

Department of Computer Science, UET Lahore

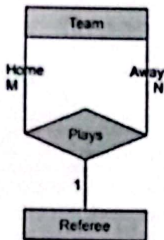
Program: BSc Computer Science	Course: CS-362	Database Systems
Examination: Final term	Session: 2020	
Maximum Marks: 40+HW(7)	Semester: 4 th	
Time Allowed: 100 minutes/1 hr 40 mins	Date: 09-05-2022	
Registration Number: Solution	O.Marks	

Instructions:

- This is a closed book exam. All working/Explanation must be clearly shown to receive full credit.
- All questions carry one point unless mentioned otherwise.

No	Database Systems Theory Exam			Marks
1	CLO1	ER Diagram/Real World Problems	PLO3: Design/Understand	8

1. Model the following relationships in ER. Two teams play football against each other. A referee makes sure the rules are followed. (2 marks)



2. Write XML Schema and DTD for the following XML snippet.(4 marks)

```

<?xml version="1.0"?>
<University Name="UET">
  <Department>
    <Name>Computer Science</Name>
    <Sections>4</Sections>
    <Accreditation>NCEAC</Accreditation>
  </Department>
  <Department>
    <Accreditation>PEC</Accreditation>
    <Sections>1</Sections>
    <Name>Computer Engineering</Name>
  </Department>
</University>
    
```

DTD:
DTD is not possible as it does not allow random order of elements

XML Schema:

```

<?xml version="1.0" ?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="University">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="Department" type="DepartmentType"
          minOccurs="0" maxOccurs="unbounded" />
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
    
```

```

</xsd:sequence>
<xsd:attribute name="Name" type="xsd:string" use="required" />
</xsd:complexType>
</xsd:element>

<xsd:complexType name="DepartmentType">
<xsd:all>
<xsd:element name="Name" type="xsd:string" />
<xsd:element name="Accreditation" type="xsd:string" />
<xsd:element name="Sections" type="xsd:integer" />
</xsd:all>
</xsd:complexType>
</xsd:schema>

```

3. Which of the following most certainly implies the need for an entire table to implement?
- A binary relationship
 - A ternary relationship**
 - A recursive relationship
 - An identifying relationship

Explanation: _____

4. The attribute AGE is calculated from DATE_OF_BIRTH. The attribute AGE is
- Single valued
 - Multi valued
 - Composite
 - Derived**

Explanation: _____

2	CLO3	Normalization/Anomalies	PLO3: Design/Apply	2
<p>5. Let relation R(A,B,C,D,E,F,G,H) satisfy the following functional dependencies: $A \rightarrow B$, $CH \rightarrow A$, $B \rightarrow E$, $BD \rightarrow C$, $EG \rightarrow H$, $DE \rightarrow F$. Which of the following FDs is also guaranteed to be satisfied by R? (2 marks)</p> <p>(A) $BDG \rightarrow AE$ (B) $CDE \rightarrow AF$ (C) $ADE \rightarrow CH$ (D) $BCD \rightarrow F$</p>				

Explanation: _____

3	CLO2	SQL	PLO3: Design/Apply	14
<p>6. Given a table named Sailor(sid, sname, age, rating), complete the given trigger such that more than 60 value in the age column is rejected. (2 marks)</p> <pre> CREATE TRIGGER t1 BEFORE UPDATE ON SAILORS REFERENCING OLD ROW AS OldTuple NEW ROW As NewRow FOR EACH ROW BEGIN If new.age>60 then Set new.age = old.age; else Set new.age = new.age; end if; end; </pre> <p>In the above example we are creating trigger before update. so, if the new age is greater than 60 we should not update else</p>				
<p>12. If we want to retain all duplicates, we must write _____ in place of union.</p> <p>(A) Union all (B) Union some (C) Intersect all (D) Intersect some</p> <p>Explanation: _____</p> <p>13. DROP is a _____ statement in SQL.</p> <p>(A) Query (B) Embedded SQL (C) DDL (D) DML</p> <p>Explanation: _____</p> <p>14. What is true regarding materialized view?</p> <p>(A) Definition and data of materialized view is stored in database (B) Only definition of materialized view is stored in database</p>				

we should update We can call this trigger by using 'S' symbol.

