



EXERCISE

Q1 Select the best answer for the following MC Qs.

- I. Duplication of data in different files is called _____
 A. Data inconsistency B. Data redundancy C. Data overflow D. Invalid data
- II. If data is not updated in a file in file based data management systems, what type of problem will it cause?
 A. Data inconsistency B. Data redundancy
 C. Data overflow D. Invalid data
- III. A record is also called:
 A. Attribute B. Entity C. Property D. Tuple
- IV. An attribute is also called:
 A. Record B. Entity C. Field D. Relation
- V. What is a thing of interest to an organization called about which data is to be held?
 A. Field B. Relation C. Entity D. Attribute
- VI. In which type of database, data is held in tables and tables are linked by common field?
 A. Hierarchical Database B. Network Database
 C. Relational Database D. Object-Oriented Database
- VII. In which database model, data is organized in tree-like structure?
 A. Hierarchical Database B. Network Database
 C. Relational Database D. Object-Oriented Database
- VIII. What refers to the maximum number of times an instance in one entity can be associated with instances in the related entity?
 A. Relation B. Cardinality C. Modality D. E-R diagram
- IX. What refers to the minimum number of times an instance in one entity can be associated with an instance in the related entity?
 A. Relation B. Cardinality C. Modality D. E-R diagram
- X. What is a key field called that is used in relationship between tables whose value matches a primary key in the other table?
 A. Candidate key B. Secondary key C. Alternate key D. Foreign key

Answers

I.	B	II.	A	III.	D	IV.	C	V.	C
VI.	C	VII.	A	VIII.	B	IX.	C	X.	D

Q2 Give short answers of the following questions.

- I. Differentiate between data and information.

Answer

Difference between data and information:

Data	Information
Data refers to facts and figures. It is raw material that is input in the computer for processing.	Information is organized form of data that is meaningful and easily understandable by computer users. It is the output produced by computer after processing the data.

- II. Define a database.

Answer

Database

A database is a collection of related data. A database consists of one or more data files and data

files contain fields.

iii. What is a Database Management System (DBMS)?

Answer

Database Management System (DBMS)

A database management system (DBMS) is a set of programs that allow users to create, maintain and manipulate database, and store or retrieve data from those database files. It provides user-friendly access and controls between user and database. Its main purpose is to improve data sharing, data access, decision making and increase end user productivity.

iv. What is a relational database?

Answer

Relational Database

A relational database is a collective set of multiple data sets organized by tables, records and columns. Relational database establish a well-defined relationship between database tables. Tables communicate and share information, which facilitates data searchability, organization and reporting.

v. Define primary and secondary key.

Answer

Primary Key

Each entity in a database must have a unique key field known as primary key to identify a record.

Secondary Key

Sometimes a record in a table need to be searched on field other than primary key such a field is known as secondary key.

vi. Define attribute and entity.

Answer

Attribute

Attribute is a property or characteristics of an entity.

Entity

Entity is a thing of interest to an organization about which data is to be held.

vii. What is meant by Instance?

Answer

Instance

An instance in a database is the actual content of the database at a particular point in time.

viii. Differentiate between cardinality and modality.

Answer

Difference between cardinality and modality

Cardinality	Modality
Cardinality refers to the maximum number of times an instance in one entity can be associated with instances in the related entity.	Modality refers to the minimum number of times an instance in one entity can be associated with an instance in the related entity.

ix. Define schema.

Answer

Schema

The word schema is a Greek word, which means shape or plan.

x. Why is it necessary to normalize a relational database?

Answer

To avoid data redundancy and inconsistency, it is necessary to normalize a relational database.

Q3. Write long answers of the following questions.

i. What are the advantages of using a DBMS over file management system?

Answer

DBMS has many advantages over the file management system because it provides solutions to all the problems faced in file management system.

