UNIT-1

2 Marks

1. Define Distributed Database.
   A logically interrelated collection of shared data (and a description of this data) physically distributed over a computer network.

2. Define Distributed DBMS.
   The software system that permits the management of the distributed database and makes the distribution transparent to users.

3. Define Distributed processing.
   A centralized database that can be accessed over a computer network.

4. Define Parallel DBMS.
   A DBMS running across multiple processors and disks that is designed to execute operations in parallel, whenever possible, in order to improve performance.

5. What are the main architectures for parallel DBMS?
   - Shared Memory
   - Shared Disk
   - Shared nothing

6. What is Homogeneous and Heterogeneous DDBMS?
   In Homogeneous system, all sites use the same DBMS product. In Heterogeneous system, sites may run different DBMS products.

   Global Conceptual Schema (GCS) is a logical description of the whole database, as if it were not distributed. It contains definitions of entities, relationships, constraints, security and integrity information. It provides physical data independence from the distributed environment.

8. What are the types of Fragmentation?
   The two types of fragmentation are
   1. Horizontal fragmentation: subset of tuples.
   2. Vertical fragmentation: subset of attributes.

9. What are the objectives of definition and allocation of fragments?
   The main objectives are:
Locality of reference.
Improved reliability and availability
Acceptable performance
Balanced storage capacities and cost
Minimal communication costs.

10. What are the four alternative strategies regarding placement of data?
The four alternative strategies regarding placement of data are:
   - Centralized
   - Fragmented (partitioned)
   - Complete replication
   - Selective replication

11. What are the rules that have to be followed during fragmentation?
The three correctness rules are:
   - Completeness: ensure no loss of data
   - Reconstruction: ensure Functional dependency preservation
   - Disjointness: ensure minimal data redundancy.

12. What are the objectives of concurrency control?
   - To be resistant to site and communication failure.
   - To permit parallelism to satisfy performance requirements.
   - To place few constraints on the structure of atomic actions.

13. What is multiple-copy consistency problem?
   It is a problem that occurs when there is more than one copy of data item in different locations and when the changes are made only in some copies not in all copies.

14. What are the two approaches in distributed environment?
   - Locking
   - Timestamp
   Locking guarantees that the concurrent execution is equivalent to some unpredictable serial execution of those transactions.
   Timestamp guarantees that the concurrent execution is equivalent to specific serial execution of those transactions, corresponding to the order of timestamps.

15. What are the types of Locking protocols?
   - Centralized 2 phase locking
   - Primary 2 phase locking
   - Distributed 2 phase locking
   - Majority locking

16. What are the failures in distributed DBMS?
   - The loss of message
   - The failure of a communication link
   - The failure of a site
Network partitioning

17. Define coordinator and participant.
   Every global transaction has one site that acts as the coordinator (or transaction manager) for that transaction, which is generally the site at which the transaction was initiated. Sites at which the global transaction has agents are called participants (or resource managers).

18. Define unilateral abort.
   If a participant votes to abort, then it is free to abort the transaction immediately; in fact any site is free to abort a transaction at any time up until it votes to commit. This type of abort is known as unilateral abort.

19. Give the states of the coordinator
   The four states of the coordinator are:
   - INITIAL
   - WAITING
   - DECIDED
   - COMPLETED

   Pessimistic protocols choose consistency of the database over availability and hence do not allow transactions to execute in a partition if there is no guarantee that consistency can be maintained. The protocol uses pessimistic control algorithm.

21. What is replication?
   The process of generating and reproducing multiples copies of data at one or more sites is called replication.

22. What are synchronous and asynchronous replications?
   In synchronous replication the replicated data is updated immediately when the source data is updated. This is done by using 2PC protocol. In asynchronous replication the target data is updated after the source database has been modified. The delay in regaining consistency may range from a few seconds to several hours or even days.

23. What is data ownership? What are the types of ownership?
   Data ownership relates to which site has the privilege to update the data. The main types of ownership are master/slave, workflow, and update-anywhere.

