DATABASE MANAGEMENT SYSTEM
1 (Basic SQL)

Khryss Leanne B. Omnes
Mc Joshua de Lima

Bachelor of Science in Computer Science
College of Arts, Sciences and Technology

Table of Contents

Module 4: Introduction to MySQL

Course Code: **CS 2112**

Course Description: This course covers organization of files, database concepts, analysis and design with emphasis on relational databases. It covers the conceptual design of databases using the concepts of the entity-relationship model and normalization. This also discusses the different types of formal and commercial query languages with emphasis on SQL. It also covers security and integrity, query optimization and processing.

Course Intended Learning Outcomes (CILO):

At the end of this program, graduates will have the ability to:

1. To design database using data models.
2. To use database programming languages to manipulate or process data efficiently.
3. To understand the fundamental in designing and implementing database systems.
4. Learn the basic language of SQL
5. Exhibit the values of love, respect and humility as an individual in his workplace and community.

Course Requirements:

- **Major Exams** **- 40%**

Periodic Grade 100%

Final Grade = Total CS + Final Exam x 70% + 30% of the Midterm

MODULE 1
INTRODUCTION TO DATABASE

Introduction

An agency ought to have correct and dependable data (facts) for powerful choice making. Data (facts) is the spine and maximum vital useful resource of an agency that allows managers and agencies to advantage an aggressive edge. In this age of facts explosion, in which human beings are bombarded with data, getting the proper facts, withinside the proper amount, on the proper time isn't always a clean task. So, most effective the ones agencies will live to tell the tale that correctly control facts.

An information system simplifies the tasks of managing the information and extracting helpful information in a very timely fashion. An information system is associate integrated assortment of connected files, in conjunction with the small print of the interpretation of the information. a knowledge Base Management System could be a software package or program that enables access to data contained in a very information. the target of the DBMS is to supply a convenient and effective methodology of shaping, storing, and retrieving the knowledge hold on within the information.

The database and database management systems have become essential for managing business, governments, schools, universities, banks etc.

Learning Outcomes

At the end of this module, students should be able to:

1. establish the various elements of a info.
2. acknowledge the benefits and drawbacks of direction System.
3. Differentiate the users of the info.
4. Examine the various languages of the info

Lesson 1. The Database

An info is associate degree organized assortment of structured information, or data, usually keep electronically during a computing system. An info is typically controlled by a direction system (DBMS). Together, the info and also the software, in conjunction with the applications that are related to them, are brought up as an info system, usually shortened to merely info (Gupta & Mittal, 2017, pp. 1).

Data at intervals the foremost common forms of knowledgebases operative nowadays is often shapely in rows and columns during a series of tables to create process and data querying economical. the info will then be simply accessed, managed, modified, updated, controlled, and arranged. Most knowledgebases use structured search language (SQL) for writing and querying data (Gupta & Mittal, 2017, pp. 1).

According to the study of Gupta & Mittal (2017), A Database consists of four components as shown in Figure 1.1.

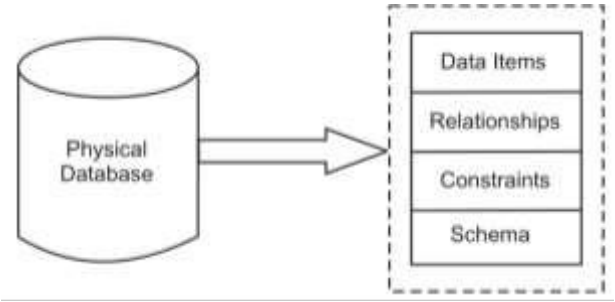


Figure 1.1 Components of a Database

1. **Data item**: It is defined as a distinct piece of information and is explained in the previous section.

2. **Relationships**: It represents a correspondence between various data elements.

3. **Constraints**: These are the predicates that define correct database states.

4. **Schema**: It describes the organization of data and relationships within the database. The schema consists of definitions of the various types of record in the database, the data-items they contain and the sets into which they are grouped. The storage structure of the database is described by the storage schema. The conceptual schema defines the stored data structure. The external schema defines a view of the database for particular users.

Lesson 2. The Database Management System

A database management system (DBMS) is that the software package used to outline, create, use, and maintain an information. It usually consists of many software package modules, every with their own practicality (Lemahieu et al., 2018, p. 63).

Popular software package vendors are Oracle, Microsoft, and IBM. MySQL could be a well-better-known ASCII text file software package. the mixture of a software package and an information is then usually known as a database system (Lemahieu et al., 2018, p. 63).

The DBMS performs the following five primary functions:

1. Define, create and organize a database: The DBMS establishes the logical relationships among different data elements in a database and also defines schemas and subschemas using the DDL.

2. Input data: It performs the function of entering the data into the database through an input device (like data screen, or voice activated system) with the help of the user.

3. Process data: It performs the function of manipulation and processing of the data stored in the database using the DML.

4. Maintain data integrity and security: It allows limited access of the database to authorized users to maintain data integrity and security.

5. Query database: It provides information to the decision makers that they need to make important decisions. This information is provided by querying the database

using SQL.

Lesson 3. Components of DBMS Environment

According to Gupta & Mittal (2017), a DBMS has three main components. These are Data Definition Language (DDL), Data Manipulation Language and Query Facilities (DML/SQL) and software for controlled access of Database as shown in Figure 1.2.

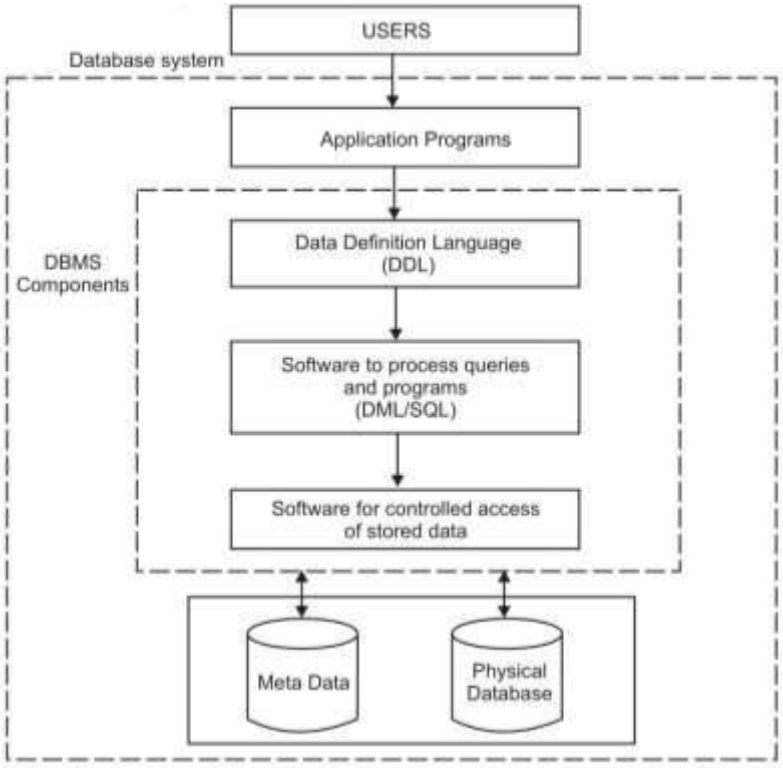


Figure 1.2 Components of DBMS

Data Definition Language (DDL): It allows the users to define the database, specify the data types, data structures and the constraints on the data to be stored in the database.

Data Manipulation Language (DML) and Query Language: DML allows users to insert, update, delete and retrieve data from the database. SQL provides general query facility.

Software for Controlled Access of Database: This software provides the facility of controlled access of the database by the users, concurrency control to allow shared access of the database and a recovery control system to restore the database in case of hardware or software failure.

Lesson 4. Database Users

The users of a database system can be classified into various categories depending upon their interaction and degree of expertise of the DBMS (Gupta & Mittal, 2017, p. 12).

End users or Naive users: The end users or naive users utilize the database framework through a menu-situated application program, where the sort and scope of reaction is constantly shown on the screen. The client need not know about the nearness of the database framework and is told through each progression. A client of an ATM falls in this class (Gupta & Mittal, 2017, p. 12).

Online users: These types of users communicate with the database legitimately through an online terminal or by implication through an application program and UI. They think about the presence of the database framework and may have some information about the constrained communication they are allowed (Gupta & Mittal, 2017, p. 12).

Application Programmers: A direction system (DBMS) is that the software package used to outline, create, use, and maintain a information. It usually consists of many software package modules, every with their own practicality (Gupta & Mittal, 2017, p. 12).

Popular software package vendors are Oracle, Microsoft, and IBM. MySQL could be a well-better-known ASCII text file software package. the mixture of a software package and a information is then usually known as a database system (Gupta & Mittal, 2017, p. 12).

Database Administrator: Database Administrator (DBA) is a person who have complete control over database of any enterprise. DBA is responsible for overall performance of database. He is free to take decisions for database and provides technical support. He is concerned with the Back-End of any project (Gupta & Mittal, 2017, p. 12).

Some of the main responsibilities of DBA are as follows:

1. Deciding the abstract schema or contents of database: DBA decides the info fields, tables, queries, data types, attributes, relations, entities otherwise you will say that he's accountable for overall logical style of info.

2. Deciding the interior schema of structure of physical storage: DBA decides however the info is truly keep at physical storage, however knowledge is depicted at physical

storage.

3. Deciding users: DBA provides permission to users to use info. while not having correct permission, nobody will access knowledge from info.
4. Deciding user view: DBA decides different views for various users.
5. Granting of authorities: DBA decides that user will use which portion of info. DBA provides authorities or rights to knowledge access. User will use solely that knowledge on that access right is granted to him.
6. Deciding constraints: DBA decides numerous constraints over info for maintaining consistency and validity in database.
7. Security: Security is that the major concern in info. DBA takes numerous steps to create data safer against various disasters and unauthorized access of knowledge.
8. watching the performance: DBA is accountable for overall performance of info. DBA frequently monitors the info to take care of its performance and check out to enhance it.
9. Backup: DBA takes regular backup of info, so it is often used throughout system failure. Backup is additionally used for checking knowledge for consistency.
10. Removal of dump and maintain free house: DBA is accountable for removing inessential knowledge from storage and maintain enough free space for daily operations. He also can increase storage capability once necessary.
11. Checks: DBA additionally decides numerous security and validation checks over info to make sure consistency.
12. Liaisioning with users: Another task of the DBA is to liaisioning with users and make sure the availableness of the info they need and write the required external schemas

Advantages and Disadvantages of Database Systems

According to Gupta & Mittal (2017), the database systems provide the following advantages over the traditional file system:

1. Controlled redundancy: In a conventional report gadget, every utility software has its personal facts, which reasons duplication of not unusual place facts gadgets in a

couple of report. This duplication/redundancy calls for more than one updates for a unmarried transaction and wastes a number of garage space. We can't remove all redundancy because of technical reasons. But in a database, this duplication may be cautiously managed, meaning the database gadget is privy to the redundancy and it assumes the obligation for propagating updates.

2. Data consistency: The trouble of updating more than one documents in conventional report gadget results in erroneous facts as distinctive documents might also additionally comprise distinctive statistics of the identical facts object at a given factor of time. This reason wrong or contradictory statistics to its users. In database structures, this trouble of inconsistent facts is mechanically solved through
   controlling the redundancy.
3. Program facts independence: The conventional report structures are normally facts dependent, which means that the facts agency and get admission to techniques are dictated through the desires of the precise utility and the utility packages are advanced accordingly. However, the database structures offer an independence among the report gadget and alertness software, that permits for modifications at one degree of the facts without affecting others. This asset of database structures permits to alternate facts without converting the utility packages that method the facts.
4. Sharing of facts: In database structures, the facts is centrally managed and may be shared through all legal users. The sharing of facts method now no longer bests the prevailing packages also can percentage the facts withinside the database however new utility packages may be advanced to function on the prevailing facts. Furthermore, the necessities of the brand-new utility packages can be glad without
   developing any new report.
5. Enforcement of requirements: In database structures, facts being saved at one crucial place, requirements can without problems be enforced through the DBA. This guarantees standardized facts codecs to facilitate facts transfers among structures. Applicable requirements would possibly encompass any or all the following— departmental, installation, organizational, industry, corporate, country wide or international.

6. Improved data integrity: knowledge integrity means the information contained within the info is each correct and consistent. The centralized management property

   enable adequate checks will be incorporated to supply knowledge integrity. One integrity sure} ought to be incorporated within the info is to make sure that if there's a respect to certain object, that object should exist.

7. Improved security: info security means that protective the information contained within the database from unauthorized users. The DBA ensures that correct access procedures are followed, as well as correct authentical schemes for access to the database management system and extra checks before allowing access to sensitive knowledge. The extent of security can be totally different for varied varieties of knowledge and operations.

8. Knowledge access is efficient: The info system utilizes totally different refined techniques to access the keep data terribly with efficiency.

9. Conflicting needs will be balanced: The DBA resolves the conflicting requirements of assorted users and applications by knowing the general requirements of the organization. The DBA will structure the system to supply an overall service that's best for the organization.

10. Improved backup and recovery facility: Through its backup and recovery system, the info system provides the facilities for convalescent from hardware or software package failures. The recovery system of the info system ensures that the database is reconditioned to the state it had been in before the program started death penalty, just in case of system crash.

11. Marginal program maintenance: in an exceedingly ancient classification system, the applying programs with the outline of knowledge and therefore the logic for

    accessing the information are designed severally. Thus, changes to the information formats or access strategies leads to the necessity to switch the applying programs. Therefore, high maintenance effort is needed. These are reduced to marginal in info systems because of independence of knowledge and application programs.

