Database Management System

Date 03/06/2021
Page

→ Introduction to DBMS :-

* **Data :-**
  The data is defined as the collection of facts and figures that when processed then generates information.

  Eg :- text, numbers, images, audio, video, speech, map

* The data may be either related to any particular place, person, object or entity

* The data is the plural of "datum".

* The hierarchy of data storage in database is as per follows -

i) **Bit :-**

→ The smallest item of data in a single binary digit either in form of 0 or 1.

→ The '0' denotes to stop bit, low volt, open switch, black in colour whereas '1' represents start bit, high volt, closed switch.

→ Both '0' and '1' represents machine code of the system.

Date
Page

i) Charater -

→ Multiple related bits are combined ~~each other~~ to form a character.

→ A bit is the basic unit of primary and secondary storage.

→ The character is the basic unit of human perception.

ii) Field :-

→ A field is the meaningful collection of related character.

→ It is the smallest logical data entity which is treated as single unit in data processing.

iv) Record :-

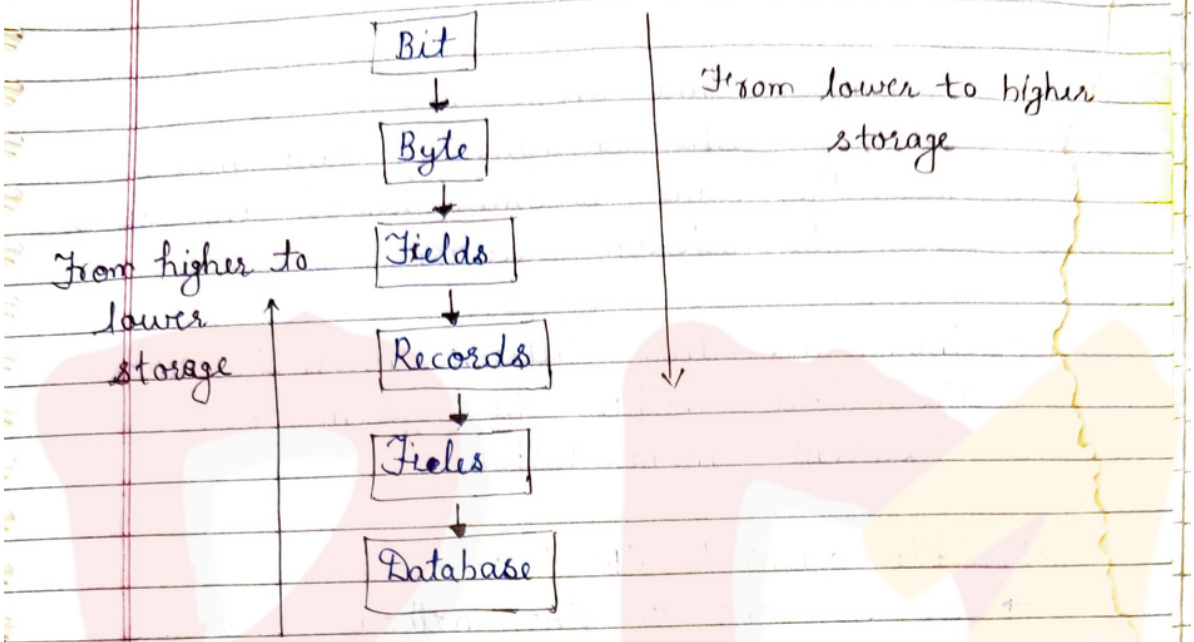
A record is a collection of related fields, which are treated as single unit.

v) File :-

Multiple related records are combined to form a file.

vi) Database

Multiple related files are integrated to form a database.



* Database (DB) -
  Collection of interrelated data.
  Eg:- i) Traditional Database (TDB) - text and numbers
  ii) Multimedia Database (MDB) - Video, speech, song, movie
  iii) Geographic information system (GIS) - Image of earth

* Database management system (DBMS) -
  Set of programs or software used to define, store and manipulate the database in such a way that it becomes easier to retrieve.
  Eg:- Oracle, MS-access, FoxPro, MySQL, IBM DB2, MS SQL, dBASE, SQLite

* DB + DBMS → Database system (DBS)

Date
Page

Database approaches -

→ Organizing data in computer -
The organizing of data means how the data becomes stored and accessed from the storage location.

→ There are two different ways to store the data within the system for different manipulations -

i) File oriented approch :-

→ It is one of the methods that is implemented in application where the data is organized in form of file.

→ In it the data becomes stored in columns

→ To support the file based operation the operating system provides one of the method called "file management system".

→ The "FMS" is responsible to open, close, ~~uptac~~ update, delete and manipulate the data of the file.

→ There are certain demerits of file oriented approach that is -

i) Data redundancy (Duplicate value input) and inconsistency.

ii) No mathemetical operation.

iii) No any type of function.

ii) Data base oriented approach -

→ When the data is being stored in tabular formate then this approach is to be used.