7. Which of the following views are updatable?
- (A) Views that have only one occurrence of R in from clause
 - (B) Views that have relation R in subquery
 - (C) Views that have only one relation R in from clause
 - (D) View that are simple join of two relation R and S

Explanation: _____

8. What is true regarding constraints?
- (A) Assertions are table level constraints
 - (B) Attribute level constraints are guaranteed to hold in case it refers to other relation
 - (C) Tuple Level constraints can be violated in referential constraints
 - (D) To make assertion on a relation, it is written in schema definition of table.

Explanation: _____

9. Solution of circular foreign key constraint named Foo is to write statement
- (A) SET CONSTRAINT Foo DEFERRABLE;
 - (B) UPDATE CONSTRAINT Foo DEFERRABLE;
 - (C) SET CONSTRAINT Foo DEFERRABLE INITIALLY DEFERRED;
 - (D) SET CONSTRAINT Foo DEFERRED;

Explanation: _____

10. Primary, unique or candidate key to maintain the referential integrity, default policy of database systems is
- (A) SET NULL
 - (B) CASCADE
 - (C) It depends of user configuration
 - (D) Reject changes

Explanation: _____

11. Updating the value of the view
- (A) Will affect the relation from which it is defined
 - (B) Will not change the view definition
 - (C) Will not affect the relation from which it is defined
 - (D) Cannot determine

Explanation: _____

- (C) Materialized view re-compute each column when they are called
- (D) Synchronization is not an issue in materialized view

Explanation: _____

15. Which of the following operators can convert a bag into set?
- (A) JOIN
 - (B) GROUP BY
 - (C) DISTINCT
 - (D) SELECT

Explanation: _____

16. For foreign key constraint, referenced table should have
- (A) Any attribute
 - (B) Primary or Unique key
 - (C) Primary key

Explanation: _____

17. Result of two relations $M(X, Y, Z)$ and $N(X, A, Z)$ for Join based on condition $A=Y$ consists of number of columns?
- (A) 5
 - (B) 4
 - (C) 3
 - (D) 6

Explanation: _____

18. In set operators, result of intersection of two relations $(R \cap S)$ is equivalent to
- (A) $R - (R \cup S)$
 - (B) $R - (S - R)$
 - (C) $R - (R - S)$
 - (D) $S - (R \cup S)$

Explanation: _____

4	CLO4	Query Optimization	PLO3: Design/Apply	9
19. Can we have two clustering indices on the same relation for different search keys? Explain.				

In general, it is not possible to have two primary indices on the same relation for different keys because the tuples in a relation would have to be stored in different order to have same values stored together. We could accomplish this by storing the relation twice and duplicating all values, but for a centralized system, this is not efficient.

20. Suppose that blocks can hold either ten records or 99 keys and 100 pointers. Also assume that the average B-tree node is 70% full, i.e. it will have 69 keys and 70 pointers. We can use B-trees as part of several different structures. For each structure described below, (i) the total number of blocks needed for a 1,000,000 record file, and (ii) the average number of disk I/O's to retrieve a record given its search key. You may assume nothing is in memory initially, and the search key is the primary key for the records. Also, note that the data file is a sequential file, sorted on the search key, with 10 records per block and the Btree is a dense index. (4 marks)

First, there are 100,000 data blocks.
If there are an average of 70 pointers per block in the bottom-level nodes of the B-tree, then there are 1,000,000/70 = 14286 B-tree blocks at that level.
The next level of the B-tree requires 1/70th of that, or 204 blocks.
and the third level has 1/70th of that, or 3 blocks.
The fourth level has only the root block.
The number of blocks needed is therefore 100,000 + 14,286 + 204 + 3 + 1 = 114,494 blocks.
Since the B-tree has four levels, we must make a total of five disk reads to go through the Btree to the desired data block.

21. Some hash functions do not work as well as theoretically possible. Suppose that we use the hash function on integer keys i defined by $h(i) = i^2 \bmod B$, where B is the number of buckets (2 marks).

- What is wrong with this hash function if $B=10$?
- How good is this hash function if $B=16$?

Few buckets are being used. More collisions.

