UNIT I

1.Define parallel DBMS.

A DBMS running across multiple processors and disk that is designed to execute operations in parallel whenever possible inorder to improve performance.

2.What are the different parallel db architectures?

→Shared memory

→Shared disc

→Shared nothing

→hierarchial

3.Differentiate interquery and intraquery parallelism.

Interquery parallelism:

Different queries or transactions execute in parallel with one another.It increases scaleup and throughput.

Intraquery parallelism:

It refers to the execution of a single query in parallel on multiple processors and disk.It is important for speeding up long running queries.

4.Differentiate intraoperation parallelism and interoperation parallelism.

Intraoperation parallelism:

Speed up processing of a query by parallelising the execution of each individual operation.

Interoperation parallelism:

Speed up processing of a query by executing in parallel the different operations in a query expression.

2 types,1.pipelined parallelism 2.independent parallelism

5.Define distributed DBMS.

The software system that permits the management of distributed database and makes the distribution transparent to the user.

6.What is the fundamental principle of distributed DBMS?

The fundamental principle of DDBMS is to make the distributed system transparent to the user that is to make the distributed system appear like a centralised system.

7.List any four advantages and disadvantages of DDBMS.

Advantages:

→capacity and incremental growth

→reliability and availability

→efficiency and flexibility

→sharing

Disadvantages:

→ managing and controlling is complex

→less security because data is at different sites.

8.Define homogenous and heterogenous DDBMS.

Homogenous DDBMS:

In all sites the same DBMS product will be used.It is easier to design and manage..advantage:Easy communication,possible to add more sites,provides increased performance.

Heterogenous DDBMS:

Sites may run different DBMS product which need not be based on same data model.Translations are required for communication between different DBMS.Data may be required from another site that may have different hardware,different DBMS product,different hardware and different DBMS product.

9.What are the major components of DDBMS?

There are four major components in DDBMS,

(1)Local DBMS component(LDBMS)
(2)Data Communication component(DC)
(3)Global System Catalog(GSC)
(4)Distributed DBMS component

10.What are the correctness rules for fragmentation?

Any fragment should follow the correctness rules.There are 3 correctness rules.They are,

(1)Completeness
(2)Reconstruction
(3)Disjointness

11. Define multiple copy consistency problem?

Multiple copy consistency problem is the problem occurs when there is more than one copy of a data item in different locations. To maintain consistency of the global database, when a replicated data item is updated at one site all other copies of the data item must also be updated. If a copy is not updated, the database becomes inconsistent.

12. Define distributed serializability?

If the schedule of transaction execution at each site is serializable, then the global schedule is also serializable provided local serialization orders are identical. This is called distributed serializability.

13. What are the different types of locking protocols in DDBMS?

The different types of locking protocols employed to ensure serializability in DDBMS are centralized 2PL, primary copy 2PL, distributed 2PL and majority locking.

14. What are the types of deadlock detection in DDBMS?

There are three common methods for deadlock detection in DDBMSs: centralized, hierarchical and distributed deadlock detection.

15. What is the general approach for timestamping in DDBMS?

The general approach for timestamping in DDBMS is to use the concatenation of the local timestamp with a unique identifier, <local timestamp, site identifier>. The site identifier is placed in the least significant position to ensure that events can be ordered according to their occurrence as opposed to their location.

16. What are the phases of 2PC protocol?

The two phases of 2PC protocol are:

a voting phase and
a decision phase.

17. Define cooperative termination protocol?

Cooperative termination protocol is defined as the termination protocol which blocks the participant without any information. However the participant could contact each of the other participants attempting to find one that knows the decision.

18. What is the use of election protocols?

If the participants detect the failure of the coordinator they can elect a new site to

act as coordinator by using election protocols. This protocol is relatively efficient.

19. Define 3PC?

The three phase commit is an alternative non blocking protocol. It is non blocking for all site failures, except in the event of the failure of all sites. The basic idea of 3PC is to remove the uncertainty period for participants that have voted COMMIT from the coordinator. 3PC introduces a third phase, called pre-commit, between voting and the global decision.

20. Define Distributed Query Processing?

The process of converting high level language query into low level language with effective execution strategy depends in order to achieve good performance is called query.processing. In distributed query processing the query was distributed and processed in different locations.

