

**BITS, PILANI - DUBAI CAMPUS  
KNOWLEDGE VILLAGE, DUBAI**

**Database Systems : CSUC 352      Comprehensive Exam**  
Duration : 3 Hours      Max marks : 40      Weightage : 40%      Date : 24-05-05

---

**Note : Attempt all questions**

**PART A**

**All questions in this section carry 3 marks**

1. Explain the architecture of a DBMS with a neat figure and mention the functionality of each component.
2. Draw an ER diagram for the following problem  
We watch movies in cinema halls or on television. Stars act in movies, shooting of movies is usually done in studios. When we consider a movie we would like to know about the title of the movie, year of release, length (in hours) and film type. As regards to a star, we are interested in his or her name and address. For a studio we want to know its name and address.
3. List four important properties of ACID transactions
4. Explain the significance of Bitmap indexes and how they are efficient in storing 1 data in multidimensional databases.
5. Consider the relational schema  $R = \{ A, B, C, D, E, F, G, H, I, J \}$  Consider the following FD's  $G = \{ \{ A, B \} \rightarrow \{ C \}, \{ B, D \} \rightarrow \{ E, F \}, \{ A, D \} \rightarrow \{ G, H \}, \{ A \} \rightarrow \{ I \}, \{ H \} \rightarrow \{ J \} \}$

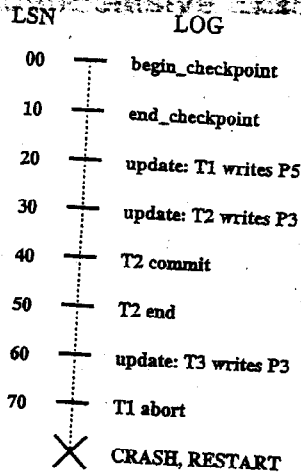
What is the key for R. Decompose R into 2NF and 3NF relations

**PART B**

**All questions in this section carry 4 marks**

6. What is the need for index structures ? Clearly mention any three types of indexes with neat figures.

7. Briefly answer the following questions  
 Consider the given fig and answer the given questions



- What is done in analysis ? (Be precise about the points at which analysis begins and ends )
- What is done during redo ? (Be precise about the points at which analysis begins and ends )
- What is done during Undo ? (Be precise about the points at which analysis begins and ends )

8. Explain the following with respect to distributed databases
- Replication of data
  - Horizontal and Vertical fragmentation
  - Distributed Deadlock

### PART C

9. Write short notes on the following
- Participation constraints
  - Candidate key
  - Types of swizzling
  - Conflict equivalent schedule
  - Deadlock in databases
  - Contents of the update Log record ( 2 X 6 = 12 marks )

Fill in the blanks

- a) Oracle is an example of a \_\_\_\_\_ data base - 1 mark

\*\*\*\*\*

BITS, PILANI - DUBAI CAMPUS  
KNOWLEDGE VILLAGE, DUBAI

**Database Systems CSUC 352 Test 2(Open book)**  
Duration : 50 Mins Max marks : 20 Weightage : 20% Date : 20-03-05

---

1. Answer the following with respect to extendible hashing
  - a. Explain why local and global depth is needed ?
  - b. After an insertion that causes the directory size to double, how many buckets have exactly one directory entry pointing to them? If an entry is then deleted from one of these buckets , what happens to the directory size ? Explain your answers briefly
  - c. Does extendible hashing guarantee at most one access to retrieve a record with a given key value ?
  - d. If the hash function distributes data entries over the size of the space of bucket numbers in a very skewed (non – uniform) way, what can you say about the size of the directory ? What can you say about the space utilization in data pages
  - e. Does doubling the directory require us to examine all buckets with the local depth equal to the global depth ?

8 marks

2. Which of the three basic file organizations would you choose for a file where the most frequent operations are as follows
  - a. Search for a record based on the range of field values
  - b. Perform inserts and scans where the order of the records does not matter.
  - c. Search for a record based on a particular field value.

1 X3 marks

3. a. Explain on how data is inserted into a grid file should be ideally partitioned.  
b. Explain the difference between 3NF and BCNF

2X 2 = 4 marks

P.T.O

4. Given the following universal relational schema R where

$R = \{A, B, C, D, E, F, G, H, I, J\}$

The functional dependencies specified are  $\{\{A, B\} \rightarrow \{C\}, \{A\} \rightarrow \{D, E\}, \{B\} \rightarrow \{F\}, \{F\} \rightarrow \{G, H\}, \{D\} \rightarrow \{I, J\}\}$  What is the key of R. Decompose R into 2 NF, 3 NF relations.