24. What are the implementation issues in data replication?
   The main implementation issues are
   - Transactional updates
   - Snapshots and database triggers
   - Conflict detection and resolution

25. What are the mechanisms proposed for conflict resolution?
The mechanisms proposed for conflict resolution are
  o Earliest and latest timestamps
  o Site priority
  o Additive and average updates
  o Minimum and maximum values
  o User defined
  o Hold for manual resolution

26. What is generic relational algebra tree?
   The relational algebra tree formed by applying the reconstruction algorithm is known as the is generic relational algebra tree.

27. What are the Benefits of replication?
   Replication provides a number of benefits, including improved performance, increased reliability & data availability, and support for mobile computing & data warehousing.

28. What are the types of replication?
   Read-only snapshots
   Updateable snapshots
   Multimaster replication
   procedural replication

29. Define Snapshot logs.
   A snapshot log is a table that keeps track of changes to a master table. A snapshot log can be created using the CREATE SNAPSHOT LOG statement.

30. Define database links.
   Database links define a communication path from one Oracle database to another database.

16 Marks

1. Explain Distributed Relational Database design in detail.
   • Objectives of data allocation and fragmentation
   • Data Allocation:
     ▪ Strategic Objectives: Centralized, Fragmented, Complete Replication, Selective Replication.
   • Fragmentation:
     ▪ Need for fragmentation
     ▪ Correctness of fragmentation
     ▪ Types of fragmentation

2. Explain about Transparencies in a DDBMS.
   Distribution transparency
3. Explain the Functions and Architecture of a DDBMS.

Functions of DDBMS
- Extended communication services
- Extended system catalog
- Distributed Query processing
- Extended security control
- Extended concurrency control
- Extended recovery services

Reference Architecture for DDBMS
- Global conceptual schema
- Fragmentation and Allocation Schemas
- Local Schemas
  - Refer FIGURE: 22.4 / Pg no. 702

Reference Architecture for Federated MDBS
- Refer FIGURE: 22.5 / Pg no. 703

Component Architecture for DDBMS
- Local DBMS component (LDBMS)
- Data Communication component (DC)
- Global System Catalog (GSC)
- Distributed DBMS component (DDBMS)
  - Refer FIGURE: 22.6 / Pg no. 704

4. Explain in detail the recovery techniques in distributed databases.

Failures in a distributed environment
How failures affect recovery
Distributed recover protocols
Two-phase commit (2PC)
Termination protocols
Recovery protocols
Election protocols
Communication topologies for 2PC
Three-phase commit (3PC)
Network partitioning

5. Discuss about the X/Open Distributed Transaction Processing model.
• X/Open established the Distributed Transaction Processing (DTP) working group with the objective of fostering appropriate programming interface for Transaction Processing.
• X/Open interfaces with diagram
  • Transaction Manager
  • Resource Manager
  • Procedures of TX interface
• X/Open interfaces in a distributed environment

6. Explain about distributed query optimization.
   Query optimization.
   Techniques in query optimization
   Distributed query transformations
   Reconstruction Algorithms
   Generic relational algebra tree
   Reduction techniques for various types of fragmentation
   Distributed joins.

7. Explain Distributed Deadlock Management.
   What is Deadlock?
   What is Deadlock Management?
   Deadlock Detection.
   Diagram for Distributed Deadlock.
   Types of Deadlock Detection.
   Centralized
   Hierarchical
   Distributed
   Explain with Fig. 23.4 & fig.23.5 in Pg no. 739 & 740

UNIT 2

2 marks

1. List the applications of Advanced Database?
   Some of the applications are:
   - Computer-aided Design
   - Computer-aided Manufacturing
   - Computer-aided Engineering
   - Network Management Systems
   - Office Information Systems
   - Digital Publishing
   - Geographic Information System
   - Interactive and Dynamic Websites

2. Define Transitive closure.
   The Transitive closure of a relation R with attributes \((A_1, A_2)\) defined on the same domain is the relation R augmented with all tuples successively deduced by
transitivity; that is, if \((a, b)\) and \((b, c)\) are tuples of \(R\), the tuple \((a, c)\) is also added to the result.

