



This exam paper must not be removed from the venue

Venue	_____
Seat Number	_____
Student Number	_____
Family Name	_____
First Name	_____

## School of Information Technology and Electrical Engineering EXAMINATION

Semester One Final Examinations, 2019

### INFS7901 Database Principles

*This paper is for St Lucia Campus students.*

Examination Duration: 120 minutes

**For Examiner Use Only**

Reading Time: 10 minutes

Question	Mark
1	
2	
3	
4	
5	
6	
7	
8	
9	

**Exam Conditions:**

This is a School Examination

This is a Closed Book Examination - specified materials permitted

During reading time - write only on the rough paper provided

This examination paper will be released to the Library

**Materials Permitted In The Exam Venue:**

**(No electronic aids are permitted e.g. laptops, phones)**

Calculators - Any calculator permitted - unrestricted

One A4 sheet of handwritten or typed notes double sided is permitted

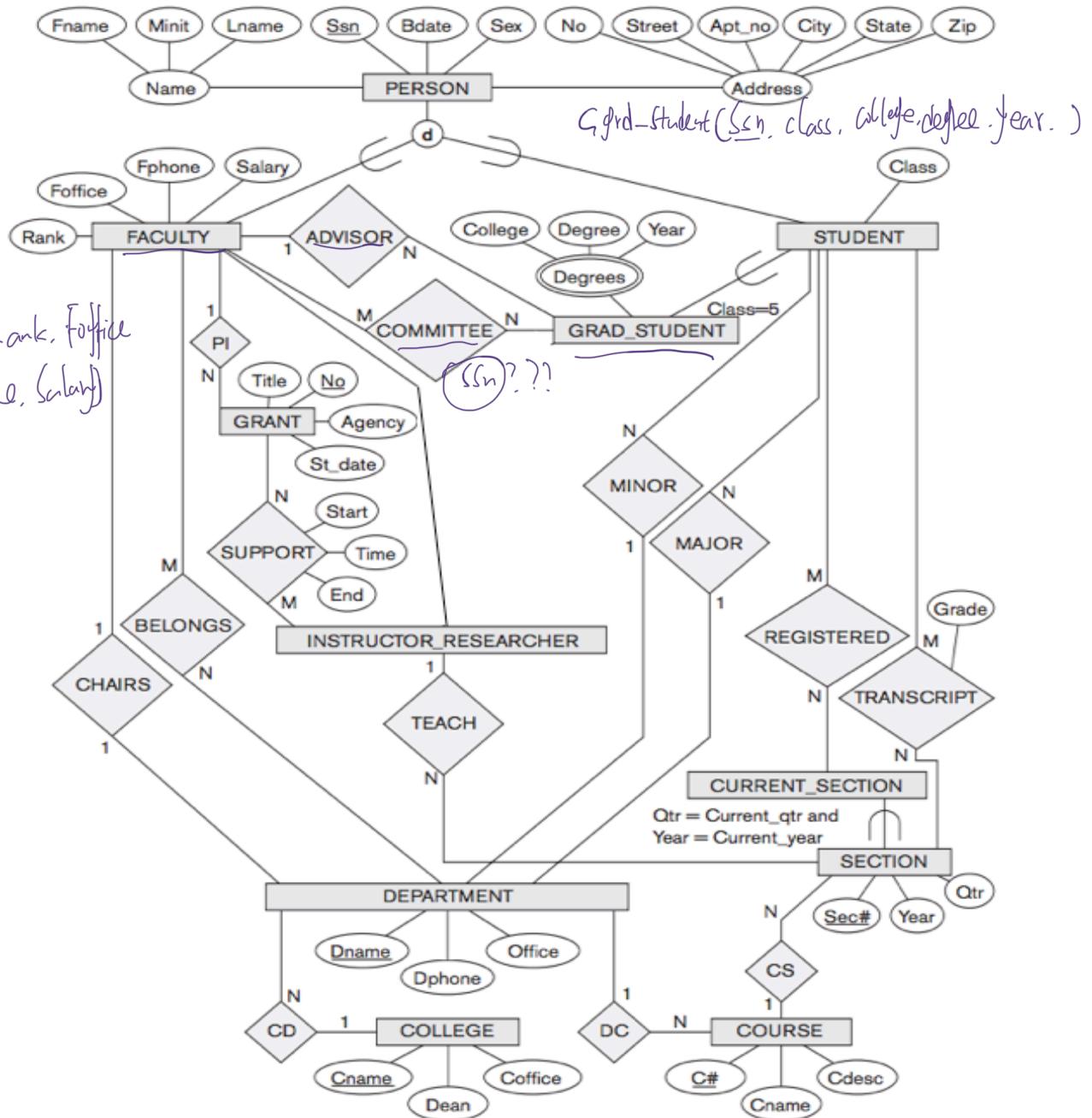
**Materials To Be Supplied To Students:**