→ In tabular (table) formate, the concept of columns and rows are available in which column represents fields whereas rows represents record of the database.

* In database, the concept of table is being categorized into two different category that is -

A) Based on storage -

A/c to the storage of the data in database there are different types of table is available -

i) Master table -

The table that contains overall records of the database related to any particular person, place, object or entity is known as master table.

→ The general user does not work on master table directly.

→ Only database administrator is responsible to work on the master table.

ii) Transaction table -

The table through which a user can work on master table is known as transaction table.

→ A database can have more than one transaction tables.

iii) Output table -

When the query statement or application programm is written and executed then some resultant becomes available which do not be stored permanently, called output table.

iv) Report table -

When the query statement or application programm is written and executed then some resultant becomes available and be stored permanently called report table.

v) Backup table -

When the a table is created in the database and for security purpose if a copy is being stored inside or outside the system as a copy then it is known as backup table.

vi) Dump table -

When the database working for a long period of time then some malfunctions occurs in their processing due to some file misplacement or dilution then this table contains the references of all such files. When the repairing of database started then it accesses the references of all such files to put in the database

B) Based on ~~storage~~ representation -

To represent the database graphically the notation of table is being performed by using different tables. these are -

i) Dependent table -

The table whose fields are depends on other table then it is known as dependent table.

ii) Independent table -

The table whose fields are not depends on another table of the database is known as independent table.

iii) Strong table -

The table without which the processing of database is not completed is known as strong table.

iv) Weak table -

The table whose availability does not have major contribution in database of data processing is known as weak table.

v) Super table -

That table that provides the property to other table is known as super table.

vi) Sub table -

The table that accesses the property from other table is known as sub table. The sub table does not have their own property.

* Database Model

* The database model is the structural representation which defines how data is connected to each other and how they are processed and stored inside the system.

* Data model gives us an idea that how the final system will look like after its complete implementation.

The type of storage device determines about the database model for working - When the storage device is magnetic tape then the sequential operation

happens and it supports "hierarchical model" operations.

* When the magnetic oxide based device is used then "relational model" is created it supports index-sequential nature of operation.

* If the optical memory is used then the concept of "network model is used because it supports direct or random access.

i) Hierarchical Data Model - (1st DBMS model)

The hierarchical model is the simplest model in which sequential nature of storage and accessing takes place.

→ In hierarchical model, data is organized into a tree like structure ~~whic~~ with each record is having one parent record and many children.

→ In this model, magnetic tape based storage device is used therefore the storage capacity of data is limited in nature.

→ Due to sequential nature of operation, the hit ratio of this model is "less".

Date
Page

* Hit ratio -

The hit ratio is calculated by dividing the total no number of records accessing by the time taken by the operation.

→ A new file is always created when a new record is inputted or deleted when the database is full.

→ The record is not adjusted due to the unavailability of overflow memory ~~data~~ area in it.

→ The binary search technique is used to search the record in the hierarchical model.

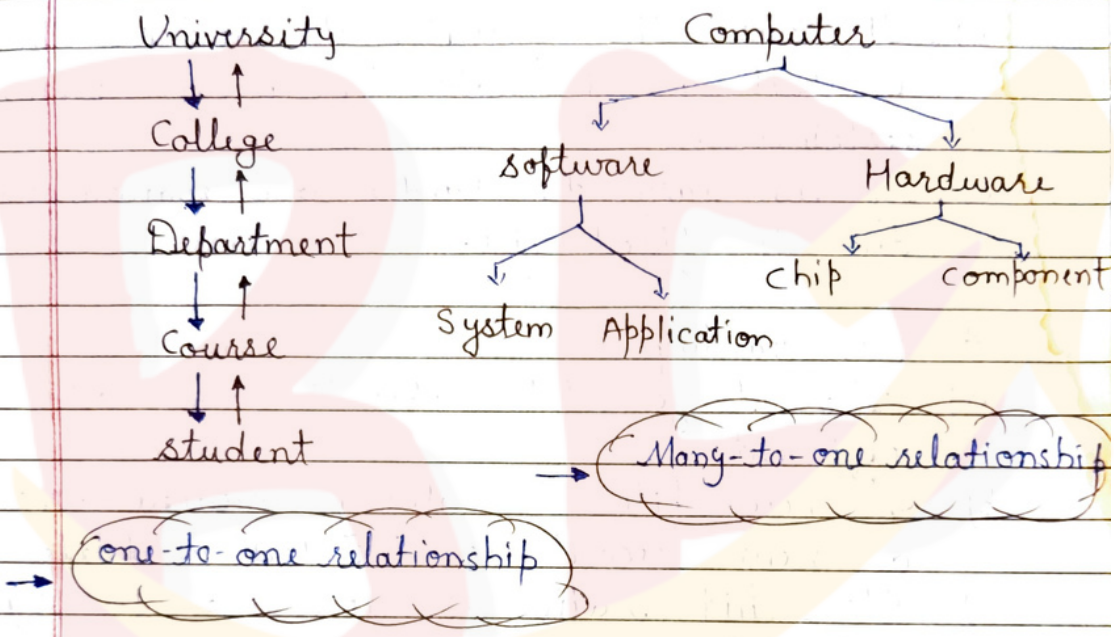
→ A single record access time is half of the average of total record search time.