12. Knowledge quality is high: the standard of knowledge in info systems are terribly high as compared to ancient file systems. This is often potential because of the presence of tools and processes within the info system.
13. Sensible knowledge accessibility and responsiveness: The info systems offer question languages or report writers that enable the users to raise accidental queries to get the required information forthwith, while not the necessity to write down application programs (as just in case of file system), that access the knowledge from the info. This is often potential because of integration in info systems.
14. Concurrency control: The info systems are designed to manage synchronic (concurrent) access of the database by several users. They additionally stop any loss of data or loss of integrity because of these simultaneous accesses.
15. Economical to scale: In info systems, the operational knowledge of a company is kept in an exceedingly central info. The applying programs that job on this knowledge will be designed with terribly less price as compared to ancient classification system. This reduces overall prices of operation and management of the info that ends up in a cheap scaling.
16. Increased coder productivity: The info system provides several commonplace functions that the programmer would usually need to write in classification system. The supply of those functions permits the programmers to focus on the particular practicality needed by the users without concern about the implementation details. This will increase the general productivity of the coder and additionally reduces the event time and price.

According to Gupta & Mittal (2017), in contrast to many advantages of the database systems, there are some disadvantages as well. The disadvantages of a database system are as follows:

1. Quality increases: the information structure might become a lot of complicated owing to the centralized info supporting several applications in a corporation. This could

cause difficulties in its management and will need professionals for management.

2. Demand of a lot of disk space: the wide practicality and more quality increase the dimensions of software package. Thus, it needs far more area to store and run than the standard classification system.
3. Further cost of hardware: the value of info system's installation is way a lot of. It depends on atmosphere and practicality, size of the hardware and maintenance prices of hardware.
4. Value of conversion: the value of conversion from previous file-system to new info system is incredibly high. In some cases, the value of conversion is therefore high that the cost of software package and further hardware becomes insignificant. It additionally includes the value of coaching force and hiring the specialized manpower to convert and run the system.
5. Would like of further and specialized manpower: any organization having info systems, got to be rent and train its force on regular basis to style and implement info and to produce database administration services.
6. Would like for backup and recovery: for an info system to be correct and obtainable

   all times, a procedure is needed to be developed and used for providing backup copies to all or any its users once harm happens.
7. Structure conflict: a centralized and shared info system needs a agreement on information definitions and possession in addition as responsibilities for correct data maintenance.
8. A lot of installation and management cost: the massive and complete info systems are pricier. They need trained force to control the system and has further annual maintenance and support prices.

Lesson 5. 3-Layer Architecture

Based from DBMS - Architecture - Tutorialspoint, n.d., a 3-tier architecture separates its tiers from each other based on the complexity of the users and how they use the data present in the database. It is the most widely used architecture to design a DBMS.

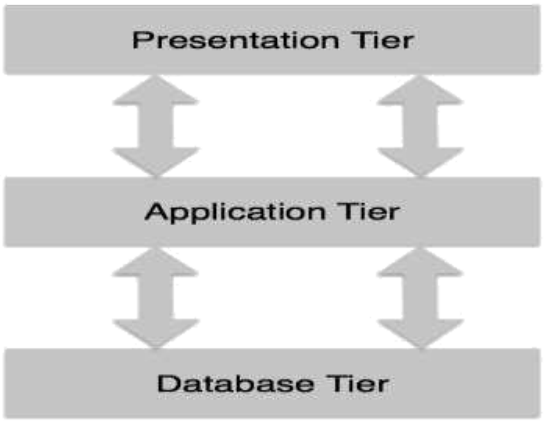


Figure 1.2 Three-Layer Architecture

Database (Data) Tier: At this level, the database lives alongside its inquiry preparing dialects. We likewise have the relations that characterize the information and their imperatives at this level.

Application (Middle) Tier: At this level live the application worker and the projects that get to the database. For a client, this application level presents a disconnected perspective on the database. End-clients are unconscious of any presence of the database past the application. At the opposite end, the database level doesn't know about some other client past the application level. Thus, the application layer sits in the center and goes about as an arbiter between the end-client and the database.

Client (Presentation) Tier: End-clients work on this level and they know nothing about any presence of the database past this layer. At this layer, various perspectives on the database can be given by the application. All perspectives are created by applications that dwell in the application level.

Lesson 6. Types of Data Models

Conforming in the study conducted by Gupta & Mittal (2017, pp. 18–21), an info model includes completely different information models, every describing the info from completely different views. It provides a transparent and unambiguous description of the info things, their relationships, and numerous information constraints from a selected perspective. many styles of information models are developed throughout an info style method.

Characteristics of knowledge Models

A data model must possess the following characteristics so that the best possible data representation can be obtained.

1. Diagrammatic representation of the data model.
2. Simplicity in designing i.e., Data and their relationships can be expressed and distinguished easily.
3. Application independent, so that different applications can share it.
4. Data representation must be without duplication.
5. Bottom-up approach must be followed.
6. Consistency and structure validation must be maintained.

Types of Data Models

The various data models can be divided into three categories, such as

1. Record Based Data Models: These models represent data by using the record structures. These models lie between the object-based data models and the physical data models (Gupta & Mittal, 2017). These data models can be further categorized into three types:
    a) Hierarchical Data Model
    b) Network Data Model
    c) Relational Data Model.

2. Object Based Data Models: These models are used in describing the data at the logical and user view levels. These models allow the users to simplicity specify the constraints in the data (Gupta & Mittal, 2017). These data models can be further categorized into four types:
    a) Entity Relationship Model (ER-Model)
    b) Object Oriented Model
    c) Semantic Data Model
    d) Functional Data Model.

3. Physical Data Models: These models provide the concepts that describes the details of how the data is stored in the computer along with their record structures, access paths and ordering. Only specialized or professional users can use these models

(Gupta & Mittal, 2017). These data models can be divided into two types:

a) Unifying Model.

b) Frame Memory Model.

Lesson 7. Data schemas

Based from DBMS - Data Schemas - (Tutorialspoint, n.d.), a database schema is the skeleton structure that represents the logical view of the entire database. It defines how the data is organized and how the relations among them are associated. It formulates all the constraints that are to be applied on the data.

A database schema defines its entities and the relationship among them. It contains a descriptive detail of the database, which can be depicted by means of schema diagrams. It's the database designers who design the schema to help programmers understand the database and make it useful (DBMS - Data Schemas - Tutorialspoint, n.d.).

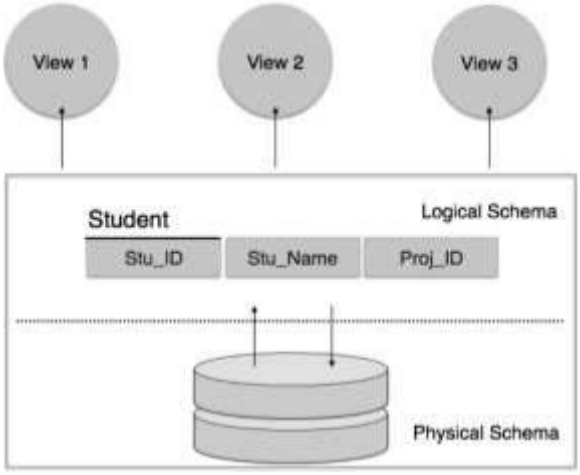


Figure 1.3 Database Schema

A database schema can be divided broadly into two categories

1. **Physical Database Schema:** This schema pertains to the actual storage of data and its form of storage like files, indices, etc. It defines how the data will be stored in a secondary storage.
2. **Logical Database Schema**: This schema defines all the logical constraints that need to be applied on the data stored. It defines tables, views, and integrity constraints.

Lesson 8. Data Independence

Data Independence is defined as the characteristics of a database system to change the schema at one level without having to change the schema at the next higher level. It can also be defined as the immunity of the application programs to change in the physical representation and access techniques of the database. The above definition says that the application programs do not depend on any particular physical representation or access technique of the database. The DBMS achieved the data independence by the use of three-level architecture (DBMS - Data Independence - Tutorialspoint, n.d.). The data independence is of TWO types:

1. **Physical Data Independence**: It indicates that the physical storage structures or devices used for storing the data could be changed without changing the conceptual view or any of the external views. Only the mapping between the conceptual and internal level is changed. Thus, in physical data independence, the conceptual schema insulates the users from changes in the physical storage of the data (DBMS - Data Independence - Tutorialspoint, n.d.)..

2. **Logical Data Independence**: It indicates that the conceptual schema can be changed without changing the existing external schemas. Only the mapping between the external and conceptual level is changed and absorbed all the changes of the conceptual schema. DBMS that supports logical data independence, changes to the conceptual schema is possible without making any change in the existing external schemas or rewriting the application programs. Logical data independence also insulates application programs from operations like combining of two records into

one or splitting an existing record into more than one records (DBMS - Data Independence - Tutorialspoint, n.d.).

Lesson 9. DBMS Languages

According to Gupta & Mittal (2017, pp. 13–14), the DBMS provides completely different languages and interfaces for every class of users to precise information queries and update. once the look of the information is complete and also the DBMS is chosen to implement it, the primary issue to be done is to specify the abstract and internal schemas for the information and also the corresponding mappings. the subsequent 5 languages are accessible to specify completely different schemas.

Data Definition Language (DDL) it's wont to specify a information abstract schema victimization set of definitions. It supports the definition or declaration of information objects. several techniques are accessible for writing DDL. One wide used technique is writing DDL into a document (Gupta & Mittal, 2017, p. 13).

Storage Definition Language (SDL) it's wont to specify the inner schema within the information. The storage structure and access ways employed by the information system is specified by the required set of SDL statements. The implementation details of the information schemas are enforced by the required SDL statements and are sometimes hidden from the users (Gupta & Mittal, 2017, p. 13).

View Definition Language (VDL) it's wont to specify user’s views and their mappings to the abstract schema. However, usually, DDL is employed to specify each abstract and external schema in several DBMS’s. There are 2 views of knowledge the logical view—that is perceived by the engineer and physical view —data keep on storage devices (Gupta & Mittal, 2017, p. 13).

Data Manipulation Language (DML) It provides a collection of operations to support the

essential data manipulation operations on the info control within the information. it's wont

to question, update or retrieve information keep in a very information (Gupta & Mittal, 2017,

p. 13). A part of DML that give information retrieval is termed command language The DML is of the 2 types:

a) **Procedural DML**: It allows the user to tell the system what data is needed and how to retrieve it.
b) **Non-procedural DML**: It allows the user to state what data are needed, rather than how it is to be retrieved.

Fourth-generation Language (4-gL): It is a compact, efficient and non-procedural programming language used to improve the efficiency and productivity of the DBMS. In this, the user defines what is to be done and not how it is to be done. The 4-GL has the following components in it (Gupta & Mittal, 2017, p. 14). These are:

a) Query languages
b) Report
c) Spread sheets
d) Database languages
e) Application generators
f) High level languages to generate application program.

The following exercises is provided to strengthen your learning.

True or False

True/False

1. A database actually consists of three parts: information, the logical structure of that information, and tables.

2. A data dictionary, or relation, is a two-dimensional table used to store data within a relational database.

3. A database management system (DBMS) allows you to specify the logical organization for a database and access and use the information within a database.

4. A physical view represents how the users view the data.

5. A database may have numerous physical views.

6. Fixed length record sometimes wastes space while variable length record does not waste space.

7. A database is any collection of related data.

8. A DBMS is a software system to facilitate the creation and maintenance of a computerized database.

9. End-users can be categorized into casual, designer, or parametric users.

10. Data redundancy exists when the same data is stored at multiple places.

11. A database always maintains a collection of unrelated data.

12. A database system is a software system to enable users to create and maintain a computerized database.

13. End-users can be categorized into casual, naïve, sophisticated, or stand-alone users.

14. Typical DBMS functionality is to define and create a particular database in terms of its data types, structures, and constraints.

15. Data redundancy exists when the same data is stored at one place.

16. A database is a very large software system used for processing related data.

17. The DBMS stores definitions of the data elements and their relationships (metadata) in a data dictionary.

18. Data about data is metadata.

19. One of the main functions of a database system is to provide timely answers to end users.

20. To work with data, the DBMS must retrieve the data from permanent storage and place it in RAM.

Answer Key (Inverted):

eurT .02	eurT .31	eslaF .6
eurT .91	eurT .21	eslaF .5
eurT .81	eslaF .11	eslaF .4
eurT .71	eurT .01	eurT .3
eslaF .61	eslaF .9	eslaF .2
eslaF .51	eurT .8	eslaF .1
eurT .41	eurT .7	

Assessment Task

Fill in the Blanks

1. A(n) _____ contains the logical structure for the information.
2. A(n)_____ represents how data is physically stored on a storage device.
3. A(n)_____ represents how knowledge users see information.
4. A data model is a collection of concepts that can be used to describe the _____ of a database.
5. _____ schema describes physical storage structures and access paths.
6. In three-schema architecture, user views are defined at _____ schema.
7. _____ data model provides concepts that are close to the way many users perceive data.
8. External schema describes the various user _____.

9. A _____ is a collection of concepts that can be used to describe the structure of a database.

10. _____ data models use concepts such entities, attributes, and relationships.

11. Data stored in database at a particular moment in time is a database _____.

12. A _____ is a unit of work that includes one or more reads or updates of database records.

13. The description of schema constructs and constraints is called _____ .

14. The description of a database is called database _____.

15. The database state is called _____ of the schema.

Multiple Choice

1. Manager’s salary details are to be hidden from Employees Table. This technique is called

as

(a) Conceptual level data hiding

(b) Physical level data hiding

(c) External level data hiding

(d) Logical level data hiding

2. Which of the following is not a type of database management system?

(a) Hierarchical

(b) Network

(c) Relational

(d) Sequential.

3. A schema describes

(a) Data elements

(b) Records and files

(c) Record relationship

(d) All of the above.

4. Which data management language component enabled the DBA to define schema components?

(a) DML

(b) Subschema DLL

(c) Schema DLL

(d) All of these.

5. Which statement is false regarding data independence?

(a) Hierarchical data model suffers from data independence.

(b) Network model suffers from data independence.

(c) Relational model suffers from logical data independence.

(d) Relational model suffers from physical data independence.

6. Databases may be more expensive to maintain than files because of

(a) backup and recovery needs personnel

(b) the complexity of the database environment

(c) the need for specialized

(d) all of the above.

