

**MILITARY COLLEGE OF SIGNALS**  
**FINAL EXAM**  
**BESE 15-A**  
**CPS 480 Database Systems**

**Instructor: A/P Dr. Imran Siddiqi**

**Time: 2.5 Hours**  
**Max Marks: 50**

**Note:** This question paper comprises **3** pages.

**(3+6)**

1. a. Consider the following relation.

| X  | Y  | Z  |
|----|----|----|
| x1 | y1 | z1 |
| x1 | y1 | z2 |
| x2 | y1 | z1 |
| x2 | y1 | z3 |

List all the functional dependencies that this instance satisfies.

b. Consider a relation R with four attributes ABCD. For each of the following sets of functional dependencies, identify the candidate key(s) for R and state the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF). (Note: All three parts below are independent of each other and you need to attempt them separately)

- i.  $C \rightarrow D, C \rightarrow A, B \rightarrow C$
- ii.  $B \rightarrow C, D \rightarrow A$
- iii.  $A \rightarrow B, BC \rightarrow D, A \rightarrow C$

**(4+4+2+6)**

2. a. Consider the *Student* relation shown in the following.

| SID | Name     | Age |
|-----|----------|-----|
| 101 | Amelia   | 25  |
| 102 | Arnaud   | 30  |
| 103 | Claudie  | 22  |
| 104 | Florence | 25  |
| 105 | Guray    | 45  |
| 106 | Nicolas  | 30  |
| 107 | Vincent  | 30  |

- i. Show the sparse index on the field 'Name' where the index file contains entries 'Ameila' and 'Guray'
- ii. Show the secondary index on the field 'Age'.

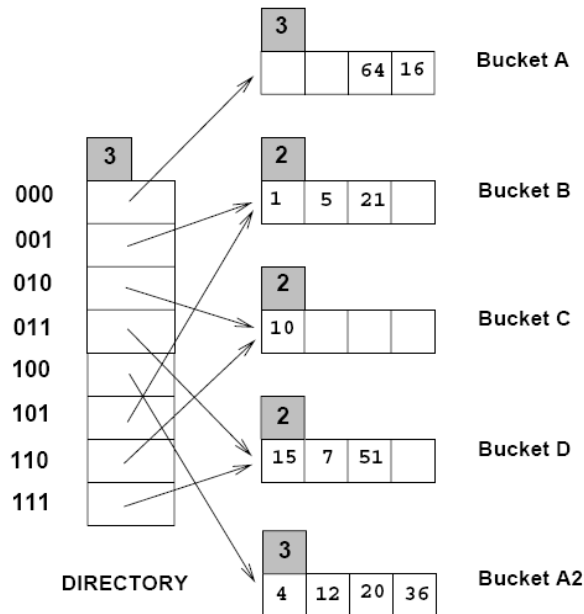
b. Organize the *Student* relation shown above using hash indexing on the search key *SID*. The following parameters are given:

$$\text{Hash Function} = (\text{Sum of Digits in SID}) \bmod 3$$

Number of buckets = 3  
Capacity of each bucket = 2

c. How the hash indexing you performed in part b (above) is different from hash file organization.

d. Consider the extendible hashing index shown in the following.



- Show the index after inserting an entry with hash value 68.
- Show the index after inserting entries with hash values 17 and 69 to the original index.

(3+3)

3. a. Using precedence graph, find if the following two schedules are conflict serializable or not.

|     |         |         |         |          |
|-----|---------|---------|---------|----------|
| T1: | READ(A) | READ(B) |         | WRITE(B) |
| T2: |         | READ(A) | READ(B) | WRITE(B) |

|     |         |          |          |                  |
|-----|---------|----------|----------|------------------|
| T1: | READ(A) | WRITE(A) | READ(B)  | WRITE(B)         |
| T2: |         | READ(A)  | WRITE(A) | READ(B) WRITE(B) |

b. Given two transactions  $T1$  and  $T2$  and two data objects  $A$  and  $B$ , give an example schedule which results in:

- A write-write conflict
- A read-write conflict

(4+6)

4. a. Assume you are given a document database of six documents containing the text as indicated in the following.

| Document | Text                                   |
|----------|--|
| 0        | pease porridge hot pease porridge cold |
| 1        | pease porridge in the pot              |
| 2        | nine days old                          |
| 3        | some like it hot some like it cold     |

|   |                         |
|---|-------------------------|
| 4 | some like it in the pot |
| 5 | nine days old           |

Show the results of creating an inverted file index on the given documents.

**b.** Consider the four documents below:

| Document | Words         |
|----------|---------------|
| 0        | Apple, Cell   |
| 1        | Doll, Goat    |
| 2        | Elephant, Ink |
| 3        | House, Fruit  |

Let the hashing function be the 5 bit binary representation of the first character of each word. For simplicity, assume A=1, B=2, C=3 and so on.

- i. Show the result of creating a signature file for the above documents.
- ii. What documents are retrieved if a user queries the term '*Elephant*'.

**(4+3+2)**

**5. a.** Assume that MCS has a total of 500 students. It is known that:

- i. 300 students play cricket
- ii. 375 students live in hostels
- iii. 200 students both play cricket and live in hostels

|               | Cricket | NO Cricket | Sum |
|---------------|---------|------------|-----|
| Hostel Living |         |            |     |
| Out Living    |         |            |     |
| Sum           |         |            | 500 |

Complete the above table and find the support and confidence of the following association rules.

- i. Cricket => Hostel Living
- ii. Cricket => Out Living

**b.** Given the following transactions:

- t1: Beef, Chicken, Milk
- t2: Beef, Cheese
- t3: Cheese, Boots
- t4: Beef, Chicken, Cheese
- t5: Beef, Chicken, Clothes, Cheese, Milk
- t6: Chicken, Clothes, Milk
- t7: Chicken, Milk, Clothes

State three association rules which satisfy  $minsup = 30\%$  and  $minconf = 80\%$ .

**c.** What is the difference between data mining and data warehousing?

+++++ Bon Courage +++++