Chapter Two

Database Model

A database model is a conceptual description of how the database works. It describes how the data elements are stored in the database and how the data is presented to the user and programmer for access; and the relationship between different items in the database. A specific DBS has its own specific Data Definition Language, but this type of language is too low level to describe the data requirements of an organization in a way that is readily understandable by a variety of users. We need a higher-level language. Such a higher-level is called database model.

Database Model: a set of concepts to describe the structure of a database, and certain constraints that the database should obey.

A database model is a description of the way that data is stored in a database. Database model helps to understand the relationship between entities and to create the most effective structure to hold data.

Database Model is a collection of tools or concepts for describing:

- Data
- Data relationships
- Data semantics
- Data constraints

The main purpose of database model is to represent the data in an understandable way.

Categories of database models include:

- Object-based
- Record-based
- Physical

Record-based Data Models

Consist of a number of fixed format records. Each record type defines a fixed number of fields, Each field is typically of a fixed length.The following are examples of this database model category.

- Hierarchical Database Model
- Network Database Model
- Relational Database Model

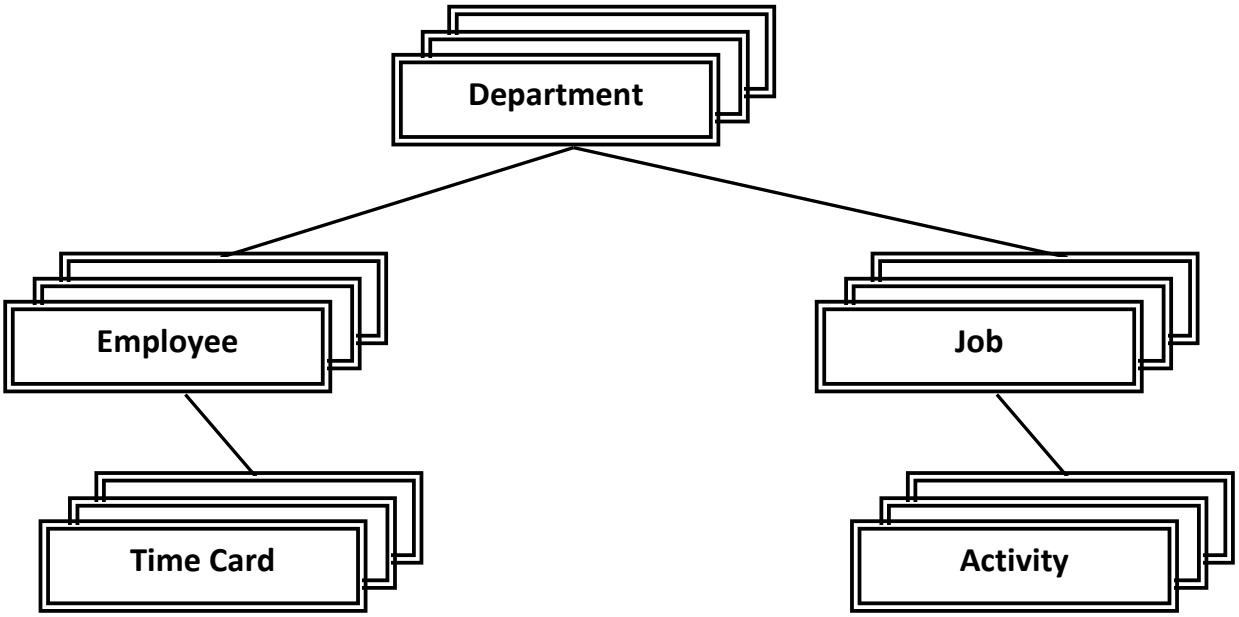
1. Hierarchical Model

In this model, the data is organized in a tree structure that originates from a root, and each class of data resides at different levels along a particular branch of the root. The data structure at each class level is called a **node**. There is always a single root node which is usually owned by the system or DBMS. Each of the pointers in the root then will point to (child) nodes there by depicting a parent-child sort of relationship. Searches are done by traversing the tree up and down with known search algorithms and modules supplied by the DBMS or may, for special cases, be designed by the application programmer. The initial structure of the database must be defined by the application programmer when the database is created. From this point on, the parent-children structure can't be changed without redesigning the whole structure.