5 marks

BITS, PILANI - DUBAI CAMPUS  
KNOWLEDGE VILLAGE, DUBAI

Database Systems CSUC 352 Test 2(Open book)

Duration : 50 Mins Max marks : 20 Weightage : 20%

ANSWERING SCHEME

---

1. Extendible hashing allows the size of the directory to increase and decrease depending on the number and variety of inserts and deletes. Once the directory size changes, the hash function applied to the search key value should also change. So there should be some information in the index as to which hash function is to be applied. This information is provided by the *global depth*.

An increase in the directory size doesn't cause the creation of new buckets for each new directory entry. All the new directory entries except one share buckets with the old directory entries. Whenever a bucket which is being shared by two or more directory entries is to be split the directory size need not be doubled. This means for each bucket we need to know whether it is being shared by two or more directory entries. This information is provided by the *local depth* of the bucket. The same information can be obtained by a scan of the directory, but this is costlier.

2. Exactly two directory entries have only one directory entry pointing to them after a doubling of the directory size. This is because when the directory is doubled, one of the buckets must have split causing a directory entry to point to each of these two new buckets.

If an entry is then deleted from one of these buckets, a merge may occur, but this depends on the deletion algorithm. If we try to merge two buckets only when a bucket becomes empty, then it is not necessary that the directory size decrease after the deletion that was considered in the question. However, if we try to merge two buckets whenever it is possible to do so then the directory size decreases after the deletion.

3. No "minimum disk access" guarantee is provided by extendible hashing. If the directory is not already in memory it needs to be fetched from the disk which may require more than one disk access depending upon the size of the directory. Then the required bucket has to be brought into the memory. Also, if alternatives 2 and 3 are followed for storing the data entries in the index then another disk access is possibly required for fetching the actual data record.
4. Consider the index in Fig. 11.1. Let us consider a list of data entries with search key values of the form  $2^i$  where  $i > k$ . By an appropriate choice of  $k$ , we can get all these elements mapped into the *Bucket A*. This creates  $2^k$  elements in the directory which point to just  $k + 3$  different buckets. Also, we note there are  $k$  buckets (data pages), but just one bucket is used. So the utilization of data pages =  $1/k$ .
5. No. Since we are using extendible hashing, only the local depth of the bucket being split needs be examined.

BITS, PILANI - DUBAI CAMPUS  
KNOWLEDGE VILLAGE, DUBAI

Database Systems CSUC 352 Test. 1(Closed Book)

Duration : 50 Mins Max marks : 20 Weightage : 20% Date : 20-03-05

1. SELECT \* FROM stock\_data;

| symbol | purch_date | Purch_price | shares | curr_date  | curr_price |
|--------|------------|-------------|--------|------------|------------|
| CAT    | 2000-02-24 | 37.8125     | 13     | 2003-08-14 | 70.49      |
| CAT    | 2000-09-21 | 34.8125     | 14     | 2003-08-14 | 70.49      |
| EK     | 2001-09-10 | 43.23       | 11     | 2003-08-14 | 26.96      |
| GM     | 2002-03-20 | 60.33       | 8      | 2003-08-14 | 37.28      |
| IP     | 2000-09-21 | 28.8125     | 17     | 2003-08-14 | 40.69      |
| JPM    | 2002-03-20 | 35.19       | 14     | 2003-08-14 | 33.85      |
| SBC    | 2000-02-24 | 35.75       | 13     | 2003-08-14 | 23.02      |

(7 rows)

Give SQL queries on the above given table

- List the amount of shares in descending order (large-to-small) in the order of purchase. 3 marks
- Find the cost of the shares by using the given formula. The given formula must display the total value of the shares of a particular symbol.

Shares \* Purchase Price = Purchase Total 3 marks

- Retrieve the symbol and purch\_date of shares where the shares is either 11,13,14. 2 marks

2. Consider the following set of requirements for a university database that is used to keep track of students transcripts.

BITS, PILANI - DUBAI CAMPUS  
KNOWLEDGE VILLAGE, DUBAI

**Database Systems CSUC 352 Test 1(Closed Book)**

Duration : 50 Mins Max marks : 20 Weightage : 20% Date : 20-03-05

**MARKING SCHEME**

- a). List the amount of shares in descending order (large-to-small) in the order of purchase.

```
SELECT * FROM stock_data ORDER BY purch_date, shares DESC;
```

b)

```
SELECT symbol, SUM(shares * purch_price) AS purch_total  
FROM stock_data GROUP BY symbol ;
```

c)

```
SELECT symbol, purch_date  
FROM stock_data  
WHERE shares IN ( 11,13,14) 2 MARKS
```

2. ER diagram – 6 marks

3. Defination of Trigger – 1 MARK  
SYNTAX OF TRIGGER - 3 MARK

4. View is used to give a user a restricted view of the database. He will be able to view only the portion of a database applicable to him. A view does not have a physical form it is more of a virtual table. 1 MARK

Syntax for a view:

```
CREATE VIEW VIEW_NAME  
AS ( SELECT QUERY ) 1 MARK
```