**CS3492 - DATABASE MANAGEMENT SYSTEMS**

**PART-A**

**UNIT1**

**1.Who is a DBA? What are the responsibilities of a DBA? April/May-2011**

A database administrator (short form DBA) is a person responsible for the design,

implementation, maintenance and repair of an organization's database. They are also known by

the titles Database Coordinator or Database Programmer, and is closely related to the Database

Analyst, Database Modeller, Programmer Analyst, and Systems Manager.

The role includes the development and design of database strategies, monitoring and improving

database performance and capacity, and planning for future expansion requirements. They may

also plan, co-ordinate and implement security measures to safeguard the database.

**2. What is a data model? List the types of data model used. April/May-2011**

A database model is the theoretical foundation of a database and fundamentally determines in

which manner data can be stored, organized, and manipulated in a database system. It thereby

defines the infrastructure offered by a particular database system. The most popular example of a

database model is the relational model.

**Types of data model used**

 Hierarchical model

 Network model

 Relational model

 Entity-relationship

 Object-relational model

 Object model

**3. Give the levels of data abstraction?**

a) Physical level

b) Logical level

c) View level

**4.What is storage manager?**

A storage manager is a program module that provides the interface between the low level data

stored in a database and the application programs and queries submitted to the system.

**5.What are the components of storage manager?**

The storage manager components include

a) Authorization and integrity manager

b) Transaction manager

c) File manager

d) Buffer manager

**6.Define single valued and multivalued attributes.**

Single valued attributes: attributes with a single value for a particular entity are called single

valued attributes.

Multivalued attributes : Attributes with a set of value for a particular entity are called

multivalued attributes.

**7.What are stored and derived attributes?**

Stored attributes: The attributes stored in a data base are called stored attributes.

Derived attributes: The attributes that are derived from the stored attributes are called derived

attributes.

**8.What does the cardinality ratio specify?**

Mapping cardinalities or cardinality ratios express the number of entities to which another entity

can be associated. Mapping cardinalities must be one of the

following:

• One to one

• One to many

• Many to one

• Many to many

9.**What is meant by normalization of data?**

It is a process of analyzing the given relation schemas based on their Functional Dependencies

(FDs) and primary key to achieve the properties

 Minimizing redundancy

 Minimizing insertion, deletion and updating anomalies

**10. What is an entity relationship model?**

The entity relationship model is a collection of basic objects called entities and relationship

among those objects. An entity is a thing or object in the real world that is distinguishable from

other objects.

**UNIT:2**

**1. Define the terms i) DDL ii) DML**

**DDL:** Data base schema is specified by a set of definitions expressed by a special language

called a data definition language.

**DML:**

A data manipulation language is a language that enables users to access or manipulate data as

organized by the appropriate data model.

**2.What is a primary key?**

Primary key is chosen by the database designer as the principal means of identifying an entity in

the entity set.

**3.What is a super key?**

A super key is a set of one or more attributes that collectively allows us to identify uniquely an

entity in the entity set.

4. **What is embedded SQL? What are its advantages? April/May-2011**

Embedded SQL is a method of combining the computing power of a programming language and

the database manipulation capabilities of SQL. Embedded SQL statements are SQL statements

written in line with the program source code of the host language. The embedded SQL

statements are parsed by an embedded SQL preprocessor and replaced by host-language calls to

a code library. The output from the preprocessor is then compiled by the host compiler. This allows programmers to embed SQL statements in programs written in any number of languages such as: C/C++, COBOL and Fortran.

**5.Write short notes on relational model**

The relational model uses a collection of tables to represent both data and the relationships among those data. The relational model is an example of a record based model.

**6. Define tuple and attribute**

• Attributes: column headers

• Tuple : Row

**7.Define the term relation.**

Relation is a subset of a Cartesian product of list domains.

8. **Define tuple variable**

Tuple variable is a variable whose domain is the set of all tuples.

**9. Define the term Domain.**

For each attribute there is a set of permitted values called the domain of that

attribute.

**10. What is a candidate key?**

Minimal super keys are called candidate keys.

**UNIT:3**

**1. What are the ACID properties? APRIL/MAY-2011**

(atomicity, consistency, isolation, durability) is a set of properties that guarantee databasetransactions are processed reliably. In the context of databases, a single logical operation on thedata is called a transaction. For example, a transfer of funds from one bank account to another,even though that might involve multiple changes (such as debiting one account and creditinganother), is a single transaction.

**2. What are two pitfalls (problem) of lock-based protocols? APRIL/MAY-2011**

 Deadlock

 Starvation

**3. What is transaction?**

Collections of operations that form a single logical unit of work are called transactions.

**4. What are the two statements regarding transaction?**

The two statements regarding transaction of the form:

 Begin transaction

 End transaction

**5. What are the properties of transaction?**

The properties o f transactions are:

 Atomicity

 Consistency

 Isolation

 Durability

**6. What are the states of transaction?**

The states of transaction are

 Active

 Partially committed

 Failed

 Aborted

 Committed

 Terminated

**7. Define deadlock?**

Neither of the transaction can ever proceed with its normal execution. This situation is called

deadlock.