7. Typically, a database consists _____ but can support _____.

(a) table, queries

(b) information, data

(c) physical view, logical view

(d) information view, data view.

8. Which view of information deals with how the information is physically arranged, stored, and accessed?

(a) Physical view

(b) Logical view

(c) Information view

(d) None of the above

9. By redundancy in a file-based system we mean that

(a) unnecessary data is stored

(b) same data is duplicated in many files

(c) data is unavailable

(d) files have redundant data.

10. Data integrity in a file-based system may be lost because

(a) the same variable may have different values in different files

(b) files are duplicated

(c) unnecessary data is stored in files

(d) redundant data is stored in files.

References:

Codd's 12 Rules - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/dbms/dbms_codds_rules.htm

DBMS - Data Independence - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_independence.htm

DBMS - Data Schemas - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_schemas.htm

ER Model - Basic Concepts - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/er_model_basic_concepts.htm

Gupta, S. B., & Mittal, A. (2017). Introdution to Database Management System (2nd ed.). University Science Press. https://doi.org/10.1201/9781315372679-7

Lemahieu, W., Broucke, S. vanden, & Baesens, B. (2018). Principles of Database Management. Cambridge University Press. https://doi.org/10.16309/j.cnki.issn.1007-

1776.2003.03.004

MySQL - Administration - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/MySQL/MySQL-administration.htm

MySQL - Data Types - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/MySQL/MySQL-data-types.htm

Relation Data Model - Tutorialspoint. (n.d.). Retrieved August 25, 2020, from https://www.tutorialspoint.com/dbms/relational_data_model.htm

Williams, H. E., & Tahaghoghi, S. M. M. (2006). Learning MySQL (A. Oram (Ed.); 1st ed.).

O’Reilly Media.

MODULE 2
ER / EER MODEL

Introduction

The ER model defines the abstract read of an information. It works around real-world entities and therefore the associations among them. At read level, the ER model is taken into account an honest possibility for planning databases.

The entity relationship (ER) model was introduced and formalized by Peter bird genus in 1976. it's one amongst the foremost well-liked information models for abstract data modeling. The ER model has a lovely and easy graphical notation. Hence, it's the best properties to make a abstract information model. it's 3 building blocks: entity varieties, attribute varieties, and relationship varieties. we have a tendency to elaborate on these in what follows. we have a tendency to conjointly cowl weak entity varieties and supply 2 samples of ER models.

Learning Outcomes

At the end of this module, students should be able to:

1. Recognize what are the parts of the ER Model.
2. Identify the different types of attributes.
3. Contrast the different relation sets.
4. Construct an ER Diagram.

Lesson 1. ER Model – Basic Concepts

Enterprise: Enterprise refers to any kind of organization.

Ex. Colleges, schools, banks, any company etc.

Entity: Entity refers to an “object” or “thing” in real world. Object may be any person, place, event etc.

Ex. Students of colleges and schools, loans in banks, employees in any company etc.

Attributes: These are the characteristics of any entity.

Ex:

(i) A student can be described by his name, age, address, height, class etc.

(ii) Loans can be described by their types such as house loan, car loan etc.

(iii) Employees in any company can be described by their Employee ID, name, department, designation etc.

(iv) A car can be described by his color, model, company etc.

Value: Value is that the info or information that is hold on in attributes of any entity.

Entity Sets: All the entities having same attributes create associate degree entity set.

Domain: Domain or worth set is that the set of all values or info concerning any attribute.

Ex. think about the coed table shown in Figure 2.1. It describes the essential ideas.

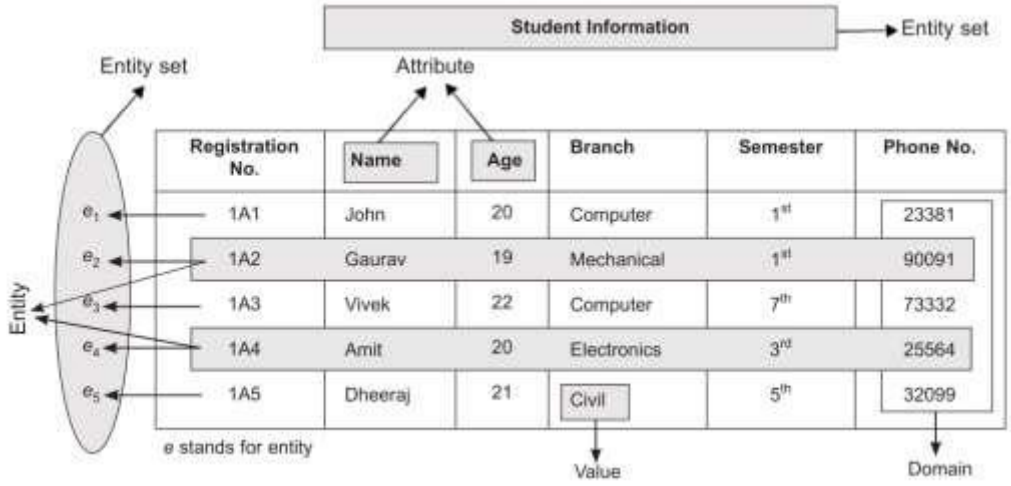


Figure 2.1 Basic Concept of ER Model

(a) **Enterprise**: Here, enterprise is college where students are studied.

(b) **Entity**: Here, entity refers to any single student with all his values.

(c) **Attributes**: The students are described by Registration No., Name, Age, Branch, Semester, Phone No. These are the attributes of students.

(d) **Value**: The values are 1A1, 21, Civil, Gaurav, 90091, 5th etc.

(e) **Entity Set**: All students are described by same set of attributes. So, all these students combine together to make an entity set “Student Information”.

(f) **Domain**: (Value set) for, attribute, Name it is John, Gaurav, Vivek, Amit, Dheeraj and for Age, it is 20, 19, 21, 22.

Lesson 2. Types of Attributes

As seen in ER Model - Basic Concepts - Tutorialspoint, n.d., entities are depicted by means that of their properties known as attributes. All attributes have values. to Illustrate, a student entity might have name, class, and age as attributes.

There exists a site or vary of values which will be assigned to attributes. to Illustrate, a student's name cannot be a numeric worth. it's to be alphabetic. A student's age cannot be negative, etc (ER Model - Basic Concepts - Tutorialspoint, n.d.).

Simple attribute: Simple attributes are atomic values, which cannot be divided further. For example, a student's phone number is an atomic value of 10 digits.

Composite attribute: Composite attributes are made of more than one simple attribute. For example, a student's complete name may have first_name and last_name.

Derived attribute: Derived attributes are the attributes that don't exist within the physical information; however their values are derived from alternative attributes gift within the information. to Illustrate, average_salary in an exceedingly department shouldn't be saved directly within the information, instead it is often derived. for one more example, age is often derived from data_of_birth. Single-value attribute:

Single-value attribute: Single-value attributes contain single value. For example: Social_Security_Number.

Multi-value attribute: Multi-value attributes may contain more than one values. For example, a person can have more than one phone number, email_address, etc.

These attribute types can come together in a way like:

simple single-valued attributes simple multi-valued attributes

composite single-valued attributes composite multi-valued attributes

Lesson 3. Relationship Sets

Based from ER Model - Basic Concepts - Tutorialspoint, n.d., the association among entities is termed a relationship. For example, associate degree worker works_at a

department, a student enrolls in an exceedingly course. Here, Works_at and Enrolls are known as relationships.

A group of relationships of comparable sort is termed a relationship set. Like entities, a relationship can also have attributes. These attributes are known as descriptive attributes (ER Model - Basic Concepts - Tutorialspoint, n.d.).

Descriptive Attributes: Attributes of any relationship set are known as descriptive attributes.

Degree of Relationship: The number of participating entities in a relationship defines the degree of the relationship.

Binary = degree
2 Ternary =
degree 3 n-ary =
degree

Cardinality defines the number of entities in one entity set, which can be associated with the number of entities of other set via relationship set.

One-to-one: One entity from entity set A can be associated with at most one entity of entity set B and vice versa.

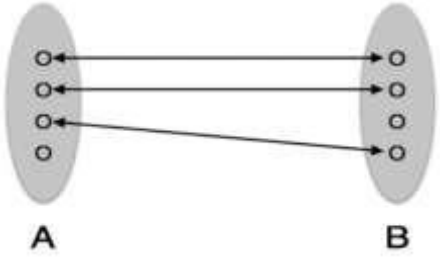


Figure 2.2 One to One Model

One-to-many: One entity from entity set A can be associated with more than one entities of entity set B, however an entity from entity set B can be associated with at most one entity.

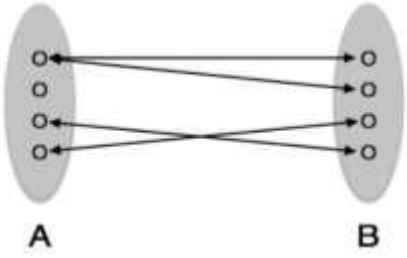


Figure 2.3 One to Many Model

Many-to-one: More than one entities from entity set A can be associated with at most one entity of entity set B, however an entity from entity set B can be associated with more than one entity from entity set A.

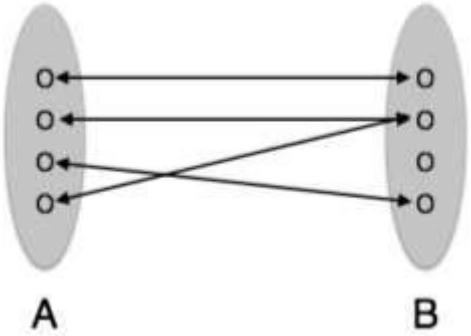


Figure 2.3 Many to One Model

Many-to-many: One entity from A can be associated with more than one entity from B and vice versa.

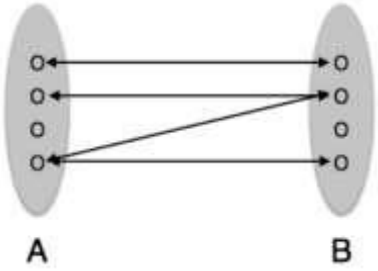
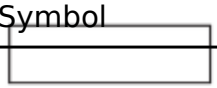
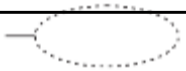
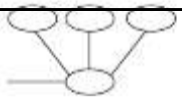


Figure 2.3 Many to Many Model

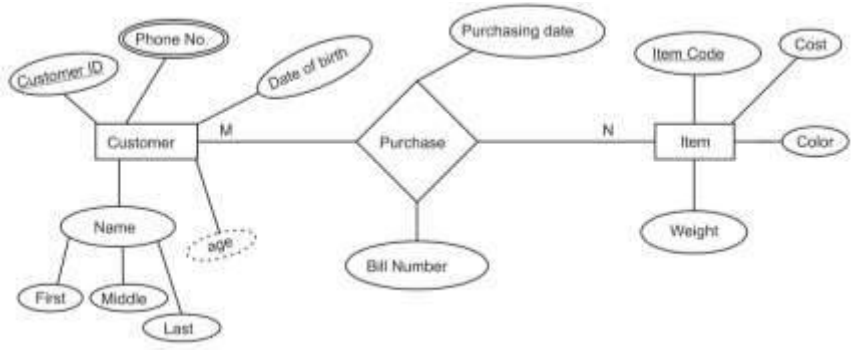
Lesson 4 ER Diagram

E-R diagrams represents the logical structure of a database. Symbols used in E-R diagrams is shown in Table 1.

Table 1. ER Diagram Symbols

Name of Symbol	Symbol	Meaning
Rectangle		Entity Set (Strong)
Double Rectangle		Entity Set (Weak)
Ellipse		Attribute
Diamond		Relationship Set
Double Diamond		Identifying Relationship Type
Double Ellipses		Multi-valued attributes
Dashed Ellipses		Derived attributes
Ellipse with line inside it		Key attribute
Ellipse joined with other ellipses		Composite attributes
Double lines		Total Participation
Single Line		Partial Participation
Triangle		Specialization or Generalization

Example: Make an E-R diagram having two entity sets, Customer and Item.



Cardinality Ratio is many to many because a customer can buy any number of items and same item can be purchased by more than one customer.

Advantages of E-R model

According to Gupta & Mittal (2017, pp. 53–54), the major advantages of E-R model are as follows:

1. Straightforward relation representation: The relation representation of the information model victimization E-R diagram are comparatively a lot of straightforward than alternative models.

2. Mapping with relative model: It are often simply mapped onto the relational model. The E-R diagrams employed in the E-R model will simply be remodeled into relative tables. The entities and attributes of E-R model will simply be remodeled into relations (tables) and columns (fields) in an exceedingly relative model.

3. Communication tool: it's terribly straightforward and simple to know with a minimum of coaching efforts needed. Therefore, the model is often utilized by the information designer to speak the planning to the top user.

4. Design tool: E-R model can even be used as a design arrange by the information developer to implement an information model in specific direction computer code.

5. Easy conversion to alternative models: E-R diagrams are often simply regenerate to a network or hierarchical information model.

6. Graphical illustration: E-R model provides graphical and represented representation of assorted entities, their attributes and relationships between entities.

7. straightforward to modify: Modifications to E-R diagram at later stage is comparatively easier than in alternative models.

Lesson 5. Enhanced Entity-Relationship (EER) Model

(EER) Model the improved Entity Relationship model or EER model is associate degree extension of the ER model. It includes all the modeling ideas (entity varieties, attribute varieties, relationship types) of the ER model, additionally as 3 new further linguistics information modeling concepts: specialization/generalization, categorization, and aggregation (Lemahieu et al., 2018, p. 174).

According to Gupta & Mittal (2017), The foremost necessary new modeling construct introduced by EER was superclass and subclass entity types. These are referred to as

supertype and subtype entities severally. By using these 2 entity types E-R model are often divided into a lot of specialized sub-models or will be part of some sub-models to create a generalized E-R model.

Superclass Entity Type (Supertype): A superclass entity type is a generic entity type that includes one or more distinct subclasses that require to be represented in a data model. It means members belong to subclass are same as the entity in the superclass. The relationship between a superclass and subclass is a one-to-one (1:1) relationship. In some cases, a superclass can have overlapping subclasses (Gupta & Mittal, 2017).