There are two different types of relationship available in this model-

i) One-to-one relationship-

In this relationship, every one table intract with only another table at a time.

ii) one to many relationship -
In this relationship, one table interacted with more than one table at a time.



* The alternate key is used to convert one relation with other relation in the database enviroment.

* If Two work with hierarchical model, the parent and child relationship is being followed -

a) A parent can have more than one child but the child does not have more than one parent.

b) To remove the parent record it is

necessary that all of the child records must be deleted from their grievance.

* Advantage of heirarchical model-

a) The cost of implementation of this model is less.

b) It has the facility to restrict the overlapping of data by giving certain restriction.

c) It is easy to handle and maintained due to fixed size.

* Disadvantage -

a) The large data operation is not efficient.

b) A new file is always created when a new record is inserted or deleted.

c) Its hit ratio is less due to sequential behavior of operation.

d) The efficiency decreased when the size of file increased.

ii) Relational Model -

In database working environment when two or more than two table are related to work then there data flow must be maintained among them. The model that describes this ~~model~~ working is known as relational model.

The relational database management system (RDBMS) uses this model to create and manipulate the data using this model.

→ It uses collection of tables for representing data ~~n~~ and the relationships among those data.

→ Data is stored in tables called Relations.

→ Each table is a group of column and rows, where columns represents attribute of an entity and rows represents records (tuples).

* Attribute or Field -

Each column in a table is called an attribute. The value of the attribute should be from the same domain.

Eg:- Student Id, student name, student age etc.

* Tuple or Record :-
Each row in the relation called tupple. A tupple defines a collection of attribute values. So each row in a relation contains unique values.

→ This model was initially described by Edgar F. codd, in 1969.

→ In this model, magnetic oxide based database used for their working.

→ The index-sequential pattern of working with the records happens in this model.

→ The hit ratio of this modes based operation is more than hierarchical model.

→ No extra file is created when a new record is ~~inst~~ inserted or previously available record being deleted from the database file. It is because that operating system provides an overflow memory area to the file.

→ The relational model supports master and transaction table concept for management of data in database.
To creat the relationship among

the table, a key known as 'primary' key is used in the database.

→ The table that having the primary key is know as master table and all other related table is known as transaction table. There is only one primary key being placed on the column of the table.

There are four different types of relationship being developed in relational model that is -

a) One-to-one -
In which all the related tables do not have the duplicate record. Each table can have primary key.

b) One-to-many -
When the first table having the primary key and other related tables do not have the primary key then it is created. The main table do not have duplicate records.

c) Many-to-one -
When the first table do not have the primary key then ~~this relation stop being assumed. It is hypothetical in nature that means it is not created inside system.~~ and all related tables have the primary key then is created.

d) Many to many -

When ~~to~~ no any ~~type~~ table have the primary key then this relationship being assumed. It is hypothetical in nature that means it is not created inside the table.

When the processing started then it is necessary that the master table must be contains the records if a user need to work on the same type of record from transaction table. The field that having primary key do not be left blank and not contains duplicate record.

When the relationship need to be ~~to~~ established then certain perspective being followd, these are -

a) Enforce referential integrity -

To creat the relationship this rule must be followd when the tables are attached.

b) Cascade update record -

When the record of master table being updated then the corresponding record from the transaction table also be updated automatically.

c) Cascade delete record -
When the record of master table being ~~upd~~ deleted then the corresponding record from the transaction tables also be deleted automatically.

* Advantage of Relational model -

a) The large capacity of data storage behavior being supported.

b) No extra file is created when a new record being inserted or deleted.

c) The hit ratio is more in nature.

d) The data flows among the tables are maintained efficiently.