None

**Instructions To Students:**

**Additional exam materials (eg. answer booklets, rough paper) will be provided upon request.**

**Question 1.** (10 marks) Answer the following questions based on the ER diagram below.



- a) (1 mark) How are faculty members uniquely identified? Explain your answer

*Ssn*

*because faculty is subclass of person.*

- b) (1 mark) Can a graduate student be advised by multiple faculty members? Explain your answer

No

- c) (1 mark) Can a faculty member belong to no departments? Explain your answer

Yes

- d) (1 mark) Can a faculty member belong to more than one department? Explain your answer.

Yes.

- e) (6 marks) Map the following entities and relationships from the ER diagram into a relational schema: GRAD\_STUDENT, FACULTY, ADVISOR, COMMITTEE. Make sure that you specify the primary keys and foreign keys of the tables.

**Question 2.** (8 marks) Answer the following questions.

Consider the relation instance on the right-hand side. Observe that  $B \rightarrow A$  appears to hold with respect to the given instance.

- a. (2 marks) Determine whether the following dependencies hold with respect to the given instance. Justify your answer.

$$A \rightarrow B \quad \text{No. } John \rightarrow 1/2 \\ Jill \rightarrow 4/5$$

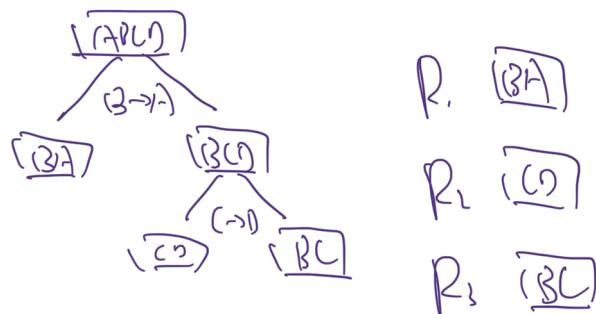
$$B \rightarrow C \quad \text{Yes.}$$

A	B	C	D
John	1	10	1
John	2	11	1
Jane	3	11	1
Jane	3	11	1
Jill	4	12	0
Jill	5	13	0

- b. (2 marks) Assuming that  $B \rightarrow A$  and  $C \rightarrow D$  hold in the given schema, find a candidate key for the relation

$$\{B, C\}$$

- c. (4 marks) Assuming that  $B \rightarrow A$  and  $C \rightarrow D$  hold in the given schema, decompose it, if necessary, so that all the resultant relations are in BCNF.



**Question 3.** (10 marks) Consider the following relational schema, in which the Catalog relation lists the prices charged for parts by suppliers.

- Supplier (sid, sname, address)
- Parts (pid, pname, colour)
- Catalog (sid, pid, cost)
  - Catalog.sid references Supplier.sid
  - Catalog.pid references Parts.pid

Write the following queries in SQL

a) (5 marks) For every supplier that supplies at least one green parts, print the name of the supplier and the total number of parts that she supplies.

b) (5 marks) Find the sname of suppliers that supply every red part

**Question 4.** (8 marks) Answer the following questions.

- a) (4 marks) For each of the following pairs of runtimes, determine which one is asymptotically faster. Justify your response.

