B.Com (Computers) II Year
RELATIONAL DATABASE MANAGEMENT SYSTEM
Unit- II

1. What is model?

A. A representation of reality that retains only selected details.

2. What is an object set?

A. Objects represent things that are important to users in the portion of reality we want to model. A set of things of the same kind are called as object sets. A particular instance of an object set is called object instance.

3. What is an Entity and Attribute?

A. An entity is a thing or object in the real world that is distinguishable from the the other objects. An entity set has a set of properties called attributes.

4. Explain the different Mapping cardinalities.

A. Mapping cardinalities express the number of entities to which another entity can be associated via a relationship set.

For a binary relationship set R between entity sets A and B, the mapping cardinalities must be one of the following.

(i). One to One: An entity in A is associated with at most one entity in B and an entity in B is associated with at most one entity in A.

(ii). One to many: An entity in A is associated with any number of entities in B. An entity in B, however can be associated with at most one entity in A.

(iii). Many to One: An entity in A is associated with at most one entity in B and an entity in B, however can be associated with any number of entities in A.

(iv).Many to many: An entity in A is associated with any number of entities in B and an entity in B is associated with any number of entities in A.

5. What is Generalization and Specialization?

A. Generalization: An object set that is a superset of another object set.
Specialization: An object set that is a subset of another object set.

6. What is Aggregation?

A. Aggregation: It is a relationship set viewed as an object set.

7. Explain Entity-Relationship diagram.

A. E-R Diagram: An entity-relationship diagram (ERD) is a data modeling technique that creates a graphical representation of the entities, and the relationships between entities, within an information system. Any ER diagram has an equivalent relational table, and any relational table has an equivalent ER diagram.

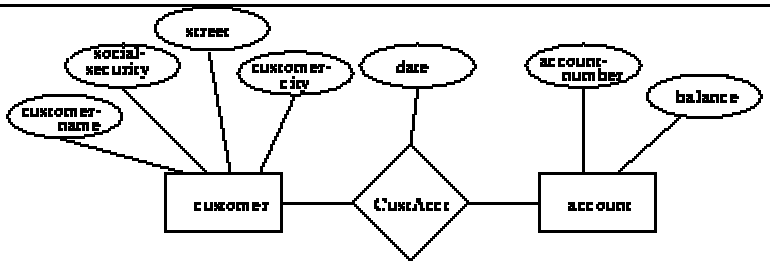
Entity: The entity is a person, object, place or event for which data is collected. It is equivalent to a database table. An entity can be defined by means of its properties, called attributes. For example, the CUSTOMER entity may have attributes for such things as name, address and telephone number.

Relationship: The relationship is the interaction between the entities.

E-R diagram components are:

- Rectangles representing entity sets.
- Ellipses representing attributes.

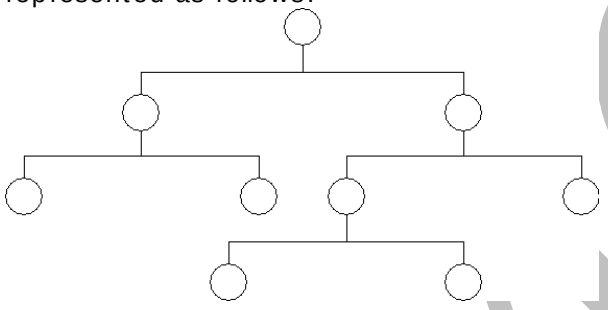
- Diamonds representing relationship sets.
- Lines' linking attributes to entity sets and entity sets to relationship sets.



8. Explain the different Implementation models.

A. There are three types of Implementation models

1. The Hierarchical Data Model: The Hierarchical Data Model can be represented as follows:

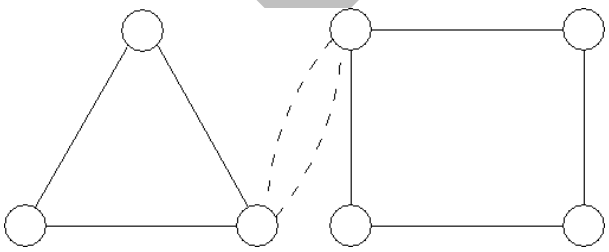


A hierarchical database consists of the following:

- It contains nodes connected by branches.
- The top node is called the root.
- If multiple nodes appear at the top level, the nodes are called root segments.
- Each node (with the exception of the root) has exactly one parent.
- One parent may have many children.