Subclass Entity Type (Subtype): A subclass entity type may be a lot of specialized entity type that incorporates a distinct role within the organization. A subclass group may be a member of taxon. it's one amongst the data-modeling abstractions employed in EER. A subclass group is also more divided and, in this case, it acts as superclass for its subclasses Gupta & Mittal, 2017).

The superclass/subclass relationship is shown in Figure 2.4

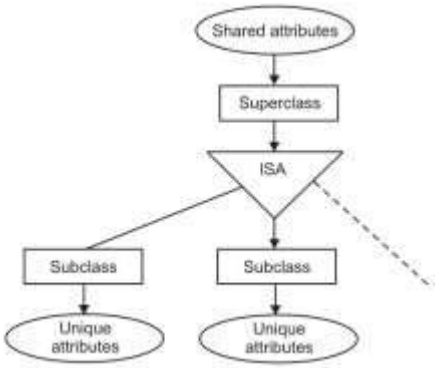


Figure 2.4 Superclass/subclass relationship

Think about the instance of a Bank as shown in Figure 2.5 during which PERSON entity is superclass entity sort which is more divided into EMPLOYEE and CUSTOMER entities. Here EMPLOYEE and CUSTOMER entities are subclass type (Gupta & Mittal, 2017, p. 56).

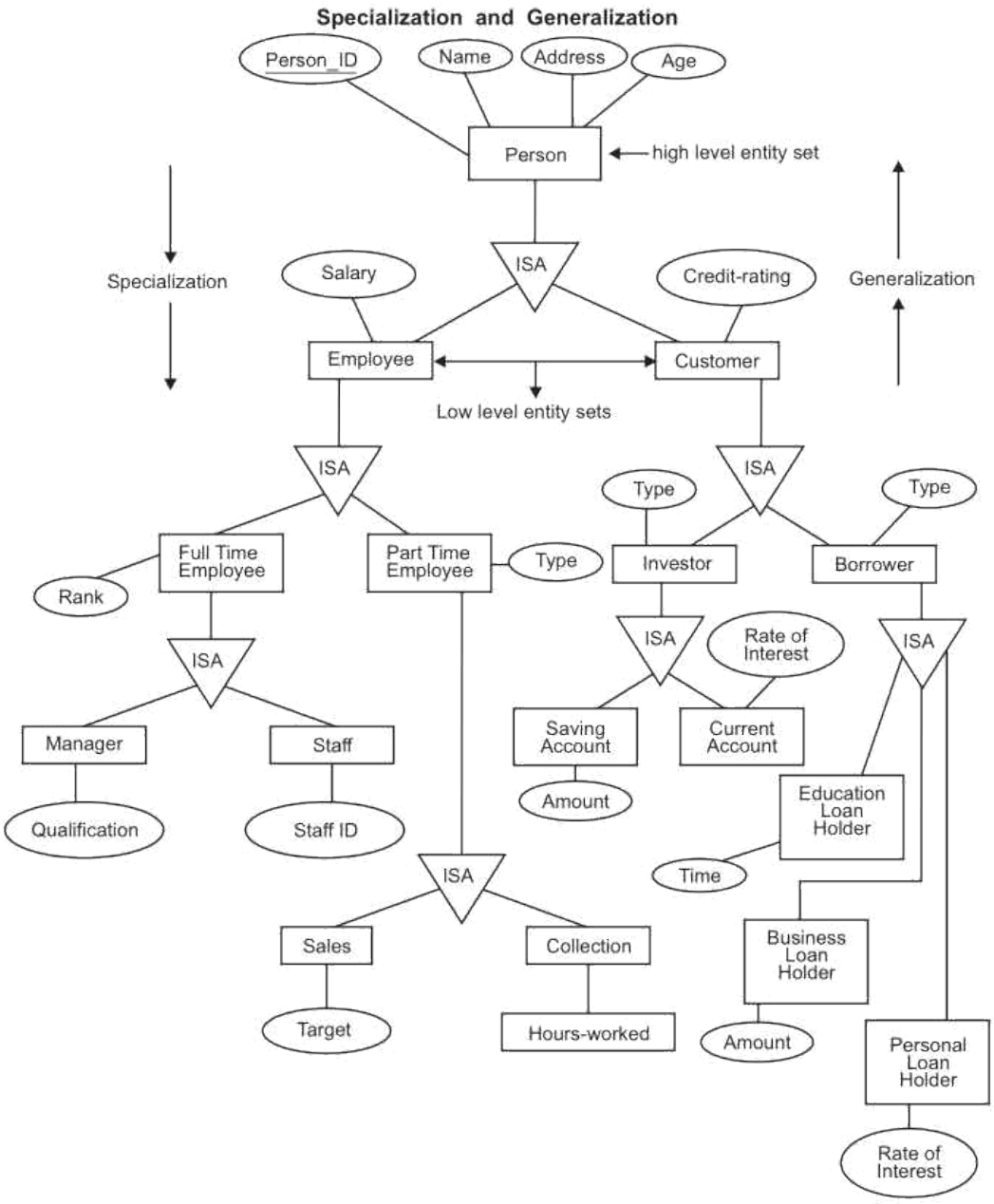


Figure 2.5 Specialization / Generalization

Basic conception is that by E-R model, an individual belongs to a Bank is understood, and by EER, however it belongs to Bank is understood means that, what's the precise relationship between Bank which person? an individual is also an EMPLOYEE or CUSTOMER which may be more categorized into Manager, EMPLOYEEs or Investor, Borrower and so on (Gupta & Mittal, 2017, p. 56).

Specialization includes subgrouping of entities within an entity set having some distinct nature then other entities. If deep info is required then go towards specialization. In other words, Specialization may be a method by that any existing entity set is split into smaller entity sets in step with the distinct or totally different nature of entities (Gupta & Mittal, 2017, p. 57).

Consider the example of Bank in Figure 2.5. Person is entity set of all folks that belongs to bank. Further Person is assessed into Employee and Customers of bank. So, Person entity set is split into EMPLOYEE entity set and CUSTOMER entity set. Employee are more classified into 2 classes full time employees and part time employees then on. Customers are classified into Investors and Borrowers then on (Gupta & Mittal, 2017, p. 57).

Generalization is a process by which two or more entity sets can be combined into a single entity set by determining similarities between the entities. It's an abstract view of any Enterprise. Generalization proceeds from the recognition that a number of entity sets share some common features. If an abstract view of information is needed then go towards generalization (Gupta & Mittal, 2017, p. 57).

Consider the example in Figure 2.5. Here Investor and Borrower are two entity sets. They have common feature that both are Customer of the Bank. Similarly, Employee entity set and Customer entity set can be combined into Person entity set.

Specialization and generalization lead to **attribute inheritance** between higher level entity set and lower level entity set. Inheritance is a process by which lower level entity set inherits (or taken) some properties of its higher-level entity set.

Consider the Figure 2.5. Here entity sets Employee and Customer inherits attributes Person_ID, Name, Address, Age from Person entity set.

Aggregation is an abstraction process in which a relationship set is considered as higher-level entity set. Consider an example of ternary relationship having three entity sets Employee, Job and Branch with relationship set works-on as shown in Figure 2.6. The information about Managers on

employees, managers of particular jobs and of different branches can be taken easily (Gupta & Mittal, 2017, p. 57).

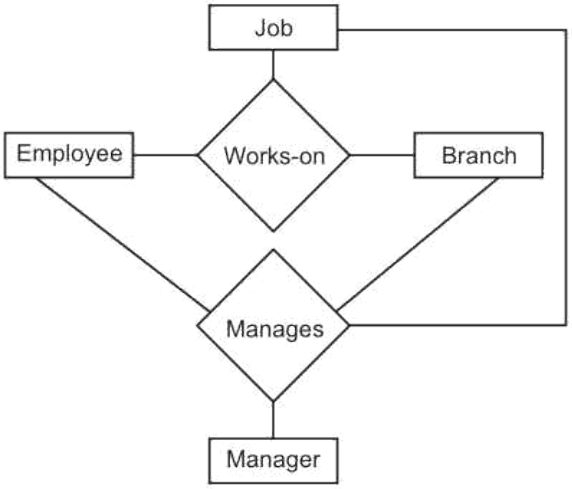


Figure 2.6 ER Model

Designing an EER Model

According to Lemahieu et al., (2018), an EER conceptual data model can be designed according to the following steps:

1. Determine the entity varieties.

2. Determine the relationship types and assert their degree.

3. Assert the cardinality ratios and participation constraints (total versus partial participation).

4. Determine the attribute varieties and assert whether or not they are straightforward or composite, single- or ambiguous, derived or not.

5. Link every attribute sort to entity set type or a relationship type.

6. Denote the key attribute sort(s) of every entity type.

7. Determine the weak entity varieties and their partial keys.

8. Apply abstractions admire generalization/specialization, categorization, and aggregation.

9. Assert the characteristics of every abstraction admire disjoint or overlapping, total or partial.

Any semantics that can't be depicted within the EER model should be documented as separate business rules and followed up using application code. though the EER model offers some new attention-grabbing modeling ideas such as specialization/generalization, categorization, and aggregation, the restrictions of the ER model sadly still apply. Hence, temporal constraints still cannot be sculpturesque, the consistency among multiple relationship varieties cannot be implemented and attribute sort domains or functions cannot be specified (Lemahieu et al., 2018).

Student Activity

Problem 1: Construct an E-R diagram for a hospital with a set of patients and a set of medical doctors. Associate with each patient, a log of various tests and examinations conducted. Construct the appropriate tables for this E-R diagram and list the tables with their attributes, primary key and foreign keys.

Solution. The E-R diagram is shown in Figure 2.7

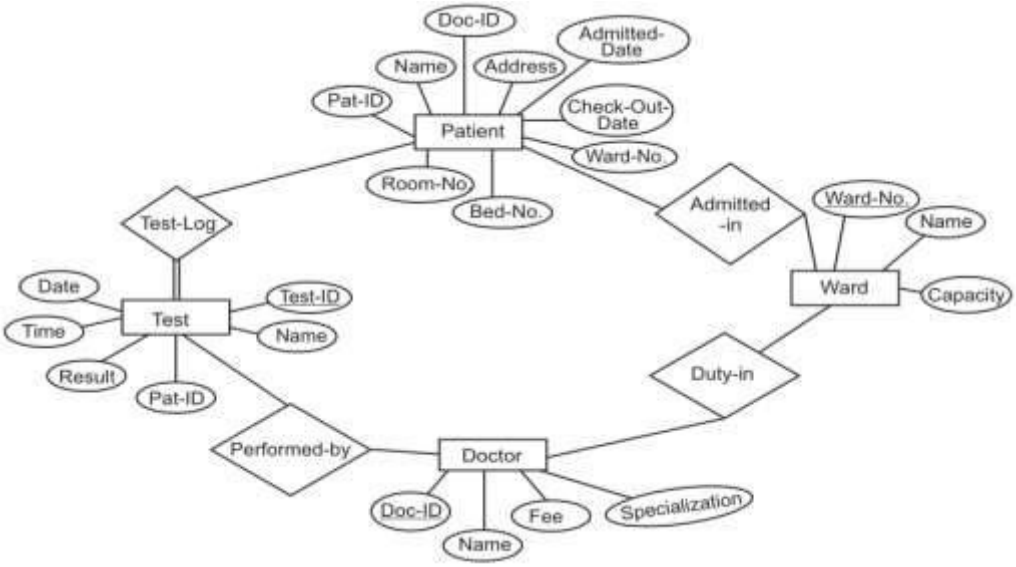


Figure 2.7 ER Diagram of Hospital

The Tables are as follows:

Patient (Pat-ID, name, address, admitted-date, check-out-date, room-no., bed-no., ward-no, doc-ID)

Ward (Ward-no., name, capacity)

Doctor (Doc-ID, name, fee, specialization)

Test (Test-ID, name, date, time, result, Pat-ID)

Primary key is shown by _____.

Foreign key is shown by _ _ _ _ _ _.

Problem 2. The people's Bank offers five type of accounts: Loan, checking, premium savings, daily interest saving, and money market. It operates a number of branches and a client of bank can have any number of accounts. Accounts can be joint i.e.; more than one client may be able to operate a given accounts. Identify the entries of interest and show their attribute. What relationship exists among these entities? Draw the corresponding E-R diagram.

Solution. The E-R diagram is shown in Figure 2.8.

The Entities are as follows:

Account (Account-no, Account-type, Balance)

Branch (Branch-ID, Branch-address, Branch-name)

Transaction (Transaction-ID, Amount, Date, Type)

Client (Account-No., Name, Age, Address)

Relationships are Account-Branch, Account-Transaction, Account-Client.