I.  $T(A) = 3n + 2n^3$  vs.  $T(B) = 2n + 25n^2$

$$n^3 \quad n^2$$

$$\overline{T(R)}$$

II.  $T(A) = n + 100n^{0.1}$  vs.  $T(B) = 2n + 10\log n^2$

$$n \quad = \quad n$$

III.  $T(A) = 16^{2\lg n}$  vs.  $T(B) = 11n^7 + 5n^2$

$$\overline{T(r)}$$

- b) (4 marks) Determine the running time of functionA in terms of big-O notation. Justify your answer.

```
def functionA(n):
    for i in range(n):
        functionB(i)

def functionB(n):
    i = 1
    while (i < n):
        i = i * 2
    print i
```

**Question 5.** (8 marks) Answer the following questions.

a) (4 marks) Consider the following function.

```
def secretFunction(aList):
    for i in range(1, len(aList)):
        tmp = aList[i]
        k = i
        while k > 0 and tmp < aList[k - 1]:
            aList[k] = aList[k - 1]
            k -= 1
        aList[k] = tmp
```

I. Which sorting algorithm is being implemented by 'secretFunction'?

merge sort

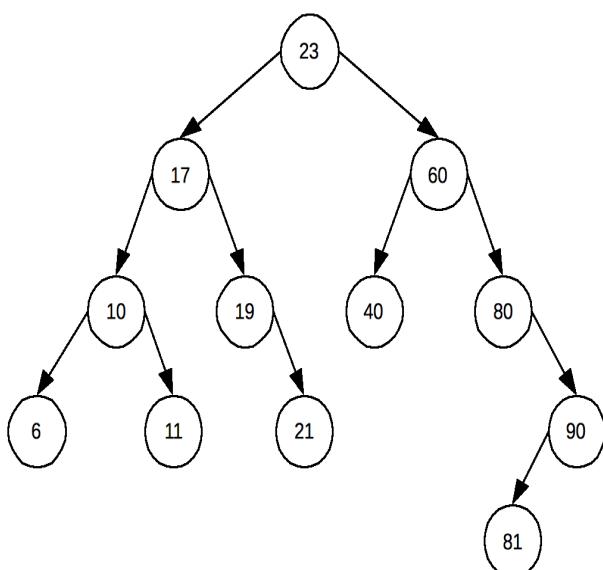
II. Briefly explain the task of the while loop and the for loop inside the function.

b) (4 marks) What is the running time of mergesort? Justify your answer by analysing the work performed by the mergesort algorithm.

n log n

**Question 6.** (8 marks) Answer the following questions.

This question makes reference to the following Binary Search Tree:



a) (1 mark) What is the height of this tree?

4

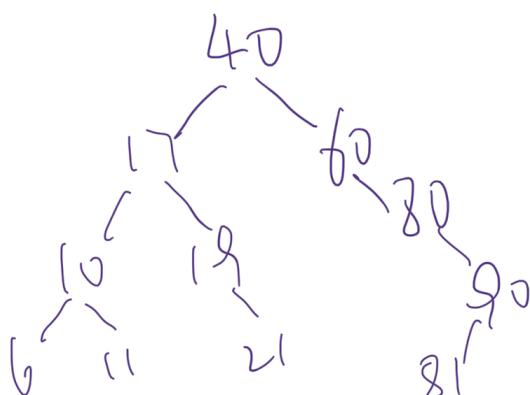
b) (1 mark) Which node is the root node?

23

c) (1 mark) What is the depth of node 19?