2. The Network Data Model:

The Network Data Model can be represented as follows:



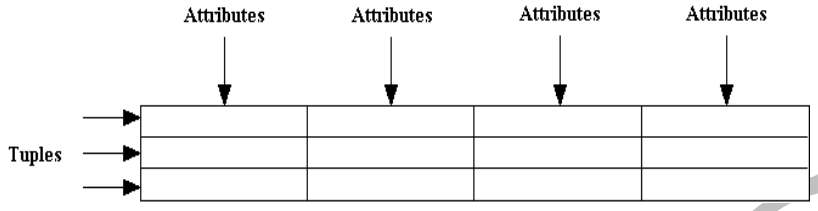
Like the Hierarchical Data Model the Network Data Model also consists of nodes and branches, but a child may have multiple parents within the network structure.

3. The Relational Data Model:

The Relational Data Model has the relation at its heart, but then a whole series of rules governing keys, relationships, joins, functional dependencies, transitive dependencies, multi-valued dependencies, and modification anomalies.

The Relation

The *Relation* is the basic element in a relational data model.



A relation is subject to the following rules:

- Relation (file, table) is a two-dimensional table.
- Attribute (i.e. field or data item) is a column in the table.
- Each column in the table has a unique name within that table.
- Each column is homogeneous. Thus the entries in any column are all of the same type (e.g. age, name, employee-number, etc).
- A Tuple (i.e. record) is a row in the table.

9. Explain the different types of keys in Relational Data Model.

A. Keys

- A simple key contains a single attribute.
- A composite key is a key that contains more than one attribute.
- A candidate key is an attribute (or set of attributes) that uniquely identifies a row.
- A primary key is the candidate key which is selected as the principal unique identifier. For example, empno is selected in EMP relation.
- A foreign key is an attribute (or set of attributes) that appears (usually) as a non key attribute in one relation and as a primary key attribute in another relation

10. What is Determinant and Dependent?

A. The terms determinant and dependent can be described as follows:

- The expression X⟶Y means 'if I know the value of X, then I can obtain the value of Y' (in a table or somewhere).
- In the expression X⟶Y, X is the determinant and Y is the dependent attribute.
- The value X determines the value of Y.
- The value Y depends on the value of X.

11. What is Functional Dependency?

A. A functional dependency can be described as follows:

- An attribute is functionally dependent if its value is determined by another attribute.
- That is, if we know the value of one data items, then we can find the value of another.
- Functional dependencies are expressed as X⟶Y, where X is the determinant and Y is the functionally dependent attribute.

12. What is Transitive Dependency?

A. A transitive dependency can be described as follows:

- An attribute is transitively dependent if its value is determined by another attribute *which is not a key*.
- If X⟶Y and X is not a key then this is a transitive dependency.
- A transitive dependency exists when A⟶B⟶C but NOT A⟶C.

13. What is Multi-Valued Dependency?

A. A multi-valued dependency can be described as follows:

- A table involves a multi-valued dependency if it may contain multiple values for an entity.
- X⟶⟶Y, ie X multi-determines Y, when for each value of X we can have more than one value of Y.

14. What is Join Dependency?

A. A join dependency can be described as follows:

If a table can be decomposed into three or more smaller tables, it must be capable of being joined again on common keys to form the original table.

15. What is Normalization?

Normalization is a process of evaluating and correcting table structures to minimize data redundancies, thereby reducing the likelihood of data anomalies. The normalization process involves assigning attributes to the tables based on the concept of Relational Data Model.

The objective of normalization is to ensure that each table conforms to the concept of well-formed relations. Normalized tables have the following characteristics.

- ❖ Each table represents a single subject. For example, a course table will contain only data that directly pertains to courses. Similarly, a student table will contain only student data.
- ❖ No data item will be unnecessarily stored in more than one table. The reason for this requirement is to ensure that the data are updated in only one place.
- ❖ All non-prime attribute in a table are dependent on the primary key. The reason for this requirement is to ensure that the data are uniquely identifiable by a primary key value.
- ❖ Each table is void of insertion, update or deletion anomalies. This is to ensure the integrity and consistency of the data.