Problems in file management system

The following problems were faced in file management system:

- Reduced Data Redundancy
- Data consistency
- Program-Data Dependency
- Lack of Flexibility
- Data not Sharable
- Backup and Recovery
- Data Security

1. Reduced Data Redundancy

Data redundancy is the duplication of data in many different files in file management system.

Example

A salesman's data may be held on a file in the Sales Department and also Personnel Department.

When data is to be updated, it must be changed in both files. This results in wastage of storage space and may lead to data inconsistency.

DBMS not faced with the problem

This problem is not faced in DBMS because all the data belonging to the entire organization is centralized in a common pool of data, accessible by all the programs.

2. Data consistency

In file management system, same data may be held in several different files, it has to be updated in each separate file when it changes. If data is not updated in any file, it causes data inconsistency.

Solution in DBMS

This problem is solved in DBMS because all the data is centralized for use by all the programs.

Example

In a school, students' information such as name, address, phone number and class are held in a file in Admission Office. The same data with tuition fee and some other data are also kept in another file in Accounts Office. If a student's address changes it must be updated on both files. In case, it is updated in only one file, the other file will have out-of-date data. This will cause data inconsistency.

3. Program-Data Dependency

In file management system, every computer program in each Department has to specify exactly what data fields constitute a record. Any change to the format to the data field of a record, such as adding a new field or changing the length of a field means that every program, which uses that file has to be changed.

Example

If a new field, place of birth of students is to be added in student data files in Admission Office and Accounts Office, run both the programs need to be changed.

4. Lack of Flexibility

In file management system, when information of a non-routine nature is needed, it can take weeks to assemble the data from the various files and write new programs produce the required reports.

Example

A report about students is required in a school that has to merge information stored in various files. This will require lot of time and effort by the programmer to write a program to print such a report in file management system, whereas, it can be feeling very easily produced in database management system.

5. Data not Sharable

When a copy of a data file is made for sharing data with other people in a file management system, it will soon lead to data inconsistency if data in one of the files is not updated. Therefore, it is difficult to share data with others in a file management system.

6. Backup and Recovery

DBMS provides facilities for backup and recovery from failures including disk crash, power failure, software errors, which may bring the database from the consistent state to a state prior to the failure.

7. Data Security

In Conventional filing systems, there is no centralized security system, which restricts users according to their role in the organization. DBMS makes it easier to enforce security restrictions since database is centralized. Users are provided permissions to access data according to their rights.

ii. Define Database Administrator and describe the tasks performed by him.

Answer

Database Administrator

Database Administrator (DBA) is the person responsible for supervising the database and the use of DBMS in an organization

Tasks performed by DBA

DBA performed the following tasks:

1. Designing the database and enforcing the operational policies and procedures for its usage.
2. Planning security measures and backup of database.
3. Controlling privileges and permissions of database users.
4. Allocating passwords to users
5. Planning recovery procedures if hardware or software failure occurs and ensuring that no data is lost.
6. Providing training to new employees about using the database.

iii. Define database model and explain its types.

Answer

Database Model

A database model is the theoretical foundation of a database and determines in which manner data can be stored, organized and manipulated. It defines a way of structuring data.

Types of Database Models

There are five types of database models, which are as follows:

- > Hierarchical Database Model
- > Network Database Model
- > Relational Database Model
- > Object-Oriented Database Model
- > Object Relational Database Model

1. Hierarchical Database Model

In a hierarchical model, data is organized into a tree-like structure as shown in Fig.7.1.

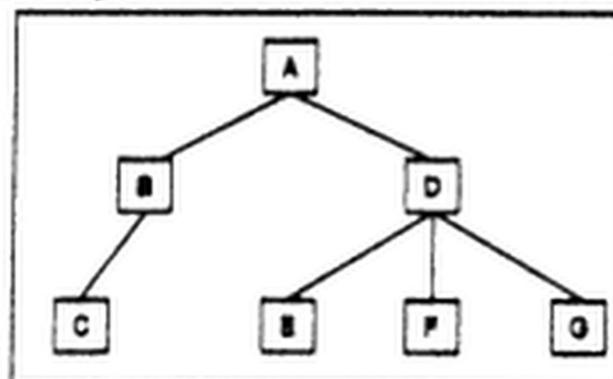


Fig. 7.1 Hierarchical Database Model

It is a logical construct with owner and subordinate relationship.