21. Write the differences between locking and non-locking protocols?

Locking protocol	Non-locking protocol
1. In this Locking guarantees that the concurrent execution is equivalent to some serial execution of those transactions.	1. In this timestamping guarantees that the concurrent execution is equivalent to a specific serial execution of those transactions, corresponding to the order of the timestamps.
2. This involves checking for deadlock at each local level and at the global level.	2. This does not involve checking for deadlock at any level.
3. This protocol does not involve generation of timestamps.	3. This protocol involves the generation of unique timestamps both globally and locally

22. Define transaction.

A transaction is a unit of program execution that accesses and possibly updates various data items. Usually, a transaction is initiated by a user program written in a high-level data-manipulation language or programming language, where it is delimited by statements of the form begin transaction and end transaction. The transaction consists of all operations executed between the begin transaction and end transaction.

23. What are the desirable properties of transaction?

a) Atomicity
b) Consistency
c) Isolation
d) Durability

24. Define the term aborted.

In the absence of failures, all transactions complete successfully. However a transaction may not always complete its execution successfully. Such a transaction is termed aborted.

25.Define the term rolled back.

Once the changes caused by an aborted transaction have been undone, we say that the transaction has been rolled back.

26. Define the term committed.

A transaction that completes its execution successfully is said to be committed.It is used to store data permanently on database

27. How can we undo the effects of a committed transaction?

Once a transaction has committed, we cannot undo its effects by aborting it. The only way to undo the effects of a committed transaction is to execute a compensating transaction.

28. Define the term terminated.

A transaction is said to have terminated if has either committed or aborted.

29. What are the transaction states?

Active - the initial state, the transaction stays in this state while it is executing

Partially committed - after the final statement has been executed.

Failed - after the discovery that normal execution can no longer proceed.

Aborted - after the transaction has been rolled back and the database has been restored to its state prior to the start of the transaction.

Committed - after successful completion.

30. Define the term conflict in transaction processing?

We say that Ii and Ij are conflict if they are operations by different transactions on the same data item, and at least one of these instructions is a write operation.

31. Define conflict equivalent.

If a schedule S can be transformed into a schedule S' by series of swaps of non conflicting instructions, we say that S and S' are conflict equivalent.

32. Define conflict serializable.

The concept of conflict equivalence leads to the concept of conflict serializability. We say that a schedule S is conflict serializable if it is conflict equivalent to a serial schedule.

UNIT II

1.Define OODM?

OODM- Object Oriented Data Model

A (Logical)data Model that captures the semantics of objects supported in object-oriented programming.

2. Define OODB?

OODB-Object Oriented Database

A persistent and sharable collection of objects defined by an OODM

3. Define OODMS?

OODBMS- Object Oriented Database Management System

OODBMS-The Manager of OODB.③↘OO refers to abstract DB plus Inheritence & object identify.It is the Combination OO capability and DB capability.

4 . What are the types of OID?

They are 2 types of OID

Logical OID

Physical OID

5. Define pointer swizzling or object faulting?

To achieve the required performance, the OOBMS must be able to convert OID to and from in memory pointer. This conversion technique is known pointer swizzing or object faulting.

6. What is the aim of pointer swizzling ?

The aim of pointer swizzling is to optimize access to objects. As we have just mentioned, reference between objects are normally represented using OIDs.

7. List the classification of pointer swizzling ?

Classification or technique for pointer swizzling:

Copy vs in place swizzling

Eager vs lazy swizzling

Direct vs indirect swizzling

8. Define persistent object ?

The object that exist even after the session is over is called Persistent object.

There is 2 types of objects

Persistent

Transient

9. Define transient object ?

The Transient object is defined as Lact only for the invocation of program..④★The Object 's memory is allocated and Deallocated by the programming language 's at the run -time system.

10. List the scheme for implementing persistence within OODBMS?

Persistent scheme

There are 3schemes for implementing persistence in OODBMS

Check pointing

Serialization

Explicit paging

11. List the two methods for creating or updating persistent objects using explicit paging?

Reachability based method

Allocation based method

12. What are the fundamental principles of orthogonal persistence ?

It is based on 3 fundamental principles

Persistence independence

Data type orthoganality

Transitive persistence

13. Define nested transaction model ?

A transaction is views as a collection of related subtransaction each of which may also containany number of subtransaction.

14. Define sagas ?

A sequence of flat transaction that can be interleaved with other transaction.

Sagas is based on the use of Compensative transaction.