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	$84.50	23.8
		101	John G. News	Database Designer	$105.00	19.4
		105	Alice K. Johnson *	Database Designer	$105.00	35.7
		106	William Smithfield	Programmer	$35.75	12.6
		102	David H. Senior	Systems Analyst	$96.75	23.8
18	Amber Wave	114	Annelise Jones	Applications Designer	$48.10	24.6
		118	James J. Frommer	General Support	$18.36	45.3
		104	Anne K. Ramoras *	Systems Analyst	$96.75	32.4
		112	Darlene M. Smithson	DSS Analyst	$45.95	44.0
22	Rolling Tide	105	Alice K. Johnson	Database Designer	$105.00	64.7
		104	Anne K. Ramoras	Systems Analyst	$96.75	48.4
		113	Delbert K. Joenbrood *	Applications Designer	$48.10	23.6
		111	Geoff B. Wabash	Clerical Support	$26.87	22.0
		106	William Smithfield	Programmer	$35.75	12.8
25	Starflight	107	Maria D. Alonzo	Programmer	$35.75	24.6
		115	Travis B. Bawangi	Systems Analyst	$96.75	45.8
		101	John G. News *	Database Designer	$105.00	56.3
		114	Annelise Jones	Applications Designer	$48.10	33.1
		108	Ralph B. Washington	Systems Analyst	$96.75	23.6
		118	James J. Frommer	General Support	$18.36	30.5
		112	Darlene M. Smithson	DSS Analyst	$45.95	41.4

The above table is not normalized. Consider the following deficiencies:

- PROJ_NUM is apparently intended to be a primary key or atleast a part of PK, but it contains nulls.
- The table entries invite data inconsistencies. For example, the JOB_CLASS value "Database Designer" or "DB Designer".
- The table displays data redundancies.
    1. update anomalies: For example modifying JOB_CLASS for employee number 101 requires many alterations.
    2. Insertion anomalies: Just to complete a row definition, a new employee must be assigned to a project. If the employee is not yet assigned, a phantom project must be created to complete the employee data entry.
    3. Deletion anomalies: Suppose that only one employee is associated with a given project. If that employee leaves the company and employee data are deleted, the project information will also be deleted. To prevent the loss of the project information, a fictitious employee must be created just to save the project information.

16. When is a table in 1NF?

A relational table doesn't contain repeating groups. The existence of repeating groups do exist, they must be eliminated by making sure that each row defines a single entity. In addition, the dependencies must be identified to diagnose the normal form.

1NF starts with a simple three-step procedure:

Step1: Eliminate the repeating groups

Start by presenting the data in a tabular format, where each cell has a single value and there are no repeating groups. To eliminate repeating groups, eliminate the nulls by making sure that each repeating group attribute contains an appropriate data value.

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	$84.50	23.8
15	Evergreen	101	John G. News	Database Designer	$105.00	19.4
15	Evergreen	105	Alice K. Johnson *	Database Designer	$105.00	35.7
15	Evergreen	106	William Smithfield	Programmer	$35.75	12.6
15	Evergreen	102	David H. Senior	Systems Analyst	$96.75	23.8
18	Amber Wave	114	Annelise Jones	Applications Designer	$48.10	24.6
18	Amber Wave	118	James J. Frommer	General Support	$18.36	45.3
18	Amber Wave	104	Anne K. Ramoras *	Systems Analyst	$96.75	32.4
18	Amber Wave	112	Darlene M. Smithson	DSS Analyst	$45.95	44.0
22	Rolling Tide	105	Alice K. Johnson	Database Designer	$105.00	64.7
22	Rolling Tide	104	Anne K. Ramoras	Systems Analyst	$96.75	48.4
22	Rolling Tide	113	Delbert K. Joenbrood *	Applications Designer	$48.10	23.6
22	Rolling Tide	111	Geoff B. Wabash	Clerical Support	$26.87	22.0
22	Rolling Tide	106	William Smithfield	Programmer	$35.75	12.8
25	Starflight	107	Maria D. Alonzo	Programmer	$35.75	24.6
25	Starflight	115	Travis B. Bawangi	Systems Analyst	$96.75	45.8
25	Starflight	101	John G. News *	Database Designer	$105.00	56.3
25	Starflight	114	Annelise Jones	Applications Designer	$48.10	33.1
25	Starflight	108	Ralph B. Washington	Systems Analyst	$96.75	23.6
25	Starflight	118	James J. Frommer	General Support	$18.36	30.5
25	Starflight	112	Darlene M. Smithson	DSS Analyst	$45.95	41.4