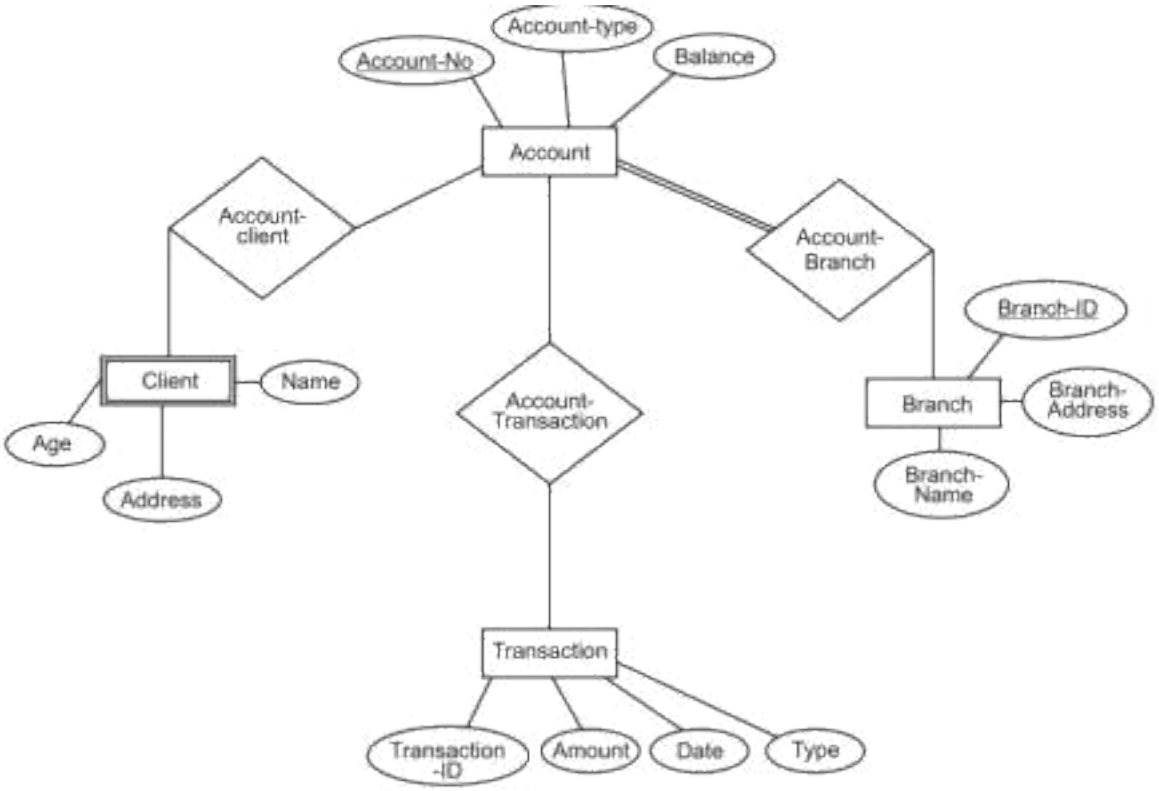


Figure 2.8 ER Diagram of a Bank

Student exercise to strengthen what you have learn.

True or False

1. When transforming an E-R model into a relational database design, each entity is represented as a table.

2. The E-R model refers to a specific table row as an entity occurrence.

3. Attributes are types of entities.

4. A composite key is a primary key composed of more than one attribute.

5. All attributes are either simple or composite.

6. All simple attributes are also single-valued.

7. Composite attributes cannot be further subdivided.

8. A multivalued attribute can have lower and upper bounds.

9. An attribute value can be derived from another attribute.

10. The names on entity types and entity sets are different.

11. An entity cannot have more than one key attribute.

12. A relationship type of degree two is called as ternary relationship.

13. Relationship types can also have attributes.

14. The attribute of a relationship type can be added to participating entity types.

15. A weak entity type can have more than one identifying entity type.

Answer Key (Inverted):

eurT .51	eslaF .01	eurT .5
eurT .41	eurT .9	eurT .4
eurT .31	eurT .8	eslaF . 3
eslaF .21	eslaF .7	eurT .2
eslaF .11	eslaF .6	eurT .1

Assessment Task

Fill in the Blanks

1. Attributes that are not divisible are called _____.

2. When the value of an attribute A is obtained from the value of an attribute B, then the attribute A is called _____.

3. _____ are characteristics of entities.

4. _____ specifies the set of values that can be assigned to the attribute.

5. The partial key attribute is underlined with a _____ line.

6. A person’s social security number would be an example of a(n) _____ attribute.

7. _____ attributes can be subdivided.

8. A(n) _____ attribute cannot be subdivided.

9. An attribute representing one or more bank accounts belonging to a person would be a(n)

_____ attribute.

10. It is better to store the date of birth and use the difference between that value and the system date as a(n) _____ attribute, rather than storing a person’s age.

Multiple Choice

1. A person’s name, birthday, and social security number are all examples of

(a) Entities (b) Attributes (c) Relationships (d) Descriptions

2. An attribute that can be broken down into smaller parts is called a(n) ____ attribute.

(a) simple (b) associative (c) complex (d) composite

3. Which of the following criteria should be considered when selecting an identifier?

(a) Choose an identifier that will not be null.

(b) Choose an identifier that doesn’t have large composite attributes.

(c) Choose an identifier that is stable.

(d) All of the above.

4. A relationship where the minimum and maximum cardinality are both one is a(n) _____

relationship.

(a) optional (b) mandatory link (c) unidirectional (d) mandatory one

5. Which statement is false?

(a) Each attribute of a relation has a name.

(b) Attribute values are (normally) required to be atomic.

(c) The special value null is a member of every domain.

(d) None of above

6. Customers, cars, and parts are examples of

(a) entities (b) attributes (c) relationships (d)
cardinals 7. Which is false?

(a) Relationship type is grouped by the same type of relationship instances.

(b) The current state of a relationship type is the relationship set.

(c) Relationship type identifies the relationship name and the participating entity types.

(d) Relationship type identifies certain relationship constraints.

8. Which is false?

(a) A relationship can have one or more attributes.

(b) A weak entity must participate in an identifying relationship type with an owner or identifying entity type.

(c) Cardinality ratio specifies maximum participation.

(d) An attribute relates two entities.

9. A student can attend 5 courses. Different professors can offer the same courses. The relationship of students to professors is a _____ relationship.

(a) many-to-many (b) one-to-many (c) one-to-one (d) many-to-one

10. In an E-R, Y is the dominant entity and X is a subordinate entity. Then which of the following is incorrect:

(a) Operationally, if Y is deleted, so is X

(b) Existence is dependent on Y

(c) Operationally, if X is deleted, so is Y

(d) Operationally, if X is deleted, and remains the same.

References:

Codd’s 12 Rules - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/dbms/dbms_codds_rules.htm

DBMS - Data Independence - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_independence.htm

DBMS - Data Schemas - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_schemas.htm

ER Model - Basic Concepts - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/er_model_basic_concepts.htm

Gupta, S. B., & Mittal, A. (2017). Introdution to Database Management System (2nd ed.). University Science Press. https://doi.org/10.1201/9781315372679-7

Lemahieu, W., Broucke, S. vanden, & Baesens, B. (2018). Principles of Database Management. Cambridge University Press. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

MySQL - Administration - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/MySQL/MySQL-administration.htm

MySQL - Data Types - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/MySQL/MySQL-data-types.htm

Relation Data Model - Tutorialspoint. (n.d.). Retrieved August 25, 2020, from https://www.tutorialspoint.com/dbms/relational_data_model.htm

Williams, H. E., & Tahaghoghi, S. M. M. (2006). Learning MySQL (A. Oram (Ed.); 1st ed.).

O’Reilly Media.

MODULE 3
FILE ORGANIZATION AND DATA MODELS

Introduction

A report business enterprise is a manner of arranging the facts in a report whilst the report is saved on secondary storage (disk, tape etc.). The one of a kind method of arranging the facts permit one of a kind operations to be done correctly over the report. A database control machine helps numerous report business enterprise techniques. The maximum essential project of DBA is to pick a nice business enterprise for every report, primarily based totally on its use. The business enterprise of facts in a report is prompted with the aid of using wide variety of things that should be considered whilst selecting a selected technique. These elements are (a) speedy retrieval, updation and switch of facts, (b) green use of disk space, (c) excessive throughput, (d) kind of use, (e) green manipulation, (f) safety from unauthorized access, (g) scalability, (h) discount in cost, (i) safety from failure.

Learning Outcomes

At the end of this module, students should be able to:

1. Identify the basic concepts of organizing files.
2. Differentiate DBMS and Relational DBMS.
3. Examine the different data models.

Lesson 1. Basic concept of File Organization

A file may be an assortment of connected sequence of records. a set of field names and their corresponding information sorts constitutes a record. a knowledge sort, related to every field, specifies the kinds of values a field will take. All records during a file are of a similar record sort (Gupta & Mittal, 2017, p. 92).

Data is usually kept within the sort of records. A record may be an assortment of fields or information things and data items is made of 1 or additional bytes. every record

encompasses a distinctive symbol referred to as record-id (Gupta & Mittal, 2017, p. 92). The records during a file are one in every of the subsequent 2 types:

Fixed length records.

Variable length records.

Fixed Length Records

Every record in the file has exactly the same size (in bytes). The record slots are uniform and are arranged in a continuous manner in the file. A record is identified using both record-id and slot number of the record. The Figure 3.1 shows a structure of fixed length STUDENT record and the Table 1 shows a portion of a file of fixed length records (Gupta & Mittal, 2017, p. 92).

```
type STUDENT = record
        NAME = char(20);
        Roll No = char(5);
    DOB = char(8);
    end
```

Figure 3.1 Structure of fixed length record Student

Table 1 Portion of a file of fixed length records.

Name	Roll No.	DOB
Naresh	3234	28-02-75
Suresh	5132	20-05-80
Ramesh	3535	24-10-77
Ashish	3987	15-09-72
Manish	4321	18-11-70
Harish	4983	09-06-73
Manoj	3590	05-01-81

Advantage of Fixed Length Records: The following are the advantages of fixed length records (Gupta & Mittal, 2017, p. 93).

1. Insertion and deletion of records in the file are simple to implement since the space made available by a deleted record is same as needed to insert a new record.

Disadvantage of Fixed Length Records: The major disadvantages of fixed length records are as follows (Gupta & Mittal, 2017, p. 93):

1. In fixed length records, since the length of record is fixed, it causes wastage of memory space. For example, if the length is set up to 50 characters and most of the records are less than 25 characters, it causes wastage of precious memory space.
2. It is an inflexible approach. For example, if it is required to increase the length of a record, then major changes in program and database are needed.

Variable Length Records

Every record within the file needn't be of identical size (in bytes). Therefore, the records within the file have completely different sizes. the foremost drawback with variable length record is that once a brand-new record is to be inserted, associate degree empty slot of the precise length is needed. If the slot is smaller, it can't be used and if it's too huge, the additional house is simply wasted. There are 2 ways to store variable length records with a hard and fast length illustration (Gupta & Mittal, 2017, p. 93). These are Reserved house and pointers. A file might have variable length records because of following reasons:

1. One or quite one fields of a record are of variable size however the records within the file are of same record sort.

2. One or quite one fields might have multiple values for individual records and are known as repetition fields however the records within the file are of same record sort.

3. One or quite one fields are nonmandatory i.e., they will have values for a few however not for all records. The file records during this case also are of identical record sort.

4. The file contains records of various record sorts and different sizes.

Advantage of Variable Length Records: the benefits of variable length records are as follows (Gupta & Mittal, 2017, p. 94):

1. It reduces manual mistakes as information mechanically change the scale of record.

2. It saves ton of memory house just in case of records of variable lengths.

3. it's a versatile approach since future enhancements are terribly simple to implement.

Disadvantage of Variable Length Records: The disadvantages of variable length records are as follows (Gupta & Mittal, 2017, p. 94):

1. It will increase the overhead of software as a result of information ought to keep record of the sizes of all records.

The following 3 styles of files are employed in database systems:

1. Master file: This file contains data of permanent nature regarding the entities. The computer file act as a supply of reference knowledge for process transactions. They accumulate the knowledge supported the group action data.

2. Transaction file: This file contains records that describe the activities allotted by the organization. This file is made as a results of process group actions and making ready transaction documents. These also are accustomed update the computer file for good.

3. Report file: This file is made by extracting knowledge from the various records to arrange a report e.g. A report file regarding the weekly sales of a selected item.

Lesson 2. File Organization Techniques

A file organization could be a means of composition the records in a very file once the file is kept on auxiliary storage (disk, tape etc). There are differing kinds of file organizations that are employed by applications. The operations to be performed and therefore the choice of memory device are the key factors that influence the selection of a specific file organization (Gupta & Mittal, 2017, p. 94). The various sorts of file organizations are as follows:

1. **Heap file organization:** during this file organization, the records are hold on within the file, within the order during which they're inserted. All the new records are hold on at the tip of the file. This file organization is additionally referred to as PILE FILE. This organization is mostly used with further access methods, like secondary indexes. Inserting a brand-new record is extremely quick and economical. however, looking out a record mistreatment any

search condition involves a linear search through the file, that is relatively longer intense. Deleting a record from a heap file isn't economical as deletion

of records resulted in wastage of cupboard space. it's usually wont to store tiny files or in cases wherever knowledge is tough to prepare. it's conjointly used once knowledge is collected at one place before process (Gupta & Mittal, 2017).

Advantages of Heap File Organization: the key blessings of Heap file organization are as follows:

1. Insertion of recent record is quick and economical.
2. The filling issue of this file organization is one hundred pc.
3. house is totally used and preserved.

Disadvantage of Heap File Organization: the key disadvantages of Heap file organization are as follows:

1. looking out and accessing of records is extremely slow.
2. Deletion of the many records lead to wastage of house.
3. Updation price of knowledge is relatively high.
4. it's restricted applications.

2. **Sequential file organization:** In ordered file organization, records are kept in a very ordered order in line with the “search key”. a hunt key's associate degree attribute or a group of attributes that are accustomed set up the records. it's not necessary that search key should be primary key. It is the only methodology of file organization. ordered methodology is predicated on tape model. Devices who support ordered access are magnetic tapes, cassettes, card readers etc. Editors and compilers conjointly use this approach to access files (Gupta & Mittal, 2017).

   Structure of an ordered file is shown in Figure 3.3. The records are kept in ordered order one when another. to succeed in at the consecutive record from any record pointers are used. The Pointers are used for quick retrieval of records (Gupta & Mittal, 2017).

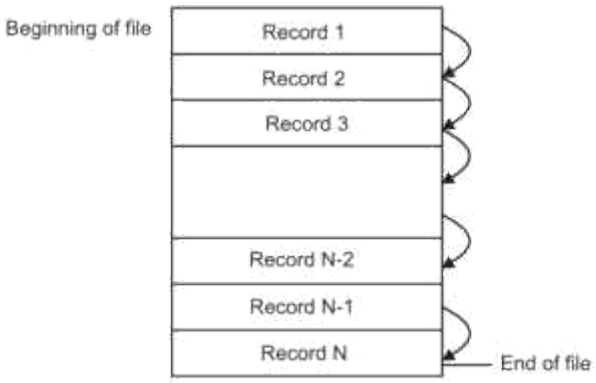


Figure 3.3 Sequential file organization.