DBMS guarantees that either all the transaction in a Sagas are Sucessfully completed or compensative Transaction are run to recover from partial exection.

15. How the Concurrency Control is implemented in OODBMS?

Concurrency control protocol is used in Multiversion control protocol.

Hence,by using this the concurrency is implemented.

16.List the basic architecture for client server DBMS?

3 basic architecture for client server DBMS is

Object Server

Page Server

Database Server

17. Define POSTGRES?

POSTGRES is the reaserch system designers of INGRES that attempts to extend the relational mode with abstract datatype procedure and rules.

18.What is a GEMSTONE?

Gemstone is a product which extend an existing object-oriented programming language with database capability.

It extend 3 Languague such as Smalltalk, C++ or Java.

19.What is OQL?

OQL –Object Query Languague

An OQL is a function that delivers an object whose type may be infrrred from the operator contributing to the query expression.OQL is Used for both associative and navigational access.

20. Advantage and Disadvantage of OODBMS?

Adv:

Enriched modeling capabilities

Extensibility

Removal of impedance mismatch.④ ★Improved performance

Disadv:

Lack of Universal Data model

Lack of Experience

Lack of standards

Complexity

21. What are the phases of two phase locking protocol.
a) Growing phase - A transaction may obtain locks, but may not release any
lock.
b) Shrinking phase - A transaction may release locks, but may not obtain any
new locks.
22. What is a lock point?
	The point in the schedule where the transaction has obtained its final lock (the end
of its growing phase) is called the lock point.
23. What are the fields that a update log record contain?
	Transaction identifier, Data-item identifier, Old value, New value
24. Write the sequence of actions when a checkpoint record is added?
1. Output on to stable storage all log records currently residing in main memory.
2. Output to the disk all modified buffer blocks.
3. Output onto stable storage a log record <check point>.16. What is LSN?
Each log record in ARIES has a log sequence number (LSN) that uniquely
Identifies the record.
24. What happen when the analysis pass of the ARIES recovery algorithm?
	This pass determines which transactions to undo, which pages were dirty at the
time of crash, and the LSN from which the redo pass should start.
25. What is a lock table?
	The lock manager uses this data structure; for each data item that is currently
locked, it maintains a linked list of records, one for each request, in the order in
which the requests arrived. It uses a hash table, indexed on the name of a data
item, to find the linked list (if any) for a data item, this table is called lock table.
26. Define deadlock
	A system is in a deadlock state if there exists a set of transactions such that every
transaction in the set is waiting for another transaction in the set.
27. What is starvation?
	Transaction gets rolled back repeatedly and it never allowed to make progress.
This is called starvation.
28. What are the different parallel database architectures?
	a) shared memory b) shared disk c) shared nothing d) Hierarchical
29. What is a local transaction?
	A local transaction is one that access data only form sites where the transaction
was initiated.
30. What is a global transaction?
	A global transaction is one that either accesses data in a site different from the one
at which the transaction was initiated, or accesses data in several different sites.
31. What are the different forms of data transparency?
a) Fragmentation transparency
b) Replication transparency.c) Location transparency.
32. In distributed transactions, name the basic system failure types.
a) Failure of a site
b) Loss of messages
c) Failure of a communication link
d) Network partition.

33. What is the difference between homogeneous and heterogeneous distributed databases?

In a homogeneous distributed databases, all sites have identical database management system software.

In a heterogeneous distributed database, different sites may use different schemas, and different database management system software.

34. What is data transparency?

The user of a distributed database system should not be required to know either where the data are physically located or how the data can be accessed at the specific local site. This characteristic called data transparency.

35. What are global transactions?

The global transactions are those that access and update data in several local databases.

36. What is the function of transaction manager?

The transaction manager manages the execution of those transactions (or sub transactions) that access data stored in a local site.

37. What is the function of transaction coordinator?

The transaction coordinator coordinates the execution of the various transactions (both local and global) initiated at that site.

38. What are the advantages of single lock-manager approach in distributed databases?

a) Simple implementation
b) Simple deadlock handling

39. What is read quorum Qr and write quorum Qw in Quorum Consensus protocol?

The quorum consensus protocol assigns each site a nonnegative weight. It assigns read and write operations on an item x two integers, called read quorum Q_r and write quorum Q_w, and that must satisfy the following condition, where S is the total weight of all sites at which x resides.

$Q_r + Q_w > S$ and $2 * Q_w > S$

40. Define robustness.

The ability to continue functioning even during failures is referred to as robustness.