Generally, Hierarchical database model is:

- The simplest database model
- Record type is referred to as node or segment
- The top node is the root node
- Nodes are arranged in a hierarchical structure as sort of upside-down tree
- A parent node can have more than one child node
- A child node can only have one parent node
- The relationship between parent and child is one-to-many and one-to-one
- Relation is established by creating physical link between stored records (each is stored with a predefined access path to other records)

- To add new record type or relationship, the database must be redefined and then stored in a new form.



Advantages of Hierarchical Database Model:

- Hierarchical Model is simple to construct and operate on.
- Corresponds to a number of natural hierarchically organized domains-e.g., assemblies in manufacturing, personnel organization in companies.
- Language is simple; uses constructs like GET, GET UNIQUE, GET NEXT, GET NEXT WITHIN PARENT etc.

Disadvantages of Hierarchical Database Model:

- Navigational and procedural nature of processing.
- Database is visualized as a linear arrangement of records.
- Little scope for "query optimization".

2. Network Model

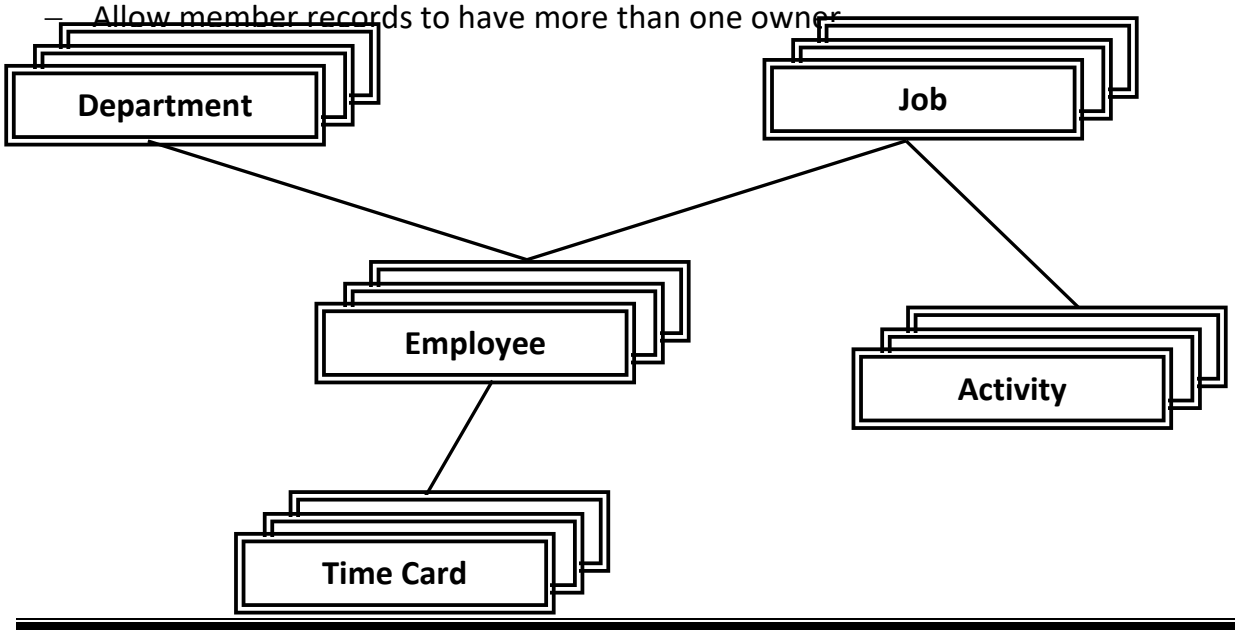
The network is a conceptual description of databases where many-to-many (multiple parent-children) relationships exist. To make this model easier to understand, the relationships between the different data items are commonly referred to as sets to distinguish them from the strictly parent-child relationships defined by the HDBM.