Members and Owners

Data elements in a subordinate relationship are called members and those having subordinates are called owners. In a hierarchical structure, subordinates can have only a single owner and there is only one data element that has no owner. This structure is very suitable in describing many relationships in the real world.

Examples of hierarchical model

Some examples of hierarchical model are table of contents, departments of an organization and types of memories.

2. Network Database Model

A network model is a logical structure in which some of the data elements can have more than one owner data element as shown in Fig 7 2

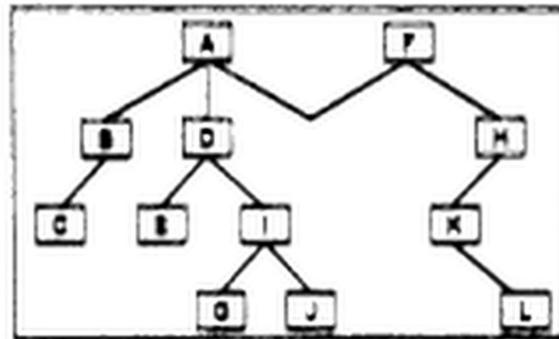


Fig. 7 2 Network Database Model

A network is a more complex structure than a hierarchical model. A hierarchical model can be thought of as a network with some discipline imposed on it. An application with which crew members, aircraft, routes and schedules are related is by its nature a network model.

3. Relational Database Model

In a relational database model, data is held in tables as shown in Fig.7.3 and the tables are linked by means of common fields.

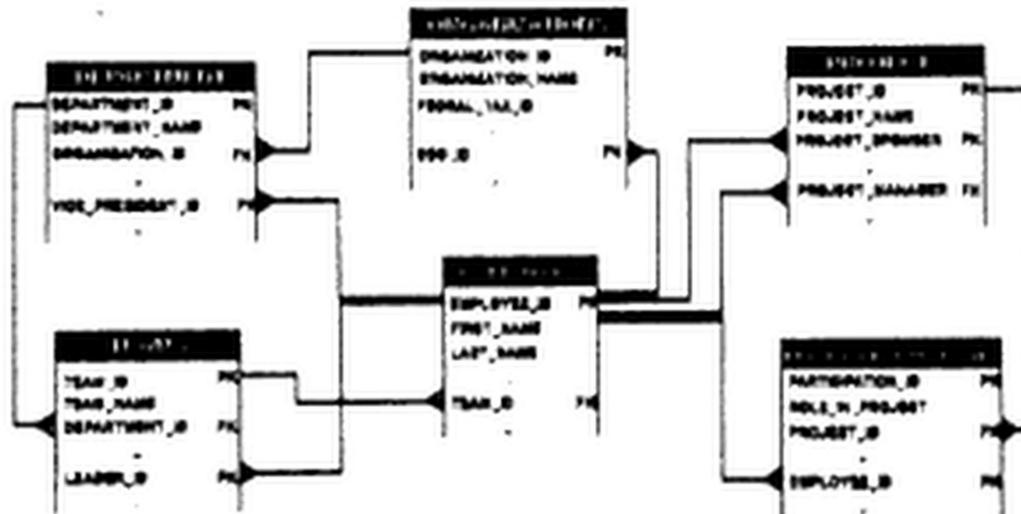


Fig.7.3 Table in a relational database

Relation

The "relation" in a relational database refers to the various tables in the database, which are linked with each other.

Table

A table in database consists of rows and columns. One row of table holds one record. Each column in the table holds one field or attribute

Use of relational database model

The relational database model is used to make database management systems more independent of any particular application. It is the most popular database model used in business applications.