22. What is the maximum number of items that a B+ tree of order 50 and depth 2 can have at its leaves?

The maximum number of child links of each internal node is $2 \cdot 50 + 1$ or 101. The maximum number of items at each leaf is 100. \therefore 101 child links (max) at the root (depth 0) \therefore 1012 (= 10,201) child links (max) across all depth 1 nodes \therefore 1012 \cdot 100 (= 1,020,100) items (max) across all depth 2 leaves

The maximum number of child links of each internal node is $2 \cdot 50 + 1$ or 101. The maximum number of items at each leaf is 100.

- \therefore 101 child links (max) at the root (depth 0)
- \therefore 101² (= 10,201) child links (max) across all depth 1 nodes
- \therefore 101² \cdot 100 (= 1,020,100) items (max) across all depth 2 leaves

23. Valid syntax to declare an index is

- Update Bar CREATE Keyindex ON Bar(A,B);
- CREATE Keyindex ON Bar(A,B);
- CREATE INDEX Keyindex ON Bar(A,B);
- CREATE INDEX Keyindex Bar(A,B);

Explanation: _____

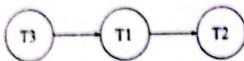
5	CLO4	Transactions	PLO3: Design/Apply	7
24. Lowest ISOLATION level in transactions is (A) Read Committed (B) Repeatable Read (C) Read Uncommitted (D) Serializable Explanation: _____ _____			27. What is true regarding isolation levels of transaction (A) Repeatable read does not allow phantom read (B) Read committed allows phantom read (C) Repeatable read does not allow repeatable reads (D) Repeatable read is the highest isolation level	

Which of these statements about serializable schedules is true?

- (A) Every serializable schedule is recoverable.
- (B) Every serializable schedule contains no conflicting actions
- (C) Every 2PL schedule is serializable.
- (D) None of the above.

Explanation: _____

26. Draw the precedence graph for the schedule $w_1(A)$; $w_2(A)$; $r_3(B)$; $w_3(B)$; $r_1(B)$; (2 marks)
Solution:



Explanation: _____

28. Describe the action of the recovery manager, including changes to both disk and the log, if there is a crash and the last log record to appear on disk is COMMIT U for the following sequence of undo/redo-log records written by two transactions T and U (2 marks).

Solution: Since U is committed, we redo its actions, setting B to 21 and D to 41. Then, since T is uncommitted, we undo its actions from the end moving backwards; we set C to 30 and A to 10.

<START T>
<T, A, 10, 11>
<START U>
<U, B, 20, 21>
<T, C, 30, 31>
<U, D, 40, 41>
<COMMIT U>
<T, E, 50, 51>
<COMMIT T>

6	CLO4	Query Optimization	PL03: Design/Apply	7
---	------	--------------------	--------------------	---

Imagine you are designing a table to store recent transactions for an online shopping platform and there are 1 trillion transactions. You want to record the following information: • user id • user name • item id • item name • transaction id • amount of money (\$) for the transaction (e.g. \$4.11, \$670.50, etc.) Assume there are 1 billion users, and 1 billion items for sale on the platform. The longest string for user and item names contains 64 characters. You should consider proper data types listed below: byte, short, int, long, float, double, Boolean, char.

- What is the size of each row in bytes? 148 bytes
- What data type should you use for each column?

UserId: int- 4 bytes
Username: char(64). 64 bytes
itemId: int.. 4 bytes/
ItemName: char(64). 64 bytes
Transaction id: long. 8 bytes
Amount: float. 4 bytes

- What is the size of the table in TB? $148 * 10^{12} = 134\text{TB}$ approx
- Assume the size of the table is 200TB, how long in seconds will it take to read the whole table from RAM? RAM transfer 100GBs/second $\Rightarrow 200\text{TB} = 200000\text{GBs} \Rightarrow 2000\text{s} \Rightarrow 33.33$ minutes approx

Assume for 1 trillion TXNs:

- How long in days (round to nearest integer) will it take to read the whole table from disk if each row of the table is stored randomly in the disk? Single disk block seek time = 10ms, Disk transfer 100MBs/second. So, total time in days = $10 * 10^{-3} * 10^{12} = 10^{10}\text{s} = 115740$ days approx
- How long in days (round to nearest integer) will it take to read the whole table from disk if the table is stored in DB blocks? (1 point) $148 * 1 \text{ trillion} / 64 \text{ MBs} =$
- What is the cost in dollars for saving the table in disk? Assume disk space costs \$100/TB. 134 * 100\$