41. What is the difference between backup coordinator and coordinator?

The only difference in function between the coordinator and its backup is that the backup does not take any action that affects other sites. Such actions are left to the actual coordinator.

42. What is the function of election algorithm?

Election algorithms enable the sites to choose the site for the new coordinator in a decentralized manner. Election algorithms require that a unique identification number be associated with each active site in the system.

43. What are the basic data partitioning strategies?

a) Round-robin
b) Hash portioning
c) Range portioning

44. What are range queries?

Locating all tuples for which the value of a given attribute lies within a specified

range (for exampple, 10000 <salary<20000): these queries are called range queries.

45. What is interquery parallelism?

In interquery parallelism, different queries or transactions execute in parallel with one another. Transaction throughput can be increased by this form of parallelism.

46. What is intraquery parallelism?

Intraquery parallelism refers to the execution of a single query in parallel on multiple processors and disks.

47. What are the two forms of interoperation parallelism?

a) Pipelined parallelism.b) Independent parallelism

48. Define Object-Oriented Database Systems.

Object-Oriented database systems are proposed as an alternative to relational systems and are aimed at application domains where complex objects play a central role. The approach is heavily influenced by object-oriented programming languages and can be understood as an attempt to add DBMS functionality to a programming language environment.

49. Define Object-Relational Database Systems.

Object-Relational database systems can be thought of as an attempt to extend relational database systems with the functionality necessary to support a broader class of applications and, in many ways, provide a bridge between the relational and object-oriented paradigms.

50. Define persistent programming language?

A persistent programming language is a programming language extended with constructs to handle persistent data.

UNIT III

1.Define Data Mining.

The process of extracting valid, previously unknown comprehensible and actionable information from large databases and using it to make crucial business decisions.

2.List the different steps in data mining.

→Data cleaning

→Data integration

→Data selection

→Data transformation

→Data mining

→Pattern evaluation

→Knowledge presentation

3.Define Classification.

It is used to establish a specific, predetermined class for each record in a database from a finite set of possible class values.

4. Define Clustering.

Clustering can be considered the most important *unsupervised learning* problem A *cluster* is therefore a collection of objects which are "similar " between them and are "dissimilar " to the objects belonging to othe r clusters.

5.Define data warehousing.

A subject oriented, integrated, time variant and non volatile collection of data in support of the management 's decision making process.

6.Define web database.

A database that is used for web applications that use an architecture called three tier architecture. It has web browser,web server, database server.

7.Define mobile database.

A database that is portable and physically separate from a centralized database server but is capable of communicating with that server from remote sites allowing the sharing of corporate data.

8.Define upflow.

Upflow means adding value to the data in the datawarehouse through summarizing, packaging and distribution of data.

9.Define downflow.

Downflow means archiving and backing up the data in the warehouse.

10.What are the different groups of end user access tools?

→Reporting and query tools.

→Application development tools.

→Executive information system tools.

→Online analytical processing tools.

→Data mining tools.

11.What are the four main operations associated with data mining techniques.

1. Predictive modeling.

2. Database segmentation.

3. Link analysis.

4. Deviation detection.

12.Define outliers.

Outliers which express deviation from some previously known expectations and norms.

13.List the benefits of data warehousing.

1. Potential high returns on investment.
2. Competitive advantage.
3. Increased productivity of corporate decision makers.

14.Define XML.

The basic object is XML in the XML document.Two main structuring concepts are used to construct an XML document:elements and attributes.Attributes in XML provide additional information that describes elements.

15.What are the uses of DTD?

DTD give an overview of XML schema. It specifies the elements and their nested structures.

16.Define data mart.

Data marts generally are targeted to a subset of the organization, such as a department and are more tightly focused.

17.Define client/server model.

Client server model is a two-tier architecture. It consists of 2 tiers namely client and server. Here the client performs presentation service and the server performs data service. The client is called fat-client because client require more resources.

18.List the use of data mining tools.

→Data preparation.
→Selection of data mining operation.
→To provide scalability and improve performance.
→Facilities for visualization of result.

19.Define OLAP.

OLAP is a term used to describe the analysis of complex data from the datawarehouse.OLAP tools use distributed computing capabilities for analysis that require more storage and processing power.

20.List the problems of data warehousing.

→Project management is an important and challenging consideration that should not be underestimated.
→Administration of a data warehouse is an intensive enterprise, Proportional to the size and complexity of the data warehouse.