4. Object-Oriented Database Model

Object-oriented databases also called Object Database Management Systems (ODBMS). Object databases store objects rather than data such as integers, real numbers and strings. Objects are used

in object oriented programming languages such as C++ and Java

Characteristics of Real world

Real world objects have two characteristics, that is, state and behaviour.

Example

The object radio has four states and five behaviours as shown in Fig.7.4

Object	Radio
States	Behaviours
On	turn on
Off	turn off
Current channel	tune channel
Volume	increase/decrease volume

Fig.7.4 States and Behaviour of object Radio

Object

An object stores its states in variables and exposes its behaviour through functions in programming languages.

5. Object Relational Database Model

An Object Relational Database (ORD) is a database management system similar to relational database, but with an object-oriented database. In object relational database, the approach is essentially that of relational database. The data resides in the database and is manipulated collectively with queries in a query language.

iv. Describe the steps involved in database design.

Answer

Steps for designing database

Following steps are performed for designing database:

Step 1: Define the Purpose of the Database (Requirement Analysis)

Gather the requirements and define the objective of your database, e.g. drafting out the sample input forms, queries and reports, often helps.

Step 2: Gather Data, Organize in tables and Specify the Primary Keys

Once you have decided on the purpose of the database, gather the data that are needed to be stored in the database. Divide the data into subject-based tables.

Choose one column (or a few columns) as the primary key, which uniquely identify the each of the rows.

Step 3: Create Relationships among Tables

A database consisting of independent and unrelated tables serves little purpose. The power of relational database lies in the relationship that can be defined between tables. The most crucial aspect in designing a relational database is to identify the relationships among tables. The types of relationship include:

- One-to-many
- Many-to-many
- One-to-one

Step 4: Refine & Normalize the Design

Refine and normalize the design by following steps:

- Adding more columns.
- Create a new table for optional data using one-to-one relationship.
- Split a large table into two smaller tables.

v. What is an E-R diagram? Explain with examples.

Answer

E-R Diagram

The Entity-Relationship (ER) diagram or model defines the conceptual view of a database. It is made up of entities and the associations among them. At view level, the ER model is considered a

good option for designing databases. Any object, for example, entities, attributes of an entity, relationship sets, and attributes of relationship sets, can be represented with the help of an ER diagram. ER Model is best used for the conceptual design of a database.

ER Model is based on:

- Entities and their attributes
- Relationships among entities

These concepts are explained in Fig. 7.8.

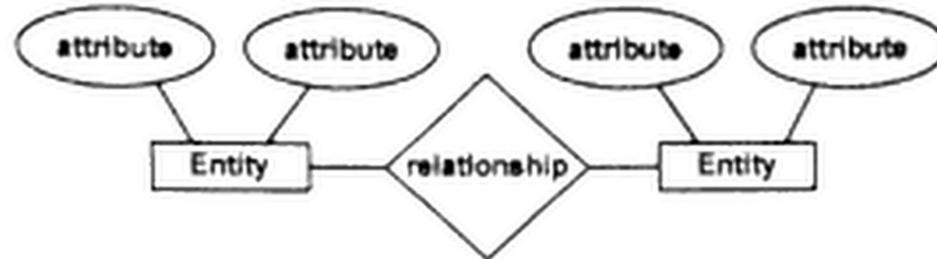


Fig.7.8 ER diagram concept

1. Entity

An entity in an ER Model is a real-world entity having properties called **attributes**. Every **attribute** is defined by its set of values called **domain**.

Example

In a school database, a 'Student' is considered as an entity. Student has various attributes like roll number, name, age, class, etc

Representation of entities

Entities are represented by means of rectangles. Rectangles are named with the entity set they represent. Fig. 7.9 shows some entities.