**8. Define the phases of two phase locking protocol**

Growing phase: a transaction may obtain locks but not release any lock.

Shrinking phase: a transaction may release locks but may not obtain any new locks.

**9. Define upgrade and downgrade?**

It provides a mechanism for conversion from shared lock to exclusive lock is known as upgrade.

It provides a mechanism for conversion from exclusive lock to shared lock is known as

downgrade.

**10. Define shadow paging**.

An alternative to log-based crash recovery technique is shadow paging. This technique needs

fewer disk accesses than do the log-based methods.

**11. Define page.**

The database is partitioned into some number of fixed-length blocks, which are referred to as

pages.

**UNIT:4**

**1. What are the advantages and disadvantages of indexed sequential file? APRIL/MAY-**

**2011**

The advantage of ordering records in a sequential file according to a key is that you can then search the file more quickly. If you know the key value that you want, you can use one of the relatively fast searches. The disadvantage is that when you insert, you need to rewrite at least everything after the insertion point, which makes inserts very expensive unless they are done at the end of the file. An indexed file approach keeps a (hopefully) small part of each row, and some kind of "pointer" to the row's location within the data file. This allows a search to use the index, which is ordered by the index and (again hopefully) much smaller and therefore much faster than scanning the entire data file for the indexed data.

**2. What is database tuning? APRIL/MAY-2011**

Database tuning describes a group of activities used to optimize and homogenize the

performance of a database. It usually overlaps with query tuning, but refers to design of the

database files, selection of the database management system (DBMS), operating system and

CPU the DBMS runs on.

**3. Give the measures of quality of a disk.**

Capacity

Access time

Seek time

Data transfer rate

Reliability

Rotational latency time.

**4. What are the types of storage devices?**

 Primary storage

 Secondary storage

 Tertiary storage

**5. Define average seek time.**

The average seek time is the average of the seek times, measured over a sequence of random

requests.

**6. Define rotational latency time.**

The time spent waiting for the sector to be accessed to appear under the head is called the

rotational latency time.

**7. What is meant by Multi-dimensional database?**

A multi-dimensional database is a computer software system designed to allow for efficient and

convenient storage and retrieval of large volumes of data that is (1) intimately related and (2)

stored, viewed and analyzed from different perspectives and these perspectives are called

dimensions.

**8. What is meant by Spatial database?**

A spatial database is a database that is optimized to store query data that represents objects

defined in geometric space. Most spatial databases allow representing simple geometric objects

such as points, lines and polygons. Some spatial databases handle more complex structures such

as 3D objects, topological coverages,etc.

**9. What is meant by Mobile database?**

A mobile database is a database that resides on mobile device such as a PDA, a smart phone, or a

laptop. Such devices are often limited in resources such as memory, computing power and

battery power.

**10. Explain how reliability can be improved through redundancy?**

The simplest approach to introducing redundancy is to duplicate every disk. This technique is called mirroring or shadowing. A logical disk then consists of two physical disks, and write is carried out on both the disk. If one of the disks fails the data can be read from the other. Data will be lost if the second disk fails before the first fail ed disk is repaired.

**UNIT:5**

**1. List the threats to databases.**

**** Loss of integrity

 Loss of availability

 Loss of confidentiality

**2. List out the control measures.**

**** Access control

 Inference control

 Flow control

 Data encryption

**3. What is meant by Data warehouse?**

A data warehouse is a repository (archive) of information gathered from multiple sources, stored

under a unified schema at a single site.

 Greatly simplifies querying, permits study of historical trends

 Shifts decision support query load away from transaction processing

systems

**4. Define Data mining.**

Data mining - knowledge discovery in database. Data mining is the process of semi-

automatically analyzing large databases to find useful patterns.

**5. List out the functionalities of Data warehouse.**

 Data cleaning

 Data transformation

 Data integration

 Data loading &

 Periodic data refreshing

**6. List the types of security mechanisms.**

 Discretionary security mechanisms

 Mandatory security mechanisms

**7. What are the database design issues?**

 Legal and ethical issues

 Policy issues

 System related issues

**8. What are the actions performed by DBA?**

 Account creation

 Privilege granting

 Privilege revocation

 Security level assignment

**9. What is meant by frequent pattern tree algorithm?**

The Frequent pattern tree algorithm reduces the total number of candidate itemsets by producing

a compressed version of the database in terms of an FP-tree. The FP-tree stores relevant

information and allows for the efficient description of frequent itemsets. The algorithm consists

of 2 steps:

1. Build FP-tree

2. Use the tree to find frequent itemsets.

**10. What is meant by Classification?**

Classification is the process of learning a model that is able to describe different classes of data.

**119. List the applications of data mining.**

 Marketing

 Finance

 Resource optimization

 Image Analysis

 Fraud detection