21.List some examples of data mining application.

→Marketing.
→Finance.
→Manufacturing.
→Health care.

UNIT IV

1.Define ER model?

The entity-relationship model (or ER model) is a top down approach to database design that begins by identifying the important data called entities and relationship between the data. The ER model was first proposed by Peter Pin-Shan Chen.

2.Define Entity type?

A group of object with same property which are identify by the enterprise as having an independence existence. In an ER model, we diagram an entity type as a rectangle containing the type name, such as *student*

3.Define Entity occurrence?

A uniquely identifiable object of an entity type is known as entity occurrence. Entity occurrence is similar to entity.

4.Define relationship type?

A relationship type is a set of meaningful associations among entity types. For example, the *student* entity type is related to the *team* entity type because each student is a member of a team.

5.Define relationship occurrence?

A uniquely identifiable association which includes one occurrence from each participating entity type.

6.Define degree of relationship?

Student MemberOf Team.The degree of a relationship type is the number of entity types that participate. If two entity types participate, the relationship type is binary. A role name indicates the purpose of an entity in a relationship.

7.Define recursive relationship with diagrammatic representation?

A recursive relationship is one in which the same entity participates more than once in the relationship. The *supervision* relationship is a recursive relationship because the same entity, a particular team, participates more than once in the relationship, as a supervisor and as a supervisee.

8.What are the types of attribute?

The types of attributes are

1. Simple and composite attribute

2. Single-valued and multi-valued attribute

Simple and composite attribute

Attributes that can 't be divided into subparts are called Simple or Atomic attribute**s.** The attribute composed of single component with independent system.

Ex: position and salary attribute of staff entity.

The attribute composed of multiple components each with an independent existence. Composite attributes can be divided into smaller subparts.For example, take Name attributes. We can divide it into sub-parts like First name,Middle name, and Last-named.

Single-valued and multi-valued attribute

Attributes that can have single value at a particular instance of time are called single valued A person can 't have more than one age value.Therefo re, age of a person is a single-values attribute. A multi-valued attribute can have more than one value at one time.

For example, degree of a person is a multi-valued attribute since a person can have more than one degree.

9.Define candidate key?

Minimal set of attributes that uniquely identifies each occurrence of an entity

type is known as primary key. For example: Branch number attribute is the candidate key for branch entity type.

10. Define primary key?

The candidate key that is selected to uniquely identify each occurrence of an entity type is called primary key. Primary keys may consist of a single attribute or multiple attributes in combination.

11. Differentiate strong and weak entity type?

An entity type that is not existence dependent on some other entity type called strong entity type. For example, the entity type *student* is **strong** because its existence does not depend on some other entity type.

An entity type that is existence dependent on some other entity type is called weak entity type. For example, a child entity is a weak entity because it relies on the parent entity in order for it to exist.

12. Define query processing?

Query processing transforms the query written in high level languages into a correct and efficient execution strategy expressed in a low level language ant to execute the strategy to retrieve the required data.

13. Define query optimization?

Query optimization means converting a query into an equivalent form which is more efficient to execute. It is necessary for high level relation queries and it provides an opportunity to DBMS to systematically evaluate alternative query execution strategies and to choose an optimal strategy.

14. What are the phases of query processing?

The phases are
1) Query Decomposition.
2) Query Optimization.
3) Code Generation.
4) Runtime Query Execution.

15. Define query decomposition and what are its stages?

The query decomposition is the firs phase of query processing whose aims are to transform a high-level query into a relational algebra query and to check whether that query is syntactically and semantically correct.

Different stages are
1) Analysis
2) Normalization
3) Semantic analysis
4) Simplification
5) Query restructuring.

16. Define conjunctive and disjunctive normal form?

Conjunctive normal form

Conjunctive normal form means sequence of conjuncts connected with an AND operator. These conjuncts contain one or more terms connected by OR operator.

Disjunctive normal form

Disjunctive normal form means sequence of disjuncts connected with an OR operator. These disjuncts contain one or more terms connected by

AND operator.

17.Differentiate Dynamic vs Static form optimization?

Dynamic optimization

Query has to be passed, validated and optimized before it can be executed.

All information required to select an optimum strategy is up to date.

Static optimization

Query is passed, validated and optimized only once.

Runtime overhead is reduced.

18.What are the problems caused by concurrency control.