Fig.7.9 Entities Attributes

2. Attributes

Attributes are the properties of entities. Attributes are represented by means of ellipses. Every ellipse represents one attribute and is directly connected to its entity (rectangle)

Example

Fig 7.10 shows three attributes 'Roll No.', 'Name' and 'BirthDate' of entity 'Student'. The primary key attribute is underlined, for example, Roll_No. attribute.

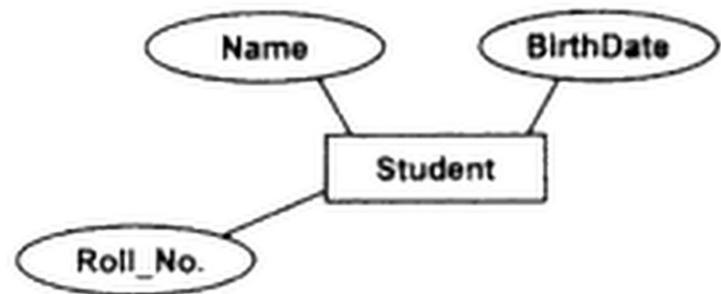


Fig.7.10 Attributes

3. Relationship

The logical association among entities is called relationship. Relationships are represented by a diamond symbol connected to the related entities. Fig 7.11 shows some relationships between entities.

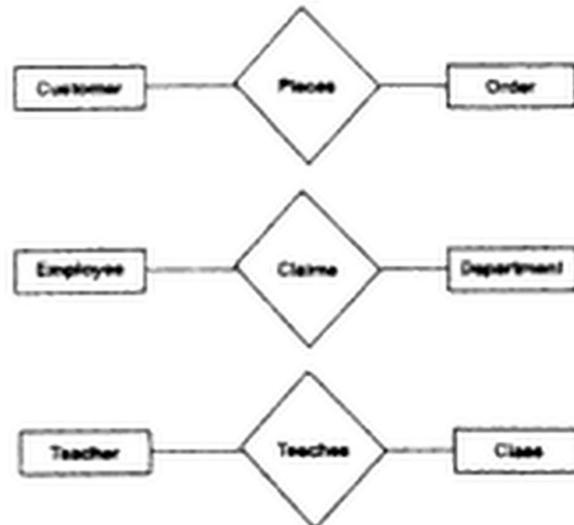


Fig.7.11 Relationships Degree of a Relationship

vi. What is Normalization? Explain the following Normal Forms 1NF, 2NF, 3NF.

Answer

Normalization

Normalization is a series of steps to create normal forms used to ensure that data anomalies are not created when information is inserted, updated or deleted.

First Normal Form

To get to first normal form, we break up our data into its related data groups.

Explanation

To understand how to obtain the first normal form, consider the table shown in Fig. 7.15.

It stores information about members of library and the books borrowed by them

MemberID	Name	TelNo	BookID	Title	Author	Pub_Date
481	Raja Imran	4815568	20056	Introduction to Java	Ali Ahmed	2008
481	Raja Imran	4815568	36211	Using Internet	Rizwan Khan	2009
532	Nasir Ali	2309118	20568	Computer Networks	Amjad Ali	1998
532	Nasir Ali	2309118	44195	Learning MS Word	Imran Khan	2010
532	Nasir Ali	2309118	31128	Web Designing	Waqas Arif	2011
669	Arif Khan	4212957	62617	Mastering MS Excel	Syed Fawad	2005

Fig. 7.15 Table in Un-normal Form

In this example, the primary key is the MemberID and we have two fields, Name and Tel_No repeating themselves across the tuples for each of the different BookID entries. The redundant data will take lot of space to store the data and more time will be required to search through.

The rules of first normal form break this table into two and relate them to each. The repeated information in the table is removed. This will create the two tables shown in Fig. 7.16.