Step2: Identify the Primary Key

PROJ_NUM is not an adequate primary key because the project number doesn't uniquely identify the rows. To maintain a proper primary key that will uniquely identify any attribute value, the new key must be composed of a combination or PROJ_NUM and EMP_NUM.

Step3: Identify all dependencies

The identification of PK is already identified the following dependency:

PROJ_NUM, EMP_NUM → PROJ_NAME, EMP_NAME, JOB_CLASS, CHG_HOUR, HOURS

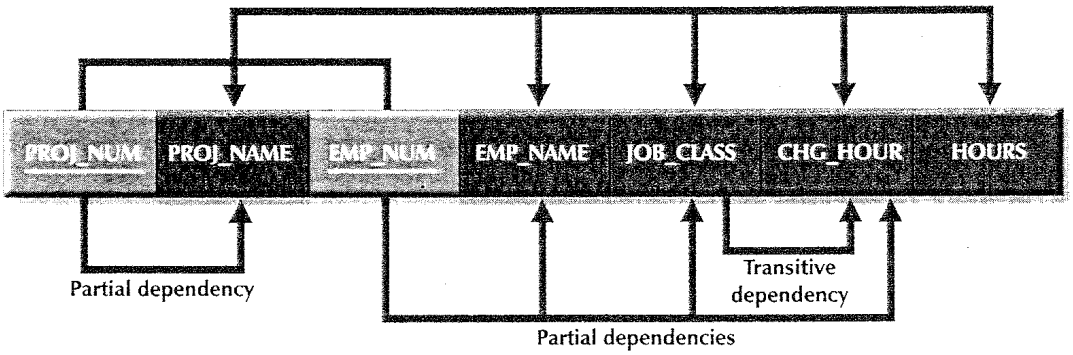
The project number identifies the project name.
PROJ_NUM → PROJ_NAME

Employee number identifies the details of employee

EMP_NUM → EMP_NAME, JOB_CLASS, CHG_HOUR

Job classification identifies the employee's charge per hour.
JOB_CLASS → CHG_HOUR

The following diagram shows the various dependencies:



Partial dependencies: A dependency based on only part of a composite key is called as partial dependency. PROJ_NUM determines the PROJ_NAME; that is PROJ_NAME is dependent on only part of the primary key. And EMP_NUM is required to find EMP_NAME, JOB_CLASS,CHG_HOUR.

Transitive dependencies: A transitive dependency is a dependency of one non-prime attribute on another non-prime attribute. CHG_HOUR is dependent on JOB_CLASS.

17. When is a table in 2NF?

Converting to 2NF is done only when the 1NF has a composite primary key. If the 1NF has a single attribute primary key, then the table is automatically in 2NF.

Step1: Write each key component on a separate line

Write each component on a separate line; then write the original(Composite key) on the last line.

PROJ_NUM
EMP_NUM
PROJ_NUM, EMP_NUM

Each component will become the key in a new table. In other words, the original table is now divided into three tables (PROJECT, EMPLOYEE and ASSIGNMENT).