The process of managing simultaneous operations on the database without having them to interfere with one another is called as concurrency control. The problems caused by concurrency control are

i. Lost update problem

ii. Uncommitted dependency problem.iii. Inconsistent analysis problem

19.Define 3NF and BCNF

Third Normal Form (3NF):

A relation that is in 1NF and 2NF, and in which no non-primary key attribute is transitively dependent on the primary key.

Boyce-Codd Normal Form (BCNF):

A relation is in BCNF, if and only if, every determinant is a candidate key.

20. Define Timestamp?

Timestamp is a unique identifier created by the DBMS that indicates the relative starting time of a transaction. Time stamping is a concurrency control protocol that orders transaction in such a way that older transaction with smaller imestamp will get priority in the event of conflict.

21.What are the properties of transaction?

The four basic properties of transactions are called as ACID properties.

A - atomicity

C - consistency

I - isolation

D - durability

ATOMICITY:

The all or nothing property. A transaction is an indivisible unit that is either performed in its entirely or not performed at all.

CONSISTENCY:

A trasaction must transform the database from one consistent state to another consisient state.

ISOLATION:

Transactions execute independently on one another. In other words, the partial effects of incomplete transaction should not be visible to other transactions.

DURABILITY:

The effects of successfuly completed transaction are permenently recorded in the db and must not be lost because of subsequent failure.

22.Define concurrency control?

The process of managing simultaneous operations on the db without having them interface with each other.

23.What are the problems caused by concrrency control?

The problems caused by concurrency control are
1. Lost update problem,
2. Uncommited dependency problem,
3. Inconsistent analysis problem.

LOST UPDATE:

An apparentlty successfully completed update operation by one user can be overriden by another user. This is known as the lost update problem.

UNCOMMITED DEPENDNCY:

An uncommited dependency problem occurs when one transaction is allowed to see the intermediate results of another transaction before it has committed.

INCONSISTENT ANALYSIS:

A problem of inconsistent analysis occurs when a transaction reads several values from the db but a second transaction updates some of them during the execution of the first.

24.Define serial schedule?

A schedule where the operations of each transaction are executed consequently without any interleaved operations from other transactions.

25.Define serializable?

If a set of trasaction execute concurrently , we say that the schedule(nonserial) is correct if it produces the same results as some serial execution. Such a schedule is called serializable.

26.Define the conservative and optimistic concurrency control methods?

CONSERVATIVE METHOD:

This approach causes the transaction to be delayed in case they conflict with other transaction at some time in future. Locking and timestamping are essentially conservative approaches.

OPTIMISTIC METHOD:.

This approach is based on the premise that conflict is rare so they allow transaction to proceed unsynchronized and only check for conflicts at the end, when a transaction commits.

27.Define shared and exclusive lock?

SHARED LOCK:A transaction has a shared lock on a data item it can only read the item but cant update.

EXCUSIVE LOCK:A transaction has exclusive lock on a data item it cannot both read and update the data.

28.Define 2PC?

A transaction follows two phase locking protocol if all locking operation precede the first unlock operation in the transaction.

29.Define ignore obsolete write rule?

The transaction T asks to write an item(x) whose value already been written by an younger transaction, that is ts(T)<write_timestamp(x). This means that a later trasaction has alreadyupdated the value of the item, and the value that the older transaction is writing must be based on an obsolete value of an item. In this case, the write operation can be safely ignored. This is sometimes knows as the ignore obsolete write rule, and allows greater consistency.

30.List out different db recovery facilities?

A DBMS should provide the following facilities to assist with recovery.

1. A backup mechanism, which makes periodic backup copies of the db.

2. Logging facilities, which keep track of the current state of transactions and db changes.

3. A checkpoint facility, which enables updates to the db that are in progress be made permenent.

4. a recovery manager, which allows the system to restore the db to an consistent state following a failure.

31.What is the need for db tuning?

The need for tuning a db are,

1. Existing tables may be joinned.

2. For a given set of tables, ther may be an alternative design choice.

32.Define normalization?

Normalization is a bottom up approach to a db design that begins by examining the relationship between attributes. It is a validation technique. It supports a db designer by presenting a series of tests, which can be applied to individual relations so that the relational schema can be normalized to a specific form to prevent possible occurence of update anomalies.

33.What is flattening the table?