File Operations on Sequentially Organized Files

Various operations which will be performed on successive files are as follows:

a. Creating a sequential file: to form a replacement file, initial free house is searched in memory. Once looking, enough house for file, assign that house to the new file. Name and physical location of recent file is entered in filing system directory (Gupta & Mittal, 2017).

b. Open an existing file: to open any file, enter the name of file. This name is searched in filing system directory. Once matching the name, records within the file are transferred from secondary memory to primary memory. Now, the file is prepared to read/write (Gupta & Mittal, 2017).

c. Closing a file: once task over file is completed then shut that file. The memory house allotted for that come in primary memory is deallocated. If file was opened mechanically with any program then it'll be closed automatically once ending that program (Gupta & Mittal, 2017).

d. Reading a sequential file: to browse any record from a sequential file it's necessary to start out from starting of file till record is use or eof is occurred. You cannot go on to any explicit record. Solely linear search is wont to search any record (Gupta & Mittal, 2017).

3. **Indexed—Sequential file organization**: to overcome the disadvantages of consecutive file

organization. It additionally preserves the benefits of consecutive access. This organization permits quick looking out of records with the utilization of index. Some basic terms related to index consecutive file organization are as follows (Gupta & Mittal, 2017):

Block: Block may be a unit of storage within which records are saved.

Index: Index may be a table with a pursuit key by that block of a record may be notice. **Pointer**: Pointer may be a variable that points from index entry to beginning address of block.

Components of an Indexed Sequential File

a. **Prime area**: throughout the creation of index consecutive file, records are written in prime space. Records are maintained in step with any key. Hence, prime space should be a consecutive file (Gupta & Mittal, 2017).

b. **Overflow space**: Overflow area is employed throughout addition of recent records.

There are 2 forms of overflow area (Gupta & Mittal, 2017):

(i) Cylinder overflow space: This area consists of free area on every cylinder that is reserved for overflow records for that individual cylinder.

(ii) freelance overflow space: This area is employed to store overflow records from anyplace.

c. **Record** characteristics: typically, mounted length records are used.

d. **Indexes**: Index may be a assortment of entries and every entry corresponds to dam of storage (Gupta & Mittal, 2017).

(i) 1st level index: the bottom level index is understood as first level index.

(ii) Higher level index: Higher level compartmentalization is employed once 1st level index becomes large.

(iii) Cylinder index: The indexes may be created in step with the hardware boundaries. Cylinder index entries consists one entry for every cylinder.

(iv) **Master index**: the very best level of index is understood as master index.

The advantages of Index consecutive files are as follows:

1. Economical classification system for medium and huge size files.

2. Straightforward to update.
3. Straightforward to keep up than direct files.
4. Economical use of cupboard space.
5. Looking out of records are quick.
6. Maintain benefits of consecutive classification system.

The disadvantages of Index consecutive files are as follows:

1. Inefficient classification system for little size files.
2. It's pricy methodology.
3. Typical structure than consecutive files.
4. Indexes would like extra cupboard space.
5. Performance degradation w.r.t. Growth of files.

4. **Hashing or Direct file organization**.

According to Gupta & Mittal (2017), to meet the necessity to access records willy-nilly direct file organization is employed. In direct file organization records will be hold on anyplace in cargo area however can be accessed directly, with none consecutive looking. It overcomes the drawbacks of consecutive, index consecutive and B-trees file organization.

Direct file organization depends upon hashing that gives the bottom of mapping procedure. To beat the drawbacks of hashing rule, collision resolution technique is required. Devices that support direct access are CD’s, Floppy etc. Direct file organization is additionally referred to as Random File Organization. Selection of hashing rule and collision resolution technique is crucial purpose in direct file organization (Gupta & Mittal, 2017).

File Operations on Direct Files

1. **M**aking an immediate file: once looking and allocating free house for file, necessary entries in system directory are created. In direct file organization, a hashing algorithmic rule for mapping procedure and any collision resolution technique to avoid collisions throughout mapping are such as. The key field is additionally such as (It might not be primary key).

2. Open associate existing file and shutting a file are same as in different file organizations.
3. Looking (Reading or retrieving) from direct file: To browse any record from direct file, simply enter the key field of that record. With the assistance of hashing algorithmic rule that key field is mapped into physical location of that record. Just in case of any collision, collision resolution technique is employed.
4. Updation of records in direct file:
    a. **Adding a new record**: To add a new record in direct file, specify its key field. With the help of mapping procedure and collision resolution technique, get the free address location for that record.
    b. **Deleting record from direct file**: To delete a record, first search that record and after searching, change its status code to deleted or vacant.
    c. **Modify any record**: To modify any record, first search that record, then make the necessary modifications. Then re-write the modified record to the same location.

The advantages of direct files are as follows:

1. Records are not needed to be sorted in order during addition.
2. It gives fastest retrieval of records.
3. It gives efficient use of memory.
4. Operations on direct file are fast so there is no need to collect same type of operations in a file, as in sequential file system.
5. Searching time depends upon mapping procedure not logarithm of the number of search keys as in B-trees.
6. Supports fast storage devices.

The disadvantages of direct files are as follows:

1. Wastage of storage space (Clustering) if hashing algorithm is not chosen properly.
2. It does not support sequential storage devices.
3. Direct file system is complex and hence expensive.
4. Extra overhead due to collision resolution techniques.

Lesson 3. Hierarchical Model

Hierarchical software packages were one amongst the primary DBMS varieties developed, and adopt a tree-like information model. As a tree-structured model it must have

1: N relationship types. Parent records can have multiple child records, but a child record can have only one parent (Lemahieu et al., 2018, p. 1695).

A hierarchical info consists of assortment of records, that are connected to every alternative by links.

Record: A record may be an assortment of attributes; every contains just one information price.

Link: A link is Associate in Nursing association between 2 records.

According to Gupta & Mittal (2017), the tree structure utilized in hierarchical model is thought as nonmoving tree. the foundation node of that tree is dummy node or Associate in Nursing empty node. So, hierarchical model may be an assortment of nonmoving trees. assortment of nonmoving trees build forest. A nonmoving tree is additionally called info tree.

Tree structure consists of 2 basic components:

1. Rectangular boxes: Rectangular boxes represent varied record varieties.
2. Line: Line represents link between 2 record varieties.

A tree structure diagram specifies the general logical structure of info (as E-R diagram in Entity Relationship model).

Relationship exists between a parent and a toddler should be one-to-many or matched relationship.

Link between parent Associate in Nursing kid is drawn by line with an arrow. A parent might have Associate in Nursing arrow inform to kid, however kid should have Associate in Nursing arrow inform to its parent. A general tree structure is shown in Figure 3.4 (Gupta & Mittal, 2017).

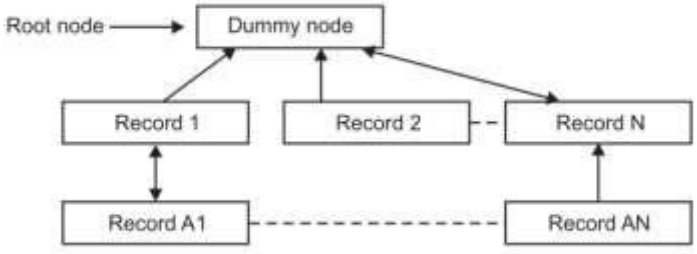


Figure 3.4 General Tree Structure.

Advantages of Hierarchical Model

According to Gupta & Mittal (2017), the following are the main advantages of hierarchical data model:

1. **Simplicity**: during this model, records are connected in style of parent/child relationship. Acting varied operations (i.e., insertion, deletion etc.) During this tree like structure is simple and easy to perform. This leads to the straightforward style of the info ensuing from this model (Gupta & Mittal, 2017).

2. **Integrity of Data**: The parent/child relationship between the varied records within the hierarchical model is drawn by a relationship or link. Every child phase is coupled to just one parent and a toddler will solely be reached through its parent; therefore, this model promotes information integrity (Gupta & Mittal, 2017).

3. **Information Security**: every child phase is coupled to just one parent and a toddler will solely be reached through its parent during this model. So, for deleting the kid phase correct data of parent segment is required. It provides information security that is implemented by the software package (Gupta & Mittal, 2017).

4. **Efficiency**: The hierarchical model contains one to several relationships between parent and kid. Once the info contains several 1: N relationships between varied records then this model handles it terribly expeditiously (Gupta & Mittal, 2017).

5. It's terribly economical to handle sizable amount of transactions exploitation this model. This can be primarily as a result of the links (or relationship) established by the pointer within the varied records are permanent and can't be changed (Gupta & Mittal, 2017).

Disadvantages of Hierarchical Model

The information is replicated in hierarchical database. The replication may occur either in different database trees or in same tree (Gupta & Mittal, 2017, p. 138)

1. Knowledge of physical level of data storage is required: The requirement for a one to many relationships between parent and child can result in redundancy of data. To get around

   the redundancy problems, data is stored in one place and referenced by links or physical pointers, which requires technical skills (Gupta & Mittal, 2017).

2. Complexity: The physical links make it very difficult to expand or modify the database, changes typically require substantial rewriting efforts (Gupta & Mittal, 2017).
3. Inflexibility: The basic problem occurs with this model is that they are not flexible enough to establish all the relationships (many-to-many etc.) Which occur in the real world. Usually there are one to many relationships between the records, established by pointers which are permanent and cannot be modified in case of other cases where relationships (like many to many etc.) Exist (Gupta & Mittal, 2017).
4. Lack of querying facilities: The lack of declarative querying facilities and need for navigation of pointers to access needed information make querying rather complex. It does not provide the query capability easily.
5. Database management problems: In this model, the modifications to the data structure leads to significant modifications to application programs that access the database. Also new relations or nodes result in complex system management tasks (Gupta & Mittal, 2017).
6. Problems with data manipulation operations: Various problems are encountered while performing various operations like insertion, deletion and updation. Moreover, the data retrieval is also very complex and asymmetric. Therefore, a better model is needed to solve these problems.
7. Lack of standards: This model does not have any specific or precise standard for database design and modelling (Gupta & Mittal, 2017).

Lesson 4. Network Model

Network model is based on graph structure. A network database consists of collection of records, which are connected to each other by links (Gupta & Mittal, 2017, p. 139).

Record: A record is a collection of attributes; each contain only one data value.

Link: A link is an association between two records.

So, network model is a collection of graphs.

Graph structure consists of two basic components.

Rectangular boxes: Rectangular boxes represent various record types.

Line: Line represents link between two records. A link cannot contain data value.

A graph structure diagram specifies the overall logical structure of database.

The basic operations that can be performed on network data model are insertion, deletion, updation and retrieval.

1. **Insertion**: The insert operation is used to insert a new record into the database.
2. **Deletion**: The deletion operation is used to delete a record from the database.
3. **Updation**: Since the record appears only once in the network model so changes made to any attributes or columns of the record can be performed easily without any inconsistency.
4. **Retrieval**: The process of searching and fetching of a record in the database is known as retrieval of a record. Unlike the hierarchical model, the retrieval operation in network data model is symmetric but complex.

Advantages of Network Model

According to Gupta & Mittal, 2017, pp. 144-145, the network model is better than the hierarchical model and eliminates some of the limitations of hierarchical model. The following are major advantages of network model:

1. Eliminate information Redundancy: In network model, we've got just one prevalence for a selected record within the info which may confer with different records exploitation links or pointers. Since there are not any multiple occurrences of records therefore it eliminates information redundancy (Gupta & Mittal, 2017).

2. Lesser Storage Requirements: Since during this model, a record happens just one occasion while not repetition therefore lesser storage necessities are there for storing the records within the info (Gupta & Mittal, 2017).

3. Higher Performance: during this model the relationships are outlined directly that ends up in better performance (Gupta & Mittal, 2017).

4. Handle sorts of Relationships: As we all know that an outsized variety of various type or relationships like one to 1 (1: 1), one to several (1: N) and plenty of to many (N: N) etc., exist within the planet things. These relationships will simply outline within the info definition exploitation network model (Gupta & Mittal, 2017).

5. Easy accessibility of Data: Since records within the network model are coupled along through use of pointers therefore it's terribly straightforward to maneuver from one owner

record to a different. Also, associate degree application will access an owner record and every one the member records within the set (Gupta & Mittal, 2017).

6. Promotes information Integrity: The network model promotes data integrity owing to the specified owner-member relationships i.e., user should initial style the owner record and so the member record (Gupta & Mittal, 2017).

7. Enforce Standards: The network model was standardized as codasyl DBTG Model therefore all the network direction systems show these standards. The standards embrace information Definition Language (DDL) and information Manipulation Language (DML) that improves the info administration and movability (Gupta & Mittal, 2017).

8. Information Independence: In network model the structure of the information is often modified while not modifying application programs that ends up in independence of information however this data independence is partly not totally in contrast to the class-conscious model wherever data independence is incredibly low (Gupta & Mittal, 2017).

Disadvantages of Network Model

According to Gupta & Mittal, 2017, pp. 145, the drawbacks or disadvantages of the Network model are as follows:

1. **Complexity**: though the abstract planning of the network model is straightforward however the look at the hardware level is incredibly advanced as a result of an outsized variety of pointers are needed to indicate the link between the owner records and therefore the member records. This makes this model obsolete for all the sensible functions (Gupta & Mittal, 2017).

2. **Problem in Querying Data**: For querying information in network model the technologist is forced to suppose in terms of links and the way to traverse them to induce the required data. Therefore, correct technical skills are needed on a part of a technologist (Gupta & Mittal, 2017).

3. **Lack of Structural Independence**: though the network model achieves information independence however it fails to realize structural independence. Since the assorted records are coupled through pointers, that forms a steering chain. Therefore, creating structural changes to the databases is incredibly troublesome. If changes are to be

created to the information base structure then all the applying programs exploitation it conjointly got to be changed before accessing the data (Gupta & Mittal, 2017).