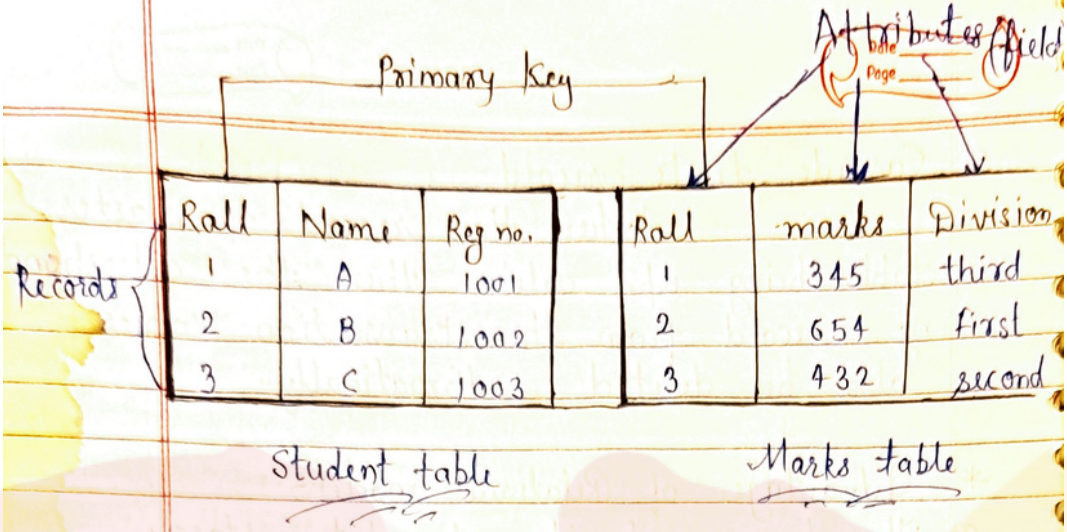
e) Multiple relation can be created among the tables.

* Disadvantage -

a) It is costly to creat and maintain.

b) The complexity is more because more than one indexes.

c) When master table do not operated, all trasaction table are dumb.



Roll	Name	Reg no.
1	A	1001
2	B	1002
3	C	1003

Student table

Roll	marks	Division
1	345	third
2	654	first
3	432	second

Marks table

1) The data being stored in column using the file oriented approach.

2) The file is the combination of records.

3) The database model supports logical entity representation.

4) The table that accesses all of their properties from other table called sub table.

5) The table that temporarily stores the resultant of query or program Output tab

6) The horizontal arrangement inside the table is know as column.

7) The smallest unit of the table is known as

8) In which approach the mathematical operations do not be performed.
→ File oriented.

9) The table is also known as Entity.

10) The vertical arrangement in the table is known as records.

11) The network model uses which type of memory - Optical memory.

12) The hit ratio is represented by -

$$\text{Hit Ratio} = \frac{\text{No. of records}}{\text{total time.}}$$

13) Which key is responsible to convert one-to-one into one-to many relation- Alternate key.

14) The collection of bits is known as - character.

15) The corrupted file of database becomes stored inside which table - Dump table.

16) The table whose property depends on another table is known as Dependent table.

17) The entity relationship diagram is created by the tables based on represen-tation.

18. A table do not have duplicate column.
19. The bit is the basic unit of primary or secondary storage.
20. The stop bit is denoted by '0'
21. The character is used for human perception.
22. The meaningful collection of related characted is known as field.
23. When child table access the property of parent table then generalization is used.
24. The data when processed then information finally is generated.
25. The information without ~~table~~ which the processing of database is not completed is known as formal.
26. The term data is derived from a greek word called Datum.
27. The highest form of data hierarchy is Database.
28. The magnetic oxide based storage support which type of processing in data model- Index- sequential.

Date
Page

iii) Network Model -

The network model was proposed by an organization named DBTG (Database Task group) in 1970.

The drawback of hierarchical model become rectified under this model.

The network model basically supports client/server architecture for its operation and maintance of database.

Both the client and server is the software that works in requester and accepter format.

The network model specify that a single child can have more than one parent either it is in hierarchy or direct connectivity.

The optical memory based storage being used to maintain the operation it in it. Due to optical memory based operation the direct or random accessing or manipulation of data can be performed by the user.

Date
Page

A function known as hash function is used to support the direct accessing of records from the database. The hash function creates a table known as hash table that contains the column value which is unique in nature. The reference being attach with the main table. Once a user try to work on any of the record then the verification happens inside the hash table instead of main table. The entire processing of record working is known as hashing. The hit ratio is too high rather than other model.

→ There are two different modes in which network model works that is -

a) Inverted list -

From a single database table when a single or multiple records need to be accessed then it is used.

b) Multilist -

From multiple database tables when more than one records based operation if performed then it is used.

Advantage -

i) The record can be directly accessed therefore efficiency is more in nature.

ii) We can use this model as alternative of hierarchical model.

iii) The large scale of data being stored and can be manipulated singly or collectively.

iv) Due to centralized behaviour the security consideration is highest among the models.

* Disadvantage -

i) The cost of working and maintenance is highest.

ii) If the server fails to work then entire environment being disturb.

iii) The unnecessary memory of database being utilized due to creation of hash table for each table of the database.

NOTE :-

The hierarchical, relational and network model is divided as per the storage and manipulation of the records in database environment.

Date
Page

iv) Entity Relationship Model

This model was proposed by a scientist named P.P chen. It not only introduced ER model but also a corresponding diagrammatical technique.

The ER data model is based on a perception of a real world which consists of a set of basic object called entity and relationship among these object.