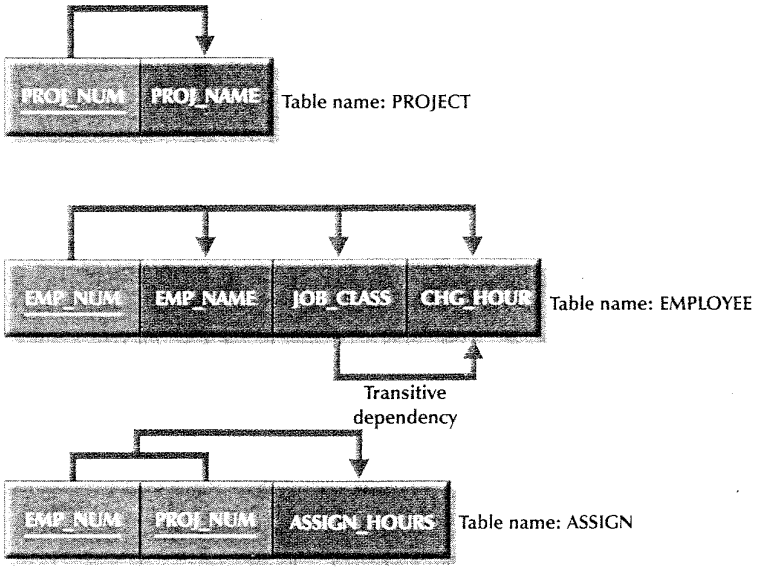
Step2: Assign corresponding dependent attributes

The three tables are described as follows:
PROJECT (PROJ_NUM, PROJ_NAME)

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS, CHG_HOUR)

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)

The number of hours spent on each project by each employee is dependent on both PROJ_NUM and EMP_NUM in the ASSIGNMENT table; place HOURS in the ASSIGNMENT table as ASSIGN_HOURS.



18. When is a table in 3NF?

A relation is said to be in 3NF if
(1). It is in 2NF
(2). There are no transitive dependencies

Step1: Identify each new determinant.

For every transitive dependency, write its determinant as a PK for a new table. A determinant is any attribute whose value determines other values within a row. If there are three dependencies, then there are three different determinants.

Write the determinant for this transitive dependency as:
JOB_CLASS

Step2: Identify the dependent attributes.

Identify the attributes that are dependent on each determinant identified in Step1 and identify the dependency.

JOB_CLASS - > CHG_HOUR

Step3: Remove the Dependent attributes from Transitive dependency.

Eliminate all dependent attributes in the transitive dependencies from each table that have such transitive relationship.

Eliminate CHG_HOUR from the employee table to leave the EMPLOYEE table dependency definition as:
EMP_NUM - > EMP_NAME, JOB_CLASS
Note that the JOB_CLASS remains in the EMPLOYEE table to serve as the FK.

The new tables are as follows:

PROJECT (**PROJ_NUM**, PROJ_NAME)

ASSIGN (**PROJ_NUM**, **EMP_NUM**, ASSIGN_HOURS)

EMPLOYEE (**EMP_NUM**, EMP_NAME, JOB_CLASS)

JOB (**JOB_CLASS**, CHG_HOUR)

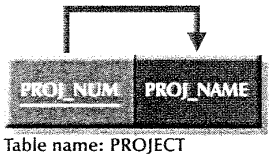


Table name: PROJECT

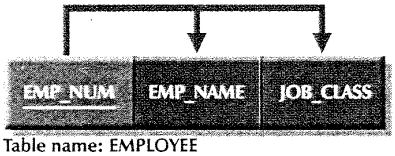


Table name: EMPLOYEE

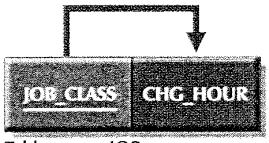


Table name: JOB

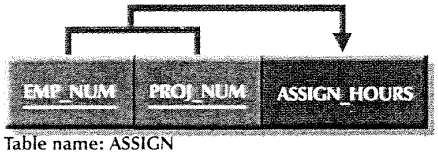


Table name: ASSIGN

19. Explain the different operators in Relational Algebra?

A. The eight relational algebra operators are

1. SELECT To retrieve specific tuples/rows from a relation.

2. PROJECT To retrieve specific attributes/columns from a relation.

3. PRODUCT To obtain all possible combination of tuples from two relations.

4. UNION To retrieve tuples appearing in either or both the relations participating in the UNION.

5. INTERSECT- To retrieve tuples appearing in both the relations participating in the INTERSECT.

6. DIFFERENCE To retrieve tuples appearing in the first relation participating in the DIFFERENCE but not the second.

7. JOIN To retrieve combinations of tuples in two relations based on a common field in both the relations.