Lesson 5. Relational Model

As explained by Gupta & Mittal, 2017, pp. 146, the relative model was discovered by Dr. E. F. Codd. The relative model is AN abstract theory of information that's supported the mathematical thought of relations. In computer database, data is kept in tabular type or in tables. So, it consists of assortment of tables. We tend to use mathematical terms in relative model. The relative model cares with 3 aspects of data: arrangement, information integrity, and information manipulation.

Relational model uses easy tables rather than advanced tree and network structures to alter the user's read of the information. It's a group of tables within which information is keep. A table may be a matrix of a series of row and column intersections. Following are the essential information structures utilized in computer database (Gupta & Mittal, 2017).

1. Relation (Entity set in E-R model): In relative model name of table is thought as relation. It consists of all potential tuples.

2. Tuple (Entity in E-R model): Single row of any relation is thought as tuple.

3. Attributes (Same in E-R model): These are the characteristics of any relation.

4. Domain (Same in E-R model): Domain is that the set of all permissible values or data regarding any attributes. Maybe in relation worker, domain for attribute Emp- Name is (Deepak, Vinay, Gaurav, Rajiv).

5. Tuple variable (Value in E-R model): Tuple variable is that the a part of tuple or row, that is data or information keep in relation for any attribute

6. Degree: variety of columns in any relation is thought as its degree.

7. Cardinality: variety of rows in any relation is thought as its cardinality.

Attributes sort mentioned in E-R model also are applicable here.

1. Keys: A secret's one attribute or a collection of attributes of a relation which might unambiguously establish a record among the relation. A key is also Primary key, Supper

key, Candidate key, Foreign key etc. Keys mentioned in E-R model also are applicable here.

2. Database instance: information instance shows the info of database at a specific instance of your time.

3. Relation schema: Relation schema is that the logical style of relations of any information. It offers the approach of illustration of relations otherwise you will say that notations used for relations in information.

4. Relation instance: it's connected with the values in relations. It shows the info keep in relations at a specific instance of your time.

Constraints - each relation has some conditions that has to hold for it to be a legitimate relation. These conditions are referred to as relative Integrity Constraints. There are 3 main integrity constraints (Relation Data Model - Tutorialspoint, n.d.):

1. Key constraints - There should be a minimum of one smallest set of attributes within the relation, which might establish a tuple unambiguously. This smallest set of attributes is named key for that relation. If there are quite one such smallest subsets, these are referred to as candidate keys (Relation Data Model - Tutorialspoint, n.d.).

Key constraints force that:

A. In a relation with a key attribute, no 2 tuples will have identical values for key attributes.

B. A key attribute cannot have NULL values.

Key constraints also are observed as Entity Constraints.

2. Domain constraints - Attributes have specific values in real-world state of affairs. Maybe, age will solely be a positive whole number. An equivalent constraint is tried to use on the attributes of a relation. Each attribute is absolute to have a particular vary of values. Maybe, age can't be below zero and phone numbers cannot contain a digit outside 0-9 (Relation Data Model - Tutorialspoint, n.d.).

3. Referential integrity constraints - denotative integrity constraints work on the thought of Foreign Keys. An overseas secret's a key attribute of a relation which will be referred in

different relation. Referential integrity constraint states that if a relation refers to a key attribute of a distinct or same relation, then that key component should exist (Relation Data Model - Tutorialspoint, n.d.).

Lesson 6. CODD’s Rules

Based from Codd’s 12 Rules - Tutorialspoint, n.d., Dr Edgar F. Codd, after his extensive research on the Relational Model of database systems, came up with twelve rules of his own, which according to him, a database must obey in order to be regarded as a true relational database. These rules can be applied on any database system that manages stored data using only its relational capabilities. This is a foundation rule, which acts as a base for all the other rules.

1. Information Rule - the information keep in a very information, might or not it's user information or information, should be a price of some table cell. Everything in a very information should be keep in a table format.

2. Guaranteed Access Rule - each single information part (value) is certain to be accessible logically with a mixture of table-name, primary-key (row value), and attribute-name (column value). No different means that, resembling pointers, are often wont to access information.

3. Systematic Treatment of NULL Values - The NULL values in a very information should run a scientific and uniform treatment. This are often a really vital rule as a result of a NULL can be understood together the following: information is missing, information isn't proverbial, or information isn't applicable.

4. Active online Catalog - The structure description of the whole information should be kept in an internet catalog, referred to as information lexicon, which might be accessed by approved users. Users will use an equivalent search language to access the catalog that they use to access the information itself.

5. Comprehensive data Sub-Language Rule - A information will solely be accessed employing a language having linear syntax that supports data definition, information manipulation, and dealings management operations. This language are often used directly or by means that of some application. If

the information permits access to information with none facilitate of this language, then it's thought of as a violation.

6. View updating Rule - All the views of a information, which might in theory be updated, should even be updatable by the system.

7. High-Level Insert, Update, and Delete Rule - A information should support high-level insertion, updation, and deletion. This should not be restricted to one row, that is, it should conjointly support union, intersection and minus operations to yield sets of information records.

8. Physical data Independence - the information keep in a very information should be freelance of the applications that access the database. Any modification within the natural object of a information should not have any impact on however the information is being accessed by external applications.

9. Logical data Independence - The logical data in a very information should be freelance of its user's read (application). Any modification in logical information should not have an effect on the applications exploitation it. As an instance, if 2 tables are incorporated or one is split into two totally different tables, there ought to be no impact or modification on the user application. This can be one among the foremost tough rule to use.

10. Integrity Independence - A information should be freelance of the appliance that uses it. All its integrity constraints are often severally changed while not the requirement of any modification within the application. This rule makes a information freelance of the front-end application and its interface.

11. Distribution Independence - The end-user should not be able to see that the information is distributed over varied locations. Users must always get the impression that the information is found at one website solely. This rule has been considered the inspiration of distributed information systems.

12. Non-Subversion Rule - If a system has an interface that has access to low-level records, then the interface should not be able to subvert the system and bypass security and integrity constraints.

Lesson 7. Comparison of DBMS and RDBMS

Table 2. Comparison of DBMS and RDBMS

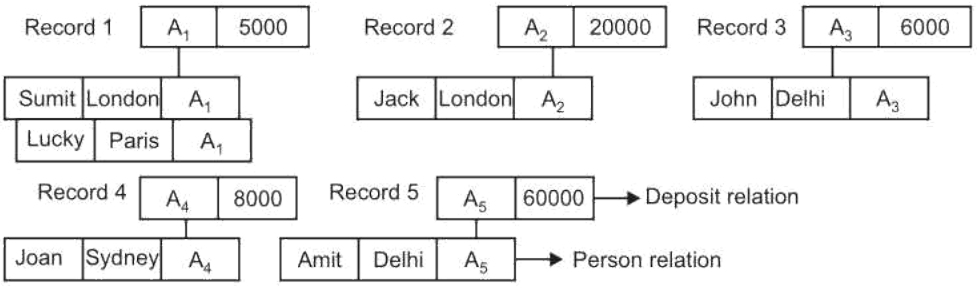
DBMS	RDBMS
DBMS is a generalized software for managing and manipulating the databases.	RDBMS is a type of DBMS that depends upon the mathematical concepts of relation.
DBMS's organize data by using data files with records and fields.	RDBMS's organize data by using tables, tuples and attributes.
Several files cannot be stored in a single file.	Several tables can be stored in a single table known as table pools.
Navigation is not so simple	Navigation is much simpler.
It is not the case here.	It creates new database tables by using basic operators.
Physical and logical data independence depends upon the structure of database.	It provides physical and logical data independence by using mappings.
Data handling, transaction processing and other features are less powerful than RDBMS.	RDBMS have more powerful data handling, transaction processing and other features to enhance speed, security and reliability

The following exercises is intended to deepen the understanding of the subject.

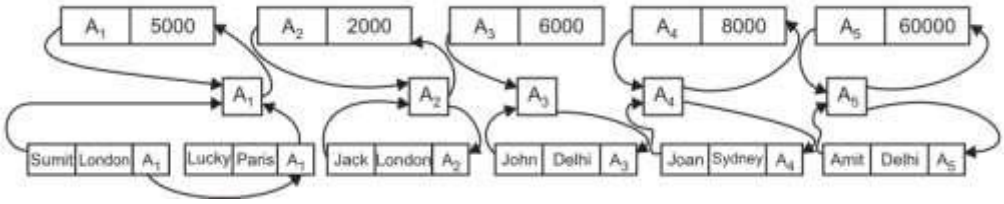
Problem: Consider the following relations given as relationship model-person (Name,City, AccNo), deposit (AccNo, Balance).

Considering at least five dummy records the database given above, show how this relational model can be converted into network model and hierarchical model.

Solution : We consider that AccNo is unique in relation deposit and combination of Name and AccNo is unique in relation person and account may be a combined account. Sample data in hierarchical model



Sample data in network model.



True or False

1. A data model is usually graphical.
2. A data model represents data structures for a specific problem domain.
3. Each row in the relational table is known as an entity instance in the ER model.
4. In relational model the M : N relationships are not appropriate.
5. The network model has structural level dependency.
6. The ER model is limited to conceptual modelling, with no implementation component.
7. The hierarchical model is software-independent.
8. The relational model is hardware-dependent and software-independent.
9. All values in a column are of same data type.
10. Each row in a table represents a collection of different data values.
11. In formal relational model terminology, a table is called a relation.
12. Tuples in a relation must have a particular order.
13. Ordering of values in a tuple is important.
14. Composite or multivalued attributes are allowed in relational model.
15. A relation schema can have more than one key.

Answer Key (Inverted):

eurT .51	eslaF .01	eurT .5
eslaF .41	eurT .9	eurT .4
eurT .31	eslaF .8	eurT .3
eslaF .21	eslaF .7	eurT .2
eurT .11	eurT .6	eurT .1

Assessment Task

Fill in the Blanks

1. _____data structure is used in the network model.

2. ___ ___ data structure is used in the hierarchical model.

3. A simple representation of complex real-world data structures is called_____.

4. In a hierarchical model_____ is equivalent of a record in a file system.

5. The _____ is the conceptual organization of the entire database as viewed by the database administrator.

6. The portion of the database as seen by the application programs that produce information from the data is called_____.

7. Each row in a relation is called a(n) _____.

8. Each column in a relation represents a(n) _____.

9. The task of creating a conceptual data model that could be implemented in any DBMS is called_____.

10. When you can change the internal model without affecting the conceptual model, it is said to have _____ independence.

11. The number of rows in a table is called its _____.

12. The number of columns in a table is called its _____.

13. In the Network model _____ is used to represent many to many relationships.

14. A(n) _____ is a named column of a relation.

15. A(n) _____ structure is a logical data model that arranges data according to some natural hierarchy on a one-to-one basis.

Multiple Choice

1. The columns of a relation (a) must be in specified order

(c) with key field in first column

(b) may be in any order

(d) with largest width column last

2. Which of the following constraint can be expressed in schema of relational model by using DDL?

(a) Schema-based

(c) Application-based

(b) Inherent model-based

(d) System-based

3. Which of the following constraint specifies that no two distinct tuples in any state of

relational schema can have same values for super keys?

(a) Entity-integrity

(c) Referential-integrity

(b) Domain

(d) Key

4. Which of the following constraint specifies that within each tuple, the value of each attribute must be atomic value from some domain?

(a) Entity-integrity

(c) Referential-integrity

(b) Domain

(d) Key

5. Which of the following key is used to identify tuples in a relation?

(a) Secondary

(c) Main

(b) Primary

(d) Number

6. When there is more than one key in a relation, then each such key is called _____.

(a) primary

(c) multiple

(b) useful

(d) candidate

7. Which of the following constraint states that no primary key value can be null?

(a) Key

(c) Referential-integrity

(b) Domain

(d) Entity-integrity

8. Which of the following constraint states that a tuple in one relation that refers to another relation must refer to an existing tuple in that relation?

(a) Key

(c) Referential-integrity

(b) Domain

(d) Entity-integrity

9. Which of the following constraint is used to maintain consistency among tuples in two relations?

(a) Key

(c) Referential-integrity

(b) Domain

(d) Entity-integrity

10. Pick the odd one out (a) Primary key

(c) Candidate key

(b) Super key

(d) Foreign key

References:

Codd’s 12 Rules - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/dbms/dbms_codds_rules.htm

DBMS - Data Independence - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_independence.htm

DBMS - Data Schemas - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_schemas.htm

ER Model - Basic Concepts - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/er_model_basic_concepts.htm

Gupta, S. B., & Mittal, A. (2017). Introdution to Database Management System (2nd ed.). University Science Press. https://doi.org/10.1201/9781315372679-7

Learning MySQL. (n.d.). Creative Commons BY-SA. https://riptutorial.com/ebook/MySQL

Lemahieu, W., Broucke, S. vanden, & Baesens, B. (2018). Principles of Database Management. Cambridge University Press. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

Tutorialspoint. (2014). MySQL Database Management.

https://doi.org/10.1017/CBO9781107415324.004

TutorialsPoint. (2015). Database Management System. Database Management System. https://doi.org/10.1017/CBO9781107415324.004

MODULE 4
INTRODUCTION TO MYSQL

Introduction

MySQL operates using client/server design during which the server runs on the machine containing the databases and shoppers connect with the server over a network. The server software systems are sometimes an operative system} or Windows 2000 operating system. Typically, MySQL is supported on Windows XP, Windows Server 2003, Red Hat lid UNIX operating system, and Debian UNIX operating system, and others. like the other client/server application, MySQL could be a multi-user info system, that means many users will access the info at the same time.

Learning Outcomes

At the end of this module, students should be able to:

1. Use MySQL administrative commands.
2. Use MySQL Monitor
3. Identify the different data types in MySQL

Lesson 1. Features of MySQL

MySQL could be a quick, easy-to-use RDBMS getting used for several little and massive businesses. MySQL is developed, marketed, and supported by MySQL AB, which could be a Swedish company. MySQL is turning into thus widespread owing to several sensible reasons:

• MySQL is free below the with ASCII text file license.