We remove the repeated groups by entering the appropiate data in the empty columns of rows containing the repeated data. In other words we fill in the blanks by duplicating the nonrepeating data where required. This approach is called as flattening the table.

Unit-V

1.Define deductive database.

A deductive database includes capabilities to define (deductive) rules, which can deduce or infer additional information from the facts that are stored in a database. Because part of the theoretical foundation for some deductive database systems is mathematical logic, such rules are often referred to as logic databases.

2.Define spatial database.

Spatial databases provide concepts for databases that keep track of objects in a multi dimensional space.

3.Define multimedia database.

Multimedia provide features that allow users to store and query different types of multimedia information, which includes images (such as photos or drawing), videoclips (such as movies, newsreels, or home videos), audio clips (such as songs, phone messages, or speeches), and documents (such as books or articles).

4.List the different spatial query language.

The different spatial query languages are
1. Range query
2. nearest neighbor query
3. Spatial joins or overlays.

5. Define inference engine.

An inference engine (or deductive mechanism) within the system can deduce new facts from the database by interpreting these rules. The model used for deductive databases is closely related to the relational data model, and particularly to the domain relational calculus formalism. It is related to the field of logic programming and the prolog language.

6.Example for spatial database.

Example for spatial database is cartographic databases that store maps include two dimensional spatial descriptions of their objects - from countries and states to rivers, cities, roads, seas and so on. These applications are also knowns as Geographical Information Systems(GIS), and are used in areas such as environmental, emergency, and battle management. Other databases, such as meterological databases for weather information, are three dimensional, - since temperatures and other Meterological information are related to three dimensional spatial points.

7. Define active database.

Active databases which provide additional functionality for specifyingactive rules. These rules can be automatically triggered by events that occur, such as   database updates or certain times being reached, and can initiate certain actions that have   been specified in the rule declaration to occur if certain conditions are met.

8. Example for multimedia database.

For example, one may want to locate all video clips in a video database that include a certain person, say Bill Clinton. One may also want to retrieve video clips based on certain activities included in them, such as a video clips where a soccer goal is scored by a certain player or team.

9. Define Quad trees.

Quad trees generally divide each space or subspace into equally sized areas, and proceed with the subdivisions of each subspace to identify the positions of various objects.

10. What are the two main methods of defining the truth values of predicates in actual

datalog programs?

There are two main methods of defining the truth values of predicates in actual datalog programs that is,

1.Fact-defined predicates (or relations)

2. Rule-defined predicates (or views).

11. What is Fact-defined predicates?

Fact-defined predicates (or relations) are defined by listing all the combinations of values (the tuples) that make the predicate true. These corresponds to base relations whose contents are stored in a database system.

12. What is Rule-defined predicates?

Rule-defined predicates (or views) are defined by being the head of one or more Datalog rules; they correspond to virtual relations whose contents can be inferred by the inference engine.

13. What is the use of relational operations?

It is straightforward to specify many operations of the relational algebra in the form of Datalog rules that define the result of applying these operations on the database relations (fact predicates). This means that relational queries and views can easily be specified in Datalog.

14. What are the characteristics of Nature of Multimedia Applications?

Applications may be categorized based on their data management characteristics as follows:

1. Repository applications

2. Presentation applications

3. Collaborative work using multimedia information.

15. What are the terms included in multimedia information systems?

Multimedia Information Systems are complex, and embrace a large set of issues, including the following:

1. Modeling

2. Design

3. Storage

4. Queries and retrieval

5. Performance

16. What are the different characteristics of Hypermedia links or hyperlinks?

1. Links can be specified with or without associated information, and they may have large descriptions associated with them.

2. Links can start from a specific point within a node or from the whole node..3. Links can be directional or nondirectional when they can be traversed in either direction.

17. What are the applications of multimedia database?

1. Documents and records management

2. Knowledge dissemination

3. Education and training

4. Marketing, advertising, retailing, entertainment, and travel

5. Real-time and monitoring.

18. What are the three main possibilities for rule consideration?

There are the three main possibilities for rule consideration:

        1. Immediate consideration
        2. Deferred consideration
        3. Detached consideration

19. What is Horn Clauses?

        In Datalog, rules are expressed as a restricted form of clauses called Horn Clauses, in which a clause can contain at most one positive literal.

20. What are the two alternatives for interpreting the theoretical meaning of rules?

        There are two main alternatives for interpreting the theoretical meaning of rules:

1. Proof-theoretic 2. Model-theoretic