3

d. (3 marks) Draw the tree resulting from deletion of node with value 23, assuming that we are using the node's successor to replace it.



e. (2 marks) Suppose that we have numbers between 1 and 1000 in a binary search tree, and we want to search for the number 363. Tick the options which could be the sequence of nodes examined?

- I. 23, 60, 17, 80, 40, 19, 10, 90, 21, 11, 6, 81
- II. 23, 81, 90, 80, 40, 60, 21, 19, 11, 6, 10, 17
- III. 23, 17, 10, 6, 11, 19, 21, 40, 60, 80, 90, 81
- IV. 6, 11, 10, 17, 21, 19, 23, 40, 60, 81, 90, 80
- V. 6, 10, 11, 17, 19, 21, 23, 40, 60, 80, 81, 90

✓

**Question 7.** (8 marks) Answer the following questions.

- a) (2 marks) A hash table of length 10 uses open addressing with hash function  $h(k)=k \bmod 10$ , and linear probing. After inserting 6 values into an empty hash table, the table is as shown. Which one of the following choices gives a possible order in which the key values could have been inserted in the table? Briefly justify your answer.

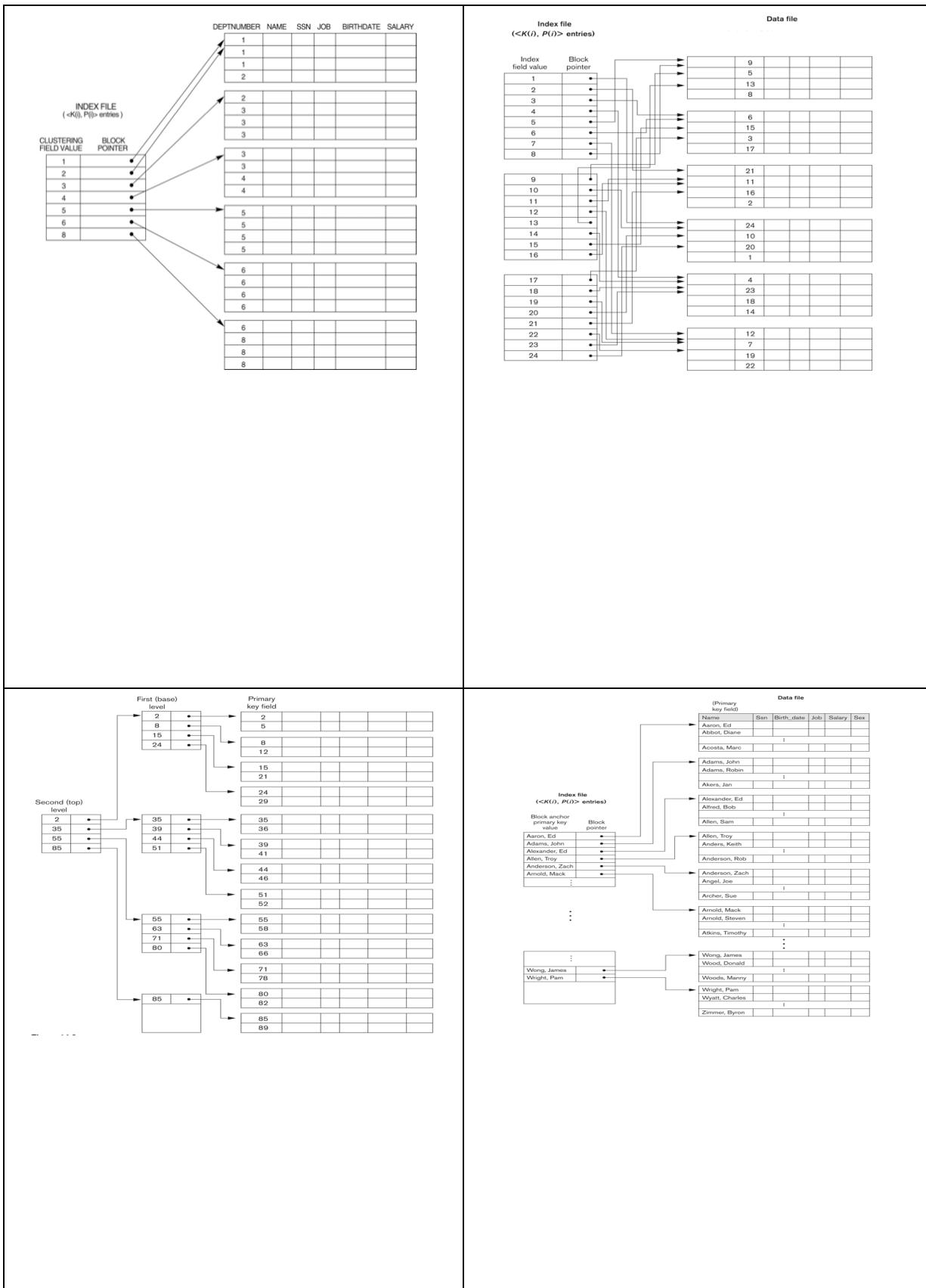
- (A) 46, 42, 34, 52, 23, 33  
(B) 34, 42, 23, 52, 33, 46  
(C) 46, 34, 42, 23, 52, 33  
(D) 42, 46, 33, 23, 34, 52  
(E) None of the above