The entity relationship is a high level conceptual model that describes data as entity, attributes and relationships.

There are three views of database once it is created for its use-

a) External view -
The number of tables decided, fields in each table, how the data being flow among the related tables, types of relationship and security etc.

b) Conceptual view -
Once the observations and perceptions being programmed.

c) Physical view -
The location of data storage, how much amount of memory required for database being decided.

Date
Page

* Entity (Table) -

The combination of rows and columns is known as table.

i) Columns -

The column is the smallest unit of table. It is also known as field (attributes). Every column having certain characteristics that is -

a) The column name must be unique in a table.

b) Every column must having a name that do not be database keyword.

c) Every column must have certain data type.

d) Every column must have the specification of storage capacity.

e) The column of one table when ~~creat~~ relate with other then referential integrity being followed.

f) The horizontal arrangement and vertical storage inside the table is known as column.

g) The column name do not contains space, symbol and digit. Only underscore is used for giving the name.

ii) Row -

The combination of fields is known as row. The row is also known as record.

→ When more than one records having the same nature then it is known as tuple.

→ The storage capacity of the row is the combination of size of columns.

→ The vertical arrangement and horizontal storage is known as row.

* An entity is an object that exists and is distinguish from one object to other.

* An entity is any object, place, person, concept or activity about which an enterprise recording of data. The entity becomes created diagrammatically in rectangle. The name of entity in ER diagram be in capital letter.

Every entity having two concepts, that is -

a) Entity type -

An entity type is a set of entities of same type which shares common properties.

Eg:- Student, teacher, customer.

b) Entity instance -

An entity instance is a specific individual thing or object. When two entities are related to one other than entity instance being written.

Eg:- Teacher teaches student, where 'teaches' is the entity instance.

* The name of entity type must be in capital letter whereas the name of entity instance is in upper - lower letters.

Relationship

A relationship is an association among the entities. Every relationship having two major parts that is -

a) Relationship type -
The type of relationship established between two entities.

→ The relationship type may be one - to - one, one - to - many, many - to - one and many to many

b) Relationship instance -
When the relation is established either with two tables or more as well as between the columns of tables.

→ If tables are related then intity integrity concept whereas if columns are attached then referential integrity being established.

* Once we think about the relationship then we think to work on RDBMS.

→ One is denoted by '1' or 'I'

→ Many is denoted by '>' or '∞'.

Attribute / Column / fields

An attribute is the property of a given entity. Every attribute having two components that is -

a) Attribute type -
An attribute type is the property of entity type.

b) Attribute instance -
An attribute instance is a particular property of an individual entity instance.

There are three different classes of attributes that is -

a) Single - valued attribute -
If the table attribute contains only one value to store is know as single valued attribute.
Eg:- Roll of student, Account no. of custo

b) Multi-valued attribute -
If the table attribute contains a number of information then it is known as multi-valued attribute.
Eg:- address of student or employee.

c) Composite attributes -
The attributes which can be decomposed into a number of separate attributes is known as composite attribute.
Eg:- Salary of employee (Basic, allowance, tax, net salary).

* Symbols used in ER digram -

i) Rectangle - It specifies the independent entity.
ii) Double rectangle - It specifies the dependent entity.
iii) Oval / Ellipse - It represents the attribute of a table.
iv) Diamond - It represents the entity instance of the relationship.

* Rules or guidelines of ER creation.
1) No duplicate table name is used in a single ER diagram.
2) The name of the entity must be written in capital letter.
3) An ER diagram can have more than one attribute having the same name but

in different tables.

4) The name of attribute also written in capital letter.

5) When two tables are related to one other then it is necessary that the entity instance being written between them.

6) When super and sub table being related to one other then it is not necessary to given the entity instance word of relationship.

7) When two attributes are attached togethes then no- of need to put the entity instance ~~sys~~ symbol b/w them.

8) When the dependent table interact with independent table then it is also not necessary to use the entity instance symbol