3. Define Object?
   A uniquely identifiable entity that contains both the attributes that describe the state of a ‘real world’ object and the actions that are associated with it.

4. Define Inheritance, Generalization and Specialization?
   Inheritance allows one class to be defined as a special case of more general cases. These special cases are known as subclasses and the more general cases are known as super classes. Process of forming a super class is known as Generalization and the process of forming the subclass is specialization.

5. Define OODM, OODB, OODBMS.
   - OODM: A logical data model that captures the semantics of objects supported in object-oriented programming.
   - OODB: A persistent and sharable collection of objects defined by an OODM.
   - OODBMS: The manager of an OODB.

   The action of converting object identifiers to main memory pointers and back again. Its aim is to optimize access to objects. It can be classified as
   - Copy versus in-place swizzling
   - Eager versus lazy swizzling
   - Direct versus indirect swizzling

7. What is meant by Version Management?
   The Process of maintaining the evolution of objects is known as Version Management. An Object version represents an identifiable state of an object.

8. What are the mandatory features proposed by Object Oriented Database System Manifesto that apply to object-oriented characteristic?
   The following rules apply to OO characteristic:
   - Complex objects must be supported
   - Object identity must be supported.
   - Encapsulation must be supported.
   - Types or classes must be supported.
   - Types or classes must be able to inherit from their ancestors.
   - Dynamic binding must be supported.
   - The DML can be computationally complete.
   - The set of data types must be extensible.

9. Define the object oriented request broker (ORB).
   The ORB handles distribution of messages between application objects in a highly inter-operable manner. In effect, the ORB is a distributed ‘software bus’ that enables objects to make and receive requests and responses from a provider.
10. What is meant by Postgres?

Postgres, an early Object-Relational DBMS, is a research system from the designers of INGRES that attempts to extend the relational model with abstract data types, procedures and rules.

11. What are the objectives of Postgres?

   The objectives are:
   1. to provide better support for complex object.
   2. to provide user extensibility for job data types, operators and access methods.
   3. to provide active database facilities (alerters and triggers) and inferencing support.
   4. to simplify the DBMS code for crash recovery.
   5. to produce a design that can take advantage of optical disks, multiple-processor work-stations and custom-designed VLSI chips.
   5. to make as few changes as possible to the relational model.

12. How the declaration of a relation is being made in Postgres?

   A relation in Postgres is declared using the following command.
   
   CREATE TableName (columnName 1 = type 1, columnName 1 = type 2, ....)
   [KEY (list of ColumnNames)]
   [INHERITS (list of TableNames)]

13. Define CORBA.

   The Common Object Request Broker Architecture (CORBA) defines the architecture of ORB-based environments. This architecture is the basis of any OMG component, defining the parts that form the ORB and its associated structures.


   Reachability-based persistence means that an object will persist if it is reachable from a persistent root object.
   Allocation-based persistence means that an object is made persistent only if it is explicitly declared as such within the application program.

15. What are the major components of ODMG architecture for an OODBMS?

   1. Object model (OM)
   2. Object definition language (ODL)
   3. Object query language (OQL)
   4. C++, Java and Smalltalk language bindings

16. What are the features which are supported by ODMG?

   - Concurrent execution
   - Distributed transaction
   - Versioning
   - Event notification
   - Internationalization
17. Explain about Wisconsin benchmark?
It is an earliest DBMS benchmark. It was developed to allow comparison of particular dbms. It consists of a set of tests as a single user covering:
- Updates and deletes involving both key and non key attributes.
- Projectons involving different degrees of duplication in the attribute and selections with different selectives on indexed, on-index, and clustered attributes.
- Joins with different selectives;
- Aggregate functions.

18. What is meant by Interface Repository?
An interface repository that stores persistent IDL (Interface Definition Language) definitions. The interface repository can be queried by a client application to obtain a description of all the registered object interface, the methods they support, the parameters they require, and the exceptions that may arise.

Repeated Inheritance: It is a special case of multiple inheritances, in which the super classes inherit from a common super class.
Selective Inheritance: It allows a subclass to inherit a limited number of properties from the super class.

