected running time occur? [2 marks]

Question 4

Study the following code example 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) In which line or lines is memory allocated on the GPU? [1 mark]

b) How many blocks of threads is the kernel run on? [1/2 mark] How many threads are there in each block? [1/2 mark]

c) What is the purpose of line 25? [1 mark]

d) In which line or lines is the kernel defined and how do we indicate to the compiler that the function is to be used as a GPU kernel? [1 mark]

e) What is the purpose of the variable blockIdx, used in line 6 and what is its type? [1 mark]

Question 5

a) What is IaaS, PaaS and SaaS? Provide some brief descriptions and examples.
b) *ssh -i file.pem ec2-user@{IP_Address}*
After setting up an AWS EC2 server, what is the use of the above shell command? What is the use of "file.pem"?
c) What is the python library Boto?
d) With which Amazon service you can create the AWS Access Key ID and AWS Secret Access Key?
e) What is the advantage of using a Content Delivery Network (CDN)?

[1 mark for each correct part]

Question 6

a) What is the difference between Amazon's EC2 and Lambda?
b) What is Multi-tenancy in software architecture?
c) Aiming at sizing the infrastructure for better performance, how can you do vertical scaling and how can you do horizontal scaling?
d) What is a container and which tool you can use for this?
e) What is polyglot persistence?

[1 mark for each correct part]

Question 7

a) What are semantic whitespaces in Python? [1 mark]

b) Which libraries, for Python, are used for working with multidimensional arrays? [1 mark]

c) In the following code, assume that **a** is an array. Explain what elements of the array each of the three lines would return. [1 marks]

```
a[-2:-1]
a[-10:]
a[10:]
```

Question 8

a) Explain what structured data is? [1 mark]

b) Explain what unstructured data is? [1 mark]

c) In the Python library **Scikit**, what is an estimator object? Name at least one specific class. [1 mark]

d) In the Python library **Scikit**, what is a transformer object? Name at least one specific class. [1 mark]

e) In the Python library **Scikit**, what are the arguments to create a machine learning pipeline? [2 marks]

Question 9

a) Explain the concept of **ensemble learning**. [2 marks]

b) How does **bootstrap aggregation** work? [2 marks]

c) How does **boosting** work? [2 marks]

Question 10

Study the following code and answer the questions that follow.

```
from keras.models import Sequential
from keras.layers import Dense

model = Sequential()
model.add(Dense(128, input_dim=128, activation='relu'))
model.add(Dense(64, activation='relu'))
model.add(Dense(128, activation='relu'))
model.add(Dense(32, activation='relu'))
model.add(Dense(2, activation='softmax'))

model.compile(optimizer='sgd', loss='binary_crossentropy')
```

a) How many layers are in the neural network **model?** [1 mark]

b) How many units does the last hidden layer contain? [1 mark]

c) What algorithm is used to train the neural network? [1 mark]

d) What error measure is used for optimizing the weights of the neural network? [1 mark]

e) What is the activation function of the output layer? [1 mark]

END OF EXAMINATION