Entity Relationship Diagram

* Employee Management System -

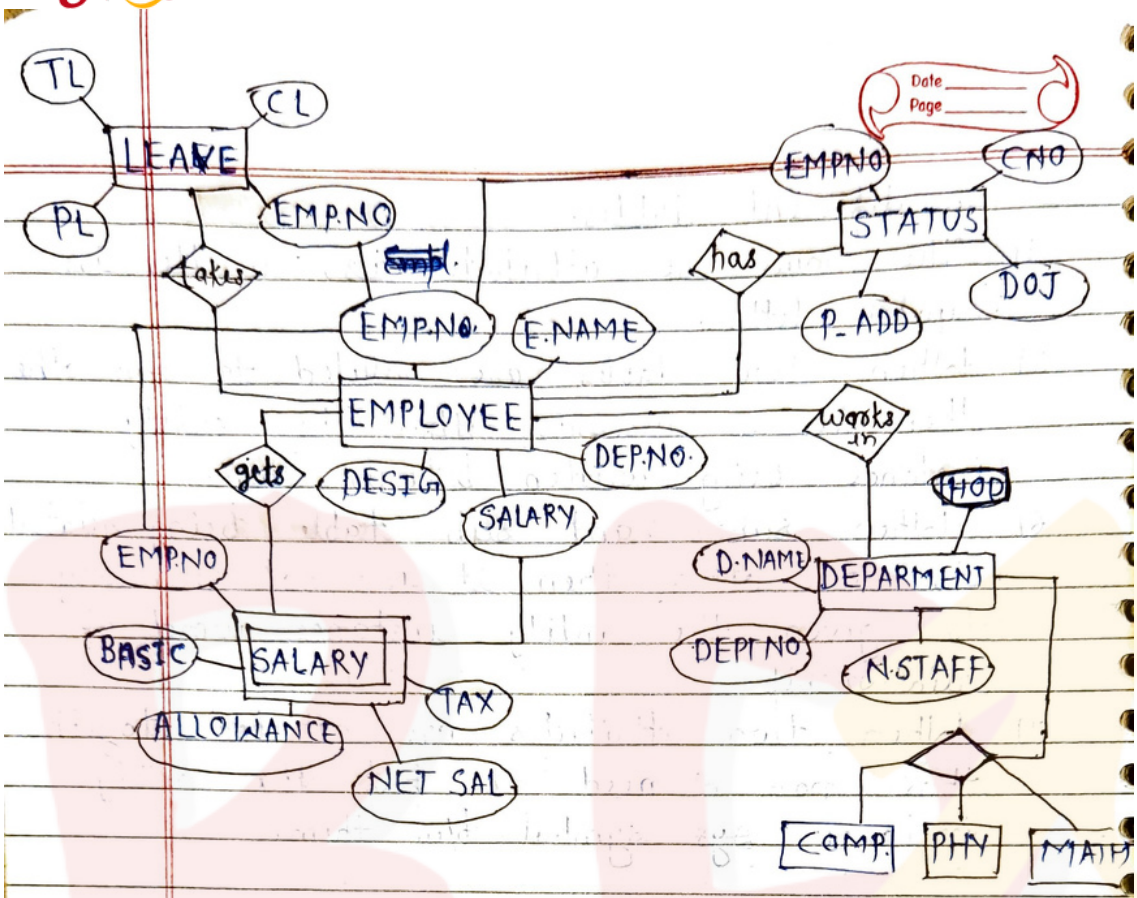
Employee - (Emp. no., E. name, Desig., Dept. no., Salary)

Salary - (Basic, allowance, tax, net salary)

Department - (Dept. no., D. name, No. of staff, HOD, Section)

Leave - (Emp. no., CL, PL, ML, EL, LWP, TI)

Status - (Emp. no., P. add, Email, Blood gr., Hobbies



* ER diagram of Hospital Management system-

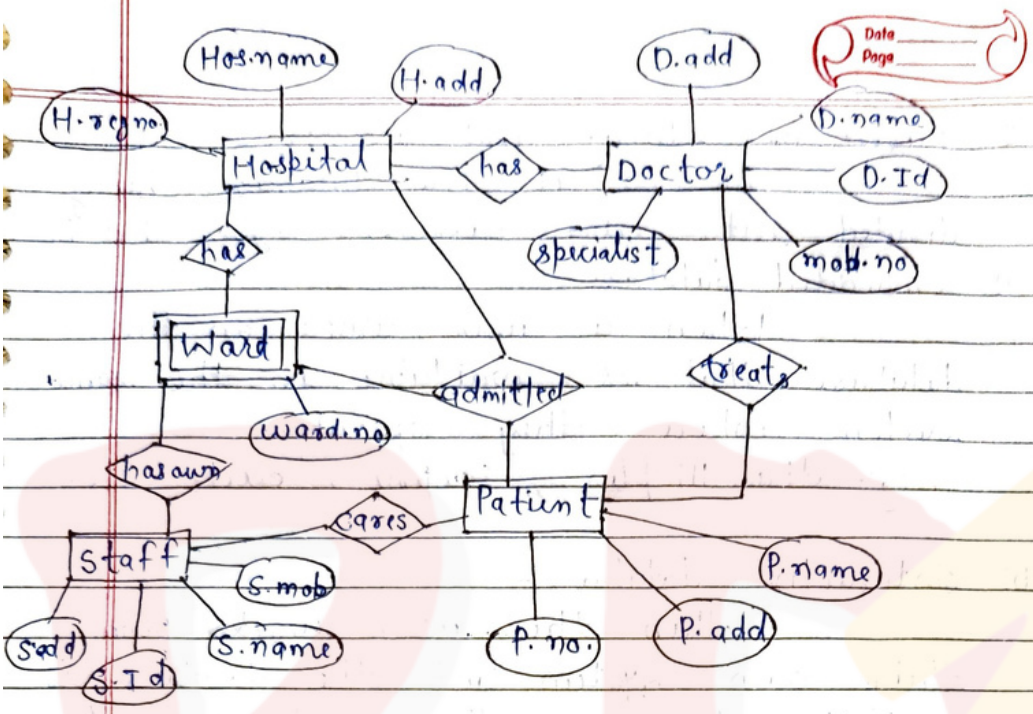
Hospital - (H.name, H.Id, H.address, H.reg.no)

Doctor - (D.name, D.Id, D.mobno., D.add., specialist)

Patient

Staff

Ward



Database Management System (DBMS)

* Database -

The database is a collection of inter-related data.