MemberID	Name	Tel_No
481	Raja Imran	4815568
532	Nasir Ali	2309118
669	Arif Khan	4212957

MemberID	BookID	Title	Author	Pub_Date
481	20056	Introduction to Java	Ali Ahmed	2008
481	36211	Using Internet	Rizwan Khan	2009
532	20568	Computer Networks	Amjad Ali	1998
532	44195	Learning MS Word	Imran Khan	2010
532	31128	Web Designing	Waqas Arif	2011
669	62617	Mastering MS Excel	Syed Fawad	2005

Fig. 7.16 Tables in First Normal Form

The other normal forms are also obtained in a similar way by further breaking up the table and removing repeating information if it exists.

Second Normal Form (Removes Partial Functional Dependencies)

To move into second normal form, all partial functional dependencies must be removed. Partial functional dependencies occur when there is a composite key and only one part of the key is needed to determine one or more other attributes. To remove partial functional dependencies, we have to project out the partial functional dependencies into its own table, creating two tables from one.

Explanation

In our example, we have to further check if we are keeping more data than we need to in our relation. If we look at Fig. 7.16, we can see that for every UnitCode, we are also storing the UnitName. It will make more sense, if we look up UnitCode 443 in another table and find the UnitName. This is how we can avoid the duplicate information entered in the UnitName field. UnitName is only dependent on UnitCode and has no dependency on StudentID. To be in second normal form, all partial dependencies must be removed. When we apply the rules of second normal form and remove the partial functional dependency in our database, we get the following tables shown in Fig. 7.17, which are in second normal form.

StudentID	UnitCode
481	UC443
481	UC471
481	UC422
532	UC443
669	UC471
669	UC468

UnitCode	UnitName
UC443	Database Design
UC471	Operating System
UC422	Digital Electronics
UC468	Accounting
UC443	Database Design
UC471	Operating System

Fig. 7.17 Table in Second Normal Form

Third Normal Form (Removes Transitive Functional Dependencies)

The next step is to move into third normal form. In order to reach third normal form, all transitive functional dependencies must be removed. Similar to partial functional dependencies, transitive functional dependencies occur when a non-key attribute can determine another non-key attribute. Transitive functional dependencies are removed by projecting out the transitively dependent attributes to another table while leaving the determinate as a foreign key in the original table.

UnitCode	UnitName	CourseCode	CourseName
UC443	Database Design	COMP2010	Computing
UC471	Operating System	COMP2010	Computing
UC422	Digital Electronics	COMP2010	Computing
UC468	Accounting	BUS2010	E-Commerce

Fig. 7.18 Table with CourseCode and CourseName attributes

Explanation

In our example above, we have CourseName that is dependent on CourseCode and CourseCode is dependent on UnitCode. Therefore, CourseName is transitively dependent UnitCode. To remove the transitional dependency, break the table into two tables, leaving the determinant as foreign key in the original table. This is shown in the tables of Fig. 7.19. Here, CourseCode is foreign key in the first table and CourseCode is primary key in the second table.

UnitCode	UnitName	CourseCode
UC443	Database Design	COMP2010
UC471	Operating System	COMP2010
UC422	Digital Electronics	COMP2010
UC468	Accounting	BUS2010

CourseCode	CourseName
COMP2010	Computing
BUS2010	E-Commerce

Fig 7.19 Table in Third Normal Form

LAB ACTIVITIES

Following lab activities are to be carried out during the practical periods.

- 1) A city has many universities and each university has many departments. Each department has many teachers. Define the attributes for the entities, university, department and teacher and draw the E-R diagram.

Answer

Practical Work

- 2) A company sells many products to their customers. There are many suppliers who supply various products. Draw the E-R diagram of entities, company, supplier and customer.

Answer

Practical Work

- 3) An institute offers many courses. A student is allowed to register in only one course at a time. An instructor teaches many courses. Draw the E-R diagram of entities, courses, student and instructor.

Answer

Practical Work

- 4) A city has many Cable TV companies. Each company has many TV channels. A customer subscribes Cable TV from only one company. Draw the E-R diagram of entities, Cable TV Company, channels and customer.

Answer

Practical Work

