

Unit 9

Database Systems

Database Introduction

A database is a collection of interrelated data. This data can be processed by one or more application systems.

Manual Database

Data is stored on a physical media.

Computerized Database

Databases kept on computers to enhance the functionality and efficiency in data storage and retrieval.

Types of computerized databases

- Flat-file model – All data stored on the same table.
- Relational model – Data are stored on different tables.
- Hierarchical model – The table structure has a parent-child relationship.
- Network model – The network model is derivative of the hierarchical model.
- Object relational model – Similar to relational model and databases are managed by using object oriented languages.

File System vs Database System

Flat file systems	Relational model
Data redundancy	Minimal data redundancy
Inconsistent data	Consistency of data
Inflexibility	Data independence
Limited data sharing	Sharing of data
Poor data control	Security and integrity control
Data isolation	Integration of data

Components of the relational database model

- Proposed by E.F.Codd in 1969.
- The relational model is a method of structuring data using tables(relations)
- Columns(attributes)
- Rows(tuples)

ANSI-SPARC Architecture

External level

How data is viewed by users

Conceptual schema/level

Describes the structure of the whole database

Physical schema/level

The way data is actually stored in physical media. This may include

1. Data structure & file organization
2. Storage space allocation for data and indexed

Mapping

The process of converting database from one level to another. There are two forms of this process

1. Conceptual/Internal mapping – Specifies how the conceptual fields and records are represented at the internal level.
2. External/Conceptual mapping – Defines correspondence between external views and the conceptual views.

Data independence

The ability to change the schema at one level of a database system without changing other levels.

1. Logical data independence – Change the conceptual schema without having to change the other schemas.
2. Physical data independence – Change the internal schema without change the conceptual or external schemas.

Databases

The database consists of data entered by the users under different relations (tables). It includes description about data and other access methods of data.

1. Raw data – data entered by users
2. Indexes – location of data on physical media.
3. Meta data – These describe the data types, validation rules and other constraints assigned to the each of the field.

Database management systems

The DBMS manage the overall operations of the database. The DBMS consists the following components.

1. Kernel
2. Design tools sub system
3. Runtime tools sub systems
4. Embedded programming language

Application program

Provides the interface for the users to interact with the database.

Database design framework

The development of the system is usually done in several phases which is known as system development life cycle.

1. Requirement analysis's – the data requirements and information requirements generated from the data is identified.
2. Conceptual database design – When the data that must be stored in database determine it is necessary to identify the dependency within the data
3. Entity relationship diagram (ERD) – The ER data model was introduced by Peter in 1976. This is

used to describe the conceptual schema of a database.

Component of an ER diagram

1. Entity – Person, place, object, event or concept in the user environment about which the organization wishes to maintain data.
 - i. Entity type – A collection or a set of entities that share common properties.
 - ii. Entity instance – Is a single occurrence of an entity type
2. Attribute – Is a property or a characteristic of an entity type that is related to the organization.
 - i. Simple attribute – Attributes that cannot be broken into simple components.
 - ii. Composite attributes – Attributes that can be divided into simple components.
 - iii. Stored attributes – Attributes that are directly store data entered by user.
 - iv. Derived attributes – Attribute whose values are calculated from stored attributes
 - v. Multi valued attributes – Attributes that can take more than one value.
 - vi. Descriptive attributes – Attributes connected to relationship types.
3. Relationship types – A meaningful association between or among entity types
 - i. Identifying relationship – Identifying relationship exists where the one entity depends on another entity
 - ii. Associative entity – Entity type which is created from a many to many relationship.

Degree of a relationship

1. Unary (degree 1) – Relationship among instances of one entity type.
2. Binary (degree 2) – Two entities are connected by one relationship.
3. Ternary (degree 3) – Three entities are connected through one relationship.
4. N-ary (degree N) – More than three entities are connected via one relationship.

Cardinality of a relationship

1. One to one (1:1)
2. One to many (1:M)
3. Many to many (M:N)

Logical Design

Relations/tables

- A 2d table consists of rows and columns of data.
- Consist of named columns and random number of rows.

Relational schema

The set of attributes within a relation along with the data type related of each instance of the attribute.

Relational instance

An occurrence of a relation

The concept of keys

A key is a field that used to sort data.

Primary key

Key which selected to uniquely identify the records in a table.

Candidate key

Attributes which uniquely identify the records in a table.

Alternate key

Other candidate keys which are not selected as the primary key of a relation.

Composite key

A key consisting of more than one attribute.

Foreign key

Foreign key is an attribute or a set of attributes in one relation that is a primary key in another relation.

Relational data integrity

The relational data model includes several types of constraints or business rules.

1. Domain integrity constraint – All of the values that appears in a column of a relation must be taken from the same domain.
 - i. Not null
 - ii. Unique
 - iii. Default value
2. Entity integrity – This is designed to assure that every relation has a primary key and it guarantees that every primary key is not null.
3. Referential integrity – In relational model associations between tables are defined through the use of foreign keys.

Database normalization

Unwanted duplication in a table may waste space and lead to loss of accuracy and consistency of data in the database.

Anomaly

An error or inconsistency that may result, when a user tries to interact with a table that contains redundant data.

1. Insert anomaly – An insert anomaly occurs when certain attributes cannot be inserted into the database without the presence of other attributes.
2. Update anomaly – When one or more instances of duplicated data is updated, but not all.
3. Delete anomaly – When certain attributes are lost because of the deletion of other attributes.

Functional dependencies

When one attribute of a relation depends on one or more attributes of the same relation.

Partial functional dependency

When an attribute depends on a part of the primary key of a relation which has a composite key.

with the primary key attributes of all the entities as foreign keys.

Transitive functional dependency

A non-key attribute depends on another non-key attribute.

Normal forms in database

Normalization is a staged process that follows a set of rules. When normalization is done on table the table is broken into several sub-tables to reduce data redundancy and increase the flexibility of the relationships.

1. Zeroth normal form (0NF) – The raw table without any rules applied.
2. First normal form (1NF) – If it does not contain repeating groups or multivalued attributes.
 - i. By separating table into two tables
 - ii. By filling the empty cells with relevant data.
3. Second normal form (2NF) – If the relation is in 1NF & every non key attribute is fully functionally depends on the primary key and does not have any partial dependencies.
4. Third normal form (3NF) – If the relation is in 2NF and there are no transitive dependencies in the table.

Transformation of conceptual design into the logical design

Transferring conceptual schema from a high level data model into the implementation data model which is known as the logical schema, is called data model mapping.

1. Step 1 – Create a relation that includes all the simple attributes.
 - i. Include only the simple component attributes of a composite attribute.
 - ii. Choose one of the key attributes of the entity as the primary key of the relation.
2. Step 2 – Create a relation and include all the
 - i. Include the primary key of the strong entity.
3. Step 3
 - i. When the two entities are connected by 1 to 1 relationship, the primary attribute of one entity is kept as the foreign key attribute in the other relation.
 - ii. When 2 entities are connected by 1-many the primary key of one side should be placed as the foreign key on the many side
 - iii. When 2 entities are connected by many to many relationship, a new relation should be created with primary keys.
4. Step 4 – Create a relation for multi valued attributes and include the primary key attribute of the entity as the foreign key.
5. Step 5 – If a relationship has more than 2 entities connected to it, a new relation should be created