There are several categories of the database

i) Based on approaches -

The pattern a/c to which the data being stored in the database.

a) File based approch -

Once the data being stored in application file then column way of storage being performed.

b) Object-oriented approach -

When the data being stored in form of tabular structure then it support object oriented behavior.

2. Based on storage and accessing -
A/c to it the database being divided into two different types that is -

a) Embedded database -
When a user work on the database that is available in the same system where they operate.
Eg:- dbase III plus, Foxpro, Excle etc.

b) Interactive database -
When the user works from their own system but the database is available elsewhere then such particular on is known as interactive database. Mostly when we need to work on it then the system must be in network.
Eg:- MS-Access, Oracle, MYSQL, SQL ser

3. Based on number of tables -
A/c to ~~the~~ it the number of tables behaves as database. It has two different types -

a) Database management system (DBMS) -
When a single table based operation happens inside the database they behave as DBMS.

b) Relational database management system (RDBMS) -
When two or more than two tables related to one other then the concept of RDBMS becomes used. It becomes classify into different types -

i) Object oriented RDBMS -
When the feature of object orientation like inheritance, polymorphism and encapsulation being attached with the RDBMS then it is know as OODBMS.

ii) Knowledge database -
When the RDBMS having the facility of decision making operation as per the comparison of process and artificial intelligence then it is known as knowledge database.

iii) Design database -
When the RDBMS used inside the CAD (computer aided drafting/designing), CAM (computer aided manufacturing) and CASE (computer aided software engineering) then it is known as design database.

* Management -
To maintain overall operation of the database is known as management.

a) Creation of database -
Think about the table name, no. of tables, column names, data types, size, relationship and location of storage.

b) Alter the database -
The change happens inside the structure of the table by adding or removing the columns, change in data types of columns, changing in the name of table and location.

c) Insertion -
To insert a record inside the table as per the columns specification.

d) Update -
To change the numeric value or replace the text value with new value will creat the update operations.

e) Delete -
To remove the data eithes selectively or collectively it is used.

f) Search -
To access any specified record from the database, it is used.

g) Short -
To arrange the records in ascending or descending order, it is used.

h) Merge -
To attach one table data with other table, it is used.

i) Copy - To copy the data or structur of a table into another table it is used.

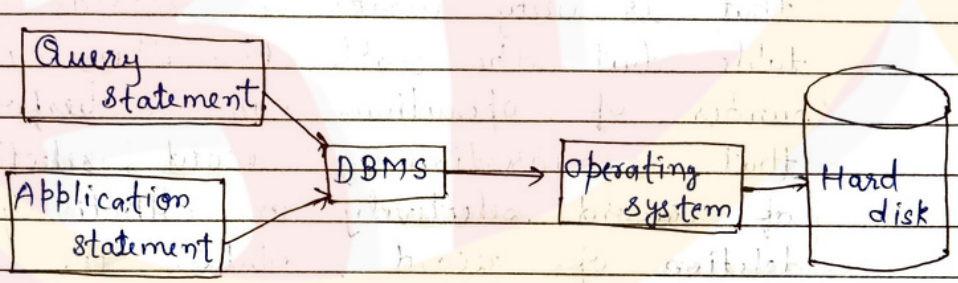
j) Rename - To change the name of the table, it is used.

Date
Page

* System -

The systematic approch of working on database by writing query statement or application program to get some resultant is by the use of system.

* "Database Management system is the collection of software and hardware of the system through which overall operations on the database being managed systametic"



→ "DBMS" is the combination of software and hardware through which overall operations of the database being performed or managed.

→ The DBMS acts as mediater b/w user written query statement or application program and hardware of the system.