20. Mention the characteristics of objects?
- structure
- identifier
- name
- lifetime

16 marks:

1.a) Explain the weakness of RDBMS?
Hints:
- Poor representation of real world entities.
- Semantic Overloading
- Poor support for integrity and enterprise constraints
- Homogenous data structure
- Limited operations
- Difficulty in handling recursive query
- Impedance Mismatch
- Other problems with RDBMS

1.b) Compare ORDBMS and OODBMS.
Hints:
The comparison is made on three perspectives
Data modelling
2. Explain Object oriented Concepts in database and storing objects in Relational Database.
Hints:
Concepts:
- Abstraction, Encapsulation, Information Hiding.
- Objects and Attributes
- Object Identity
- Methods and Messages
- Classes
- Subclasses, Superclasses, Inheritance
- Overriding and Overloading
- Polymorphism and Dynamic Binding
- Complex Objects

Storing objects in Relational Database:
- Mapping Classes to Relations
- Accessing Objects in the Relational Database
- Explain with e.g.

3. Explain about issues in OODBMS.
Hints:
- 3 problematic areas for RDBMS
  a) Long duration transaction
  b) Versions
    - transient versions
    - Working versions
    - Released version
  c) Schema Evolution.
    i) typical changes to the schema.
    ii) specify all the rules. Of four different groups
      - the resolution of conflicts caused by multiple inheritance and the redefinition of attributes and methods in a subclass.
      - The propagation of modifications to subclasses
      - The aggregation and deletion of inheritance relationships between classes and creation and removal of classes.
      - Handling of composite objects.

Architecture:
1. Client-Server:
   - Object server
   - Page server
   - Database server
2. Storing and executing methods:
   Benchmarking:
   - Wisconsin benchmark
   - TPC-A and TCP-B
   - TPC-C
   - Other benchmark
   - OO1 benchmark
   - OO7 benchmark

4. Explain briefly on Object Oriented DBMSs- Standards and Systems.
   - Object Management Group
     The common object request broker architecture
   - Object Data standards ODMG 3.0, 1999
     Object Data management group
     The Object model
     The Object Definition Language
     The Object Query Language

5. Explain the advantages and disadvantages of OODBMS.
   Advantages:
   - Enriched modeling capabilities
   - Extensibility
   - Removal of impedance mismatch
   - More expensive query language
   - Support for schema evolution
   - Support for long-duration transactions
   - Applicability to advanced db applns.
   - Improved performance
   Disadvantages:
   - Lack of universal data model
   - Lack of experience
   - Lack of stds
   - Competition
   - Query optimization compromises encapsulation
   - Locking at object level may impact performance
   - Complexity
   - Lack of support for views
   - Lack of support for security

UNIT 3

2 Marks

1. What is meant by Internet, Intranets, and Extranets?
   - Internet
A world wide collection of interconnected computer networks.

- **Intranet**
  A website or group of sites belonging to an organization accessible only by the members of the organization.

- **Extranet**
  An Intranet that is partially accessible to authorized outsiders.

2. What is meant by E-Commerce and E-Business?
   - **E-Commerce**
     Customers can place and pay for orders via the business website.
   - **E-Business**
     Complete integration of internet technology into the economic infrastructure of the business.

3. What is meant by world wide web?
   A hypermedia based system that provide a means of browsing information on the internet in a non-sequential way using hyperlinks.

4. What is meant by HTTP and HTML?
   - **HTTP**
     The protocol used to transfer web pages through the internet.
   - **HTML**
     The document formatting language used to design most web pages.

5. What is meant by URL?
   A string of alphanumeric characters that represents the location or address of a resource on the internet and that resource should be accessed

6. Give any five requirements for Web-DBMS integration
   - The ability to access valuable corporate data in a secure manner.
   - Support for transaction that spans multiple HTTP requests
   - Support for session and application-based authentication
   - Acceptable performance
   - Minimal administration overhead

7. Give advantages of 3-tier over 2-tier design
   - The need for less expensive hardware because the client is ‘thin’.
   - Application maintenance is centralized with the transfer of the business logic for many end-users into a single application server. This eliminates the concerns of software distribution that are problematic in the traditional two-tier client-server model.
   - The added modularity makes it easier to modify or replace one tier without affecting the other tiers.
   - Load balancing is easier with the separation of the core business logic from the database functions.