• MySQL could be a powerful program in its title. It handles an outsized set of the practicality of the foremost expensive and powerful info packages.

• MySQL uses a customary style of the well-known SQL information language.

• MySQL works on several operative systems and with several languages together with PHP, PERL, C, C++, JAVA, etc.

• MySQL works quickly and works well even with giant information sets.

• MySQL is extremely friendly to PHP, the foremost appreciated language for net development.

• MySQL supports giant databases, up to fifty million rows, or a lot of in an exceeding table. The default file size limit for a table is 4GB, however, you'll be able to increase this (if your package can handle it) to a theoretical limit of eight million terabytes (TB).

• MySQL is customizable. The ASCII text file GPL license permits programmers to change the MySQL code to suit their specific environments.

Lesson 2. Installing MySQL

All downloads for MySQL are located at MySQL Downloads. Pick the version number of the MySQL Community Server which is required along with the platform you will be running it on.

MySQL Download Link: https://dev.MySQL.com/downloads/installer/

The default installation on any version of Windows is now much easier than it used to be, as MySQL now comes neatly packaged with an installer. Simply download the installer package, unzip it anywhere and run the setup.exe file.

The default installer setup.exe will walk you through the trivial process and by default will install everything under C:\Program Files\MySQL.

Test the server by firing it up from the command prompt the first time. Go to the location of the MySQLd server which is probably C:\MySQL\bin, and type:

MySQLd.exe --console

For a more detailed installation guide, please visit: https://bit.ly/cardinalyt

Lesson 3. Introduction to MySQL Monitor

The monitoring program is named simply MySQL and is found in a very directory with the other MySQL programs. the precise location depends on your operating system and the way you selected to install MySQL (Williams & Tahaghoghi, 2006, p. 95);

Follow these steps to start out the monitor and hook up with your MySQL server because the MySQL administrator (the MySQL root user) by typewriting this from the command line:

MySQL --user=root

If your server has a password, you should specify the password as follows:

MySQL --user=root --password=the_MySQL_root_password

If all goes well, you'll get the monitor's MySQL> prompt

```
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 456 to server version: 5.0.22

Type 'help;' or '\h' for help. Type '\c' to clear the buffer.

mysql>
```

When interacting with a MySQL server, you'll use a mixture of SQL keywords, MySQL proprietary commands, and names of databases and database parts. We tend to follow common conventions and use a mode to create it easier to tell apart between parts of an SQL query. MySQL isn't fussy regarding whether or not you enter SQL or the monitor's proprietary statements in capital or lowercase. For instance, SELECT, select, Select, and even select are equivalent (Williams & Tahaghoghi, 2006).

However, depending on your platform, MySQL will be fussy regarding database and table names. for instance, underneath Windows, MySQL isn't fussy in the least (because Windows itself isn't fussy regarding the filenames that store those structures), whereas on Macintosh OS X its fussiness depends on what underlying filesystem you utilize to store disk files. Linux and OS systems observe the distinction between capital and minuscule strictly. A reliable approach is to adopt the convention of victimization minuscule for all info, table, and column names (Williams & Tahaghoghi, 2006).

There are some restrictions on what characters and words you'll be able to use in your info, table, and alternative names. for instance, you can't have a column named from or choose These restrictions are largely obvious since they apply to reserved keywords that confuse MySQL's parser (Williams & Tahaghoghi, 2006).

You'll be aware that we generally tend to terminate all SQL statements with the semicolon character (;). This tells MySQL that we've completed entering an assertion which it ought to presently parse and execute it. This gives you flexibility, allowing you to kind in an assertion over many lines (Williams & Tahaghoghi, 2006).

Indifference to SQL statements, you can't span the MySQL monitor's very own instructions over pretty one line. that is frequently due to the punctuation isn't truly wished for these, and absolutely urgent the Enter key has a steady result. For instance, the employment command tells MySQL which you desire to apply a selected database (Williams & Tahaghoghi, 2006). The following works fine

```
mysql> USE test
Database changed
```

However, if you try to span the command over more than one line, you won't get far:

```
mysql> USE
ERROR:
USE must be followed by a database name
```

The monitor has a handy HELP command that you can use to get more information on the monitor commands or SQL syntax. If you type HELP and press the Enter key, you'll get a list of commands the monitor understands (Williams & Tahaghoghi, 2006) :

```
MySQL> help

For information about MySQL products and services, visit:
   http://www.MySQL.com/
For developer information, including the MySQL Reference Manual, visit:
   http://dev.MySQL.com/
To buy MySQL Enterprise support, training, or other products, visit:
```

```
List of all MySQL commands:
Note that all text commands must be first on line and end with ';'
?         (\?) Synonym for `help'.
clear     (\c) Clear the current input statement.
connect   (\r) Reconnect to the server. Optional arguments are db and host.
delimiter (\d) Set statement delimiter.
ego       (\G) Send command to MySQL server, display result vertically.
exit      (\q) Exit MySQL. Same as quit.
go        (\g) Send command to MySQL server.
help      (\h) Display this help.
notee     (\t) Don't write into outfile.
print     (\p) Print current command.
prompt    (\R) Change your MySQL prompt.
quit      (\q) Quit MySQL.
rehash    (\#) Rebuild completion hash.
source    (\.) Execute an SQL script file. Takes a file name as an argument.
status    (\s) Get status information from the server.
tee       (\T) Set outfile [to_outfile]. Append everything into given outfile.
use       (\u) Use another database. Takes database name as argument.
charset   (\C) Switch to another charset. Might be needed for processing binlog
               with multi-
byte charsets.
warning   (\W) Show warnings after every statement.
nowarning (\w) Don't show warnings after every statement.
resetconnection(\x) Clean session context.
For server side help, type 'help contents'
```

Lesson 4. Administrative MySQL commands

Based from MySQL - Administration - Tutorialspoint, n.d., the following is the list of the vital MySQL commands, that you just might use time to time to work with MySQL database:

- USE Databasename: this will be used to opt for data among the MySQL work area.

• SHOW DATABASES: Lists out the databases that area unit accessible by the MySQL software system.

• SHOW TABLES: Shows the tables among the data once a piece of information has been chosen with the use command.

• SHOW COLUMNS FROM tablename: Shows the attributes, types of attributes, key data, whether or not or not NULL is allowable, defaults, and various knowledge for a table.

• SHOW INDEX FROM tablename: Presents the tiny print of all indexes on the table, furthermore because of the 1st KEY.

• SHOW TABLE standing LIKE tablename\G: Reports details of the MySQL software system performance and statistics.

Lesson 5. MySQL Datatypes

Based from MySQL - Data Types - Tutorialspoint, n.d., properly defining the fields in a table is vital to the improvement of your database. you must use only the type and size of the field you actually need to use. as an example, do not define a field ten characters wide, if you recognize you are solely aiming to use two characters. These type of fields (or columns) are referred to as knowledge types, after the type of data you will be storing in those fields.

MySQL uses many different data types broken into three categories:

Numeric

Date and Time String Types

Numeric Data Types

MySQL uses all the quality ANSI SQL numeric knowledge varieties, therefore if you are coming to MySQL from a different database system, these definitions can look acquainted with you. the following list shows the common numeric data types and their descriptions (MySQL - Data Types - Tutorialspoint, n.d.):

INT – A normal-sized whole number that may be signed or unsigned. If signed, the allowable vary is from -2147483648 to 2147483647. If unsigned, the allowable vary is from zero to 4294967295. you'll be able to specify a dimension of up to eleven digits.

TINYINT – a really tiny whole number that will be signed or unsigned. If signed, the allowable vary is from -128 to 127. If unsigned, the allowable vary is from 0 to 255. you'll be able to specify a width of up to four digits.

SMALLINT – a small whole number that may be signed or unsigned. If signed, the allowable vary is from -32768 to 32767. If unsigned, the allowable vary is from zero to 65535. you'll be able to specify a width of up to five digits.

MEDIUMINT – A medium-sized whole number that will be signed or unsigned. If signed, the allowable vary is from -8388608 to 8388607. If unsigned, the allowable vary is from zero to 16777215. you'll be able to specify a dimension of up to nine digits.

BIGINT – an outsized whole number that will be signed or unsigned. If signed, the allowable vary is from -9223372036854775808 to 9223372036854775807. If unsigned, the allowable vary is from zero to 18446744073709551615. you'll be able to specify a dimension of up to twenty digits.

FLOAT(M, D) – A number that can't be unsigned. you'll be able to outline the show length

(M) and also the variety of decimals (D). this is often not needed and can default to ten,2, wherever a pair of is that the variety of decimals and ten is that the total variety of digits (including decimals). Decimal exactitude will attend twenty-four places for a FLOAT.

DOUBLE(M, D) – A double exactitude number that can't be unsigned. you'll be able to outline the show length (M) and also the variety of decimals (D). this is often not needed and can default to sixteen,4, wherever four is that the variety of decimals. Decimal exactitude will attend fifty-three places for a DOUBLE. REAL could be a word for DOUBLE.

DECIMAL(M, D) – associate degree unpacked number that can't be unsigned. within the unpacked decimals, every decimal corresponds to 1 computer memory unit. process the show length (M) and also the variety of decimals (D) is needed. NUMERIC could be a word for DECIMAL.

Date and Time Types

DATE – A date in YYYY-MM-DD format, between 1000-01-01 and 9999-12-31. For example, December 30th 1973 would be stored as 1973-12-30.

DATETIME – A date and time combination in YYYY-MM-DD HH:MM:SS format, between 1000-01-01 00:00:00 and 9999-12-31 23:59:59. For example, 3:30 in the afternoon on December 30th 1973 would be stored as 1973-12-30 15:30:00.

TIMESTAMP – A timestamp between midnight, January 1st 1970 and sometime in 2037. This looks like the previous DATETIME format only, but without the hyphens between numbers; 3:30 in the afternoon on December 30th 1973 would be stored as 19731230153000 (YYYYMMDDHHMMSS).

TIME – Stores the time in a HH:MM:SS format.

YEAR(M) – Stores a year in a 2-digit or a 4-digit format. If the length is specified as 2 (for example YEAR(2)), YEAR can be between 1970 to 2069 (70 to 69). If the length is specified as 4, then YEAR can be 1901 to 2155. The default length is 4.

String Types

CHAR(M) – A fixed-length string between one and 255 characters long (for example CHAR(5)), right-padded with areas to the desired length once stored. defining a length isn't needed, however, the default is one.

VARCHAR(M) – A variable-length string between 1 and 255 characters long. for instance, VARCHAR(25). you want to outline a length once making a VARCHAR field.

BLOB or TEXT – A field with the most length of 65535 characters. BLOBs are "Binary massive Objects" and are accustomed to store massive amounts of binary data, like pictures or alternative types of files. Fields outlined as TEXT conjointly hold massive amounts of information. The distinction between the 2 is that the kinds and comparisons on the hold on information ar case sensitive on BLOBs and aren't case sensitive in TEXT fields. you are doing not specify a length with BLOB or TEXT.

TINYBLOB or TINYTEXT – A BLOB or TEXT column with the most length of 255 characters. you are doing not specify a length with TINYBLOB or TINYTEXT.

MEDIUMBLOB or MEDIUMTEXT – A BLOB or TEXT column with the most length of 16777215 characters. you are doing not specify a length with MEDIUMBLOB or MEDIUMTEXT.

LONGBLOB or LONGTEXT – A BLOB or TEXT column with the most length of 4294967295 characters. you are doing not specify a length with LONGBLOB or LONGTEXT.

ENUM – an enumeration, that could be a fancy term for a list. once defining an ENUM, you're making a listing of items from which the worth should be designated (or it may be NULL). for instance, if you wished your field to contain "A" or "B" or "C", you'd outline your ENUM as ENUM ('A', 'B', 'C') and solely those values (or NULL) might ever populate that field.

Assessment Task

Identification

_____ 1. A very small integer that can be signed or unsigned

_____ 2. A large integer that can be signed or unsigned

_____ 3. Datatype that the decimal precision can go to 53 places

_____ 4. DATE Format

_____ 5. A normal-sized integer that can be signed or unsigned.

_____ 6. What is the format for the TIME datatype

_____ 7. A fixed-length string between 1 and 255 characters in length

_____ 8. Stores a year in a 2-digit or a 4-digit format.

_____ 9. TEXT column with a maximum length of 255 characters.

_____ 10. Presents the details of all indexes on the table, including the PRIMARY KEY.

References:

Codd's 12 Rules - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/dbms/dbms_codds_rules.htm

DBMS - Data Independence - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from

https://www.tutorialspoint.com/dbms/dbms_data_independence.htm

DBMS - Data Schemas - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/dbms_data_schemas.htm

ER Model - Basic Concepts - Tutorialspoint. (n.d.). Retrieved August 18, 2020, from https://www.tutorialspoint.com/dbms/er_model_basic_concepts.htm

Gupta, S. B., & Mittal, A. (2017). Introdution to Database Management System (2nd ed.). University Science Press. https://doi.org/10.1201/9781315372679-7

Lemahieu, W., Broucke, S. vanden, & Baesens, B. (2018). Principles of Database Management. Cambridge University Press. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

MySQL - Administration - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/MySQL/MySQL-administration.htm

MySQL - Data Types - Tutorialspoint. (n.d.). Retrieved August 20, 2020, from https://www.tutorialspoint.com/MySQL/MySQL-data-types.htm

Relation Data Model - Tutorialspoint. (n.d.). Retrieved August 25, 2020, from https://www.tutorialspoint.com/dbms/relational_data_model.htm

Williams, H. E., & Tahaghoghi, S. M. M. (2006). Learning MySQL (A. Oram (Ed.); 1st ed.).

O’Reilly Media.

- **END OF MODULE FOR PRELIMINARY TERM PERIOD –**

END OF THE PRELIM TERM MODULE EXAMINATION FOR PRELIM PERIOD IS ON SEPTEMBER 21 TO OCTOBER 3, 2020. CHECK YOUR EXAM SCHEDULE DO NOT FORGET TO TAKE THE EXAM AS SCHEDULED