Componet of DBMS-

By the use of components of DBMS the execution of query or program happens collectively until and unless the ~~set~~ result being prompt. There are seven different components of DBMS-

Date
Page

i) DDL compiler -

The ~~dat~~ Data Definition Language compiler is one of the component that is responsible for table creation, change in the structure of table that is alter table and remove all records from the ~~the~~ table at once (truncate table).

ii) DML compiler -

The Data Manipulation Language precompiler is one of the components of DBMS that is only works on records of the table but the table too. There are a number of operations can be performed that is insertion of record, selection of record selectively or collectively, deletion of record, update the record, copy the record, copy the table, copy the structure of table, copy the file, move of table, move of file, rename of table etc.

iii) File manager

The file manager is one of the component of DBMS that is used for naming convention of table, decide the storage location of table, amount of memory (allocation / de-allocation) required for table and helps in accessing of the table on which query or application program is written.

iv) Query processor -
It is responsible to execute the query statement or application program for generating result. It gets help from the file manager of DBMS.

v) Database manager -
It is one of the software parts that control the working of database in single user environment.

a) Schema definition -
When the database table as well as their records if graphically represented then such one is known as schema. Once the schema is divided then it termed as sub-schema. There are three different types of schema that is external, conceptual and internal schema.

b) Security -
The security is required to inhibit the accessing of data from unauthorized user. There are two ways to secure the data of database that is putting proper authentication (ID + password) for external security and hiding as well as converting data into unintellectual format (~~copy~~ cryptography) for internal security in the database.

c) Consistency -
It is the combination of integrity and accuracy. ∴ The consistency of the database being regularly checks by it.

d) Backup and recovery -

When the copy of the data need to be performed and stored either external or internal to the system then backup is created. When certain data being loss due to any happening then using the backup when we get it secondary then it is known as recovery.

vi) Database administrator (DBA) -

When the database is installed inside the server or central system then to maintain the database, a person known as database administrator is used. The concept of DBA activated whenever the data related operation is form network & locations. There are a numbers of function performed by database administrator that is -

a) Database installation and maintenace -

The DBA is responsible to install the database properly in the system with the accessing facility either by single or multiple user envirsoment.

→ If the working of database is not in proper way as required then it is the responsibility of administrator to solve the problem in the minimum time.

b) Data extraction and transformation -

In the database environment there are different types of users created by DBA and each user having a specific permission to work on it for extraction. If a user need to transform the mode of operation and working enviroment of previously available database into new one then it is the responsibility of administrator to do this.

c) Database backup and recovery -

The database backup and recovery can being possible in offline and online mode by the facility of database administrator.

d) New user creation -

It is the responsibility of administrator that a new user is created with proper identification and authentication. It is also recommended that the type of permission is also facilitated at the time of their creation.

e) Security -

It is the responsibility of the DBA to secure the data from virus, unauthorised users, S/H failure to protect the data. The external and internal security is being also be facilitated with the database by DBA.

f) Performance monitoring -

The DBA is responsible to regularly verify the working of the database in both less or more workload. It is necessary that the database must be operated in efficient way of data accessing and acceptance.

g) Troubleshooting / Debugging -

If the database do not support the hardware platform on which it operates as well as the network through which it works then it is the responsibility of the DBA to sort out all the problems by their own level.

7) Data dictionary -

When the database table is created then the table itself stores inside the data dictionary or DBMS software. The data dictionary stores the data inside the table as per the ascending order of the key field (field that have primary key). Once the query statement or application program is written on any database table then it access from the data dictionary. The DDL and DML operations is indirectly happens inside the data dictionary. There are two types of data dictionary that is -

a) Active data dictionary -
When the user works on database that is available in the same location where the user works called data dictionary

b) Passive data dictionary -
The data dictionary that is available elsewhere and a users operates from their own location then such data dictionary is known as passive data dictionary.

* Advantages of DBMS -
The DBMS is come of the application programming enviornment in which the a number of operations can be performed by the users. There are a number of characteristics of DBMS that is -

i) Data sharing -
Once the database becomes stores and different users need to access it form there location then the concept of data sharing becomes performed. The database must be stored in client server working enviornment.

ii) Data integrity -
The term integrity is made up of by combination of accuracy and consistency. Whenever the data, its we inputted inside the database or ~~accurate~~ as accessed then data must be accurate as per their data type, as well as the query

statement or application program interacti-vity. The consistency specify the same pattern of data inputting in the database.

iii) Data security -

The data is vital for every organisation, so it is necessary to secure from unauthorised user of working enviorment. As per the concern of security, of the external or internal security can be put by database. The external security is by proper identification and password whereas the internal security is by using hiding of data or convert the data in unintelluctual format.

iv) Conflict resolution -

Whenever a single database table need to access from multiple locations either for the same or different operations then a conflict becomes generated during the processing. If the DBMS creates a copy of the database table as many times has the number of users access it then each and every users works on the transaction table of the Therefore no any conflict bicomes arises.

v) Data independence -

It is one of the advantage of the DBMS in which a user is totally free from both the hardware and software site. There are two different

types of independence that is -

a) Physical DI -

Whenever the data table is created then & how much amount memory being required by the database is not decided by the user as well as once the records or table being removed from the database then how memory will release is totally free from users knowledge.

b) Logical data independence -

Whenever the query statement or application program becomes written by the users then how it becomes executed by the query processor with the help of file manager and data dictionary is totally hidden from users knowledge called logical D.I.