1  
2

0	
1	
2	42
3	23
4	34
5	52
6	46
7	33
8	
9	

- b) (2 marks) Is double hashing vulnerable to secondary clustering? Briefly justify your answer.
- c) (4 marks) For good hashing performance, we need to have a good hash function. Can we create an ideal hash function that would work well in all problems independent of the dataset? If so explain how, and if not explain why?

**Question 8.** (8 marks). For each of the following images, (1) identify the type of indexing being demonstrated and (2) provide a short description of how that index operates.



**Question 9.** (12 marks) Consider the following relational schema, in which the Catalog relation lists the prices charged for parts by suppliers.

- Supplier (sid, sname, address)
- Parts (pid, pname, colour)
- Catalog (sid, pid, cost)
  - Catalog.sid references Supplier.sid
  - Catalog.pid references Parts.pid

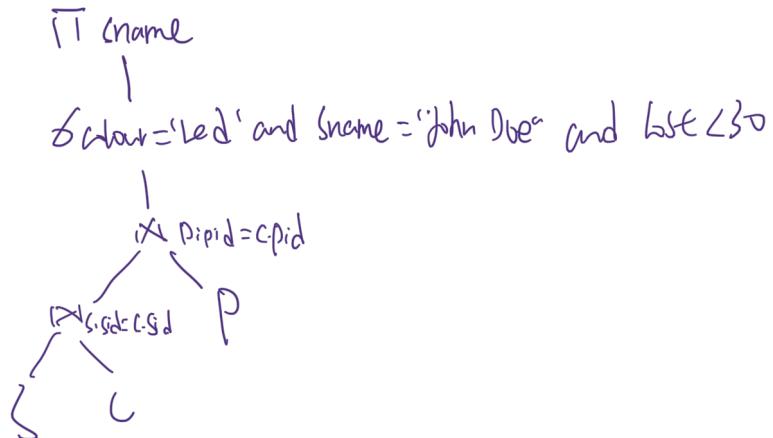
a) (6 marks) Write the following queries in relational algebra

I. Find the name of parts supplied by “Harry Potter” that are green and cost more \$100.

II. Find the name of suppliers who supply every red part

- b) (2 marks) Draw the initial query tree for the SQL query below and explain briefly why this plan is considered inefficient.

```
SELECT sname  
FROM Parts p, Supplier s, Catalog c  
Where p.pid = c.pid AND s.sid = c.sid AND  
colour = 'red' AND sname = "John Doe" and cost < 300.
```



- c) (4 marks) Transform the initial query into an equivalent final query tree that is efficient to execute.

**END OF EXAMINATION**