The network model uses pointers to map the relationships between the different data items. The flexibility of the NDB model is in showing many-to-many relationships is its greatest strength, though the flexibility comes at a price (the interrelationships between the different data sets become extremely complex and difficult to map).

Like the HDBM, NDBMs can very quickly be searched, especially through the use of index pointers that lead directly to the first item in a set being searched. The NDBM suffers from the same structural problem as the HDBM; the initial design of the database is arbitrary, and once its setup, any changes to the different sets require the programmer to create an entirely new structure. The dual problems of duplicated data and inflexible structure led to the development of a database model that minimizes both problems by making relationships between the different data items the foundation for how the database is structure.

Generally, Network database model is

- Allows record types to have more that one parent unlike hierarchical
- A network database models sees records as set members
- Each set has an owner and one or more members
- Allows/supports many to many relationship between entities
- Like hierarchical model network model is a collection of physically linked records.
- Allow member records to have more than one owner



Advantages of Network Data Model:

- Network Model is able to model complex relationships and represents semantics of add or delete on the relationships.
- Can handle most situations for modeling using record types and relationship types.
- Language is navigational; uses constructs like FIND, FIND member, FIND owner, FIND NEXT within set, GET etc. Programmers can do optimal navigation through the database.

Disadvantages of Network Data Model:

- Navigational and procedural nature of processing.
- Database contains a complex array of pointers that thread through a set of records.
- Little scope for automated "query optimization".

3. Relational Database Model

The relational database model is a way of looking at data - that is, it is a prescription for a way of representing data (namely, by means of tables), and a prescription for a way of manipulating such data (by means of operators). More precisely, the relational database model is concerned with three aspects of data: data structure (objects), data integrity, and data manipulation (operators).

The primary purpose behind the relational database model is the preservation of data integrity. To be considered truly relational, a DBMS must completely prevent access to the data by any means other than queries handled by the DBMS itself. While the relational model does not specify how the data is stored on the disk, the preservation of data integrity implies that the data must be stored in a format that prevents it from being accessed from outside the DBMS that created it.

The relational model also requires that the data be accessed through programs that don't rely on the position of the data in the database. This is in direct contrast to the other database models, where the program has to follow a series of pointers to the data it wants. A program querying a relational database simply asks for the data it wants, and it is up to the DBMS to do the necessary searches and provide the answer. Searches can be speed up

by creating an index on one or more columns in a table; however, the DBMS controls and uses the index. The user has only to ask the DBMS to create the index, and it will be maintained and used automatically from that point on.

The relational database model has a number of advantages over the other models. The most important is its complete flexibility in describing the relationships between the various data items. Once the tables are created and relationships defined then users can query the database on any of the individual columns in a table or on the relationships between the different tables.

Changing the structure of the database objects is as simple as adding or deleting columns in a table. Creating new tables, deleting old tables etc. are also very simple. The major tasks that the designers of a relational database has to make concerns the definitions of the tables and their relationships in the database.

Generally, Relational database model is

- Developed by Dr. Edgar Frank Codd in 1970 (famous paper, 'A Relational Model for Large Shared Data Banks').
- Terminologies originates from the branch of mathematics called set theory and relation.
- Can define more flexible and complex relationship.
- Viewed as a collection of tables called “Relations” equivalent to collection of record types.
- Relation: Two dimensional table.
- Stores information or data in the form of tables → rows and columns.
- A row of the table is called tuple→ equivalent to record.
- A column of a table is called attribute→ equivalent to fields.
- Data value is the value of the Attribute.
- Records are related by the data stored jointly in the fields of records in two tables or files. The related tables contain information that creates the relation.
- The tables seem to be independent but are related some how.
- No physical consideration of the storage is required by the user.

- Many tables are merged together to come up with a new virtual view of the relationship.

Alternative terminologies		
Relation	Table	File
Tuple	Row	Record
Attribute	Column	Field

- The rows represent records (collections of information about separate items).
- The columns represent fields (particular attributes of a record).
- Conducts searches by using data in specified columns of one table to find additional data in another table.
- In conducting searches, a relational database model matches information from a field in one table with information in a corresponding field of another table to produce a third table that combines requested data from both tables.