8. Define Perl and PHP
Perl stands for Practical Extraction and Report Language and is a high-level interpreted programming language with extensive, easy-to-use text processing capabilities.

PHP stands for Hypertext Preprocessor and is an open source HTML-embedded scripting language that is supported by many Web servers including Apache HTTP server and Microsoft’s Internet Information Server, and is preferred Linux Web Scripting language.

9. Compare JavaScript and Java applets

<table>
<thead>
<tr>
<th></th>
<th>JavaScript</th>
<th>Java (applets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Interpreted by client</td>
<td>Compiled on server before execution on client</td>
</tr>
<tr>
<td>2.</td>
<td>Object-based. Code uses built-in, extensible objects, but no classes or inheritance</td>
<td>Object-oriented. Applets consist of object classes with inheritance</td>
</tr>
<tr>
<td>3.</td>
<td>Code integrated with, and embedded in, HTML.</td>
<td>Applets distinct from HTML</td>
</tr>
<tr>
<td>4.</td>
<td>Variable data types not declared</td>
<td>Variable data types must be declared</td>
</tr>
<tr>
<td>5.</td>
<td>Dynamic binding. Object references checked at runtime.</td>
<td>Static binding. Object references must exist at compile-time.</td>
</tr>
<tr>
<td>6.</td>
<td>Cannot automatically write to hard disk.</td>
<td>Cannot automatically write to hard disk.</td>
</tr>
</tbody>
</table>

10. Define CGI

CGI means Common Gateway Interface. It is a specification for transferring the information between a web server and a CGI program.

11. Define URL

URL stands for Uniform Resource Locator. The string of numeric character that represents the location or address of a resource on the Internet and how that resource should be accessed.

12. Define MIME

MIME stands for Multipurpose Internet Mail Extensions. MIME specification to allow the browser to differentiate between components. This allows a browser to display a graphics file, but to save a ZIP file to disk.

13. List any four advantages of CGI

- Simplicity
- Language independence
- Web server independence
- Wide acceptance

14. List any three weakness of CGI

- Bottle neck
15. Define cookies
   A cookie is a piece of information that the client stores on behalf of the server. The information stores in these cookies comes from the server as a part of the server’s response to an HTTP request. A client may have a number of cookies associated with a particular Web page or Web site. Each time the client that visits that site, the browser packages the cookie with the HTTP request. Cookies are used to make CGI more interactive.

16. Limitations of CGI
   Limitations are in performance and the handling of shared resources. The server needs to execute a gateway program, and communicate with it using some Inter-Process communication (IPC) mechanism. This request causes heavy burden on the server.

17. What do you mean by API?
   Application Programming Interface (API) adds functionality to the server or changes server behavior and customizes the heavy burden on the server. Such additions also called non-CGI gateways.

18. What are the two main types of API?
   * Netscape server API (NSAPI)
   * Microsoft’s Internet Information Server API (ISAPI)

19. Define iAS and the versions of iAS?
   Oracle Internet Application Server (iAS) is a reliable, scalable, secure, middle-tier application server that is designed to support e_business. It provides a set of services for assembling a complete, scalable middle-tier infrastructure.
   iAS is available in three versions:
   - Standard Edition
   - Enterprise Edition
   - Wireless Edition

20. Define the various HTTP server modules?
   In addition to the combined Apache HTTP server, Oracle has enhanced Oracle specific ones:
   They are
   - Mod_ssl
   - Mod_plsql
   - Mod_perl
   - Mod_jserv
   - Mod_ose

21. What are the various Business logic services?
   Business logic supports the application logic and include:
22. What are the various presentation Services?

These services deliver dynamic content to client browsers, supporting servlets, JSP, Perl/CGI Script, PL/SQL pages, forms and business intelligence:

- Apache Jserv
- OracleJSP
- Oracle PSP (PL/SQL Server pages)
- Perl Interpreter

23. Define Semistructured Data and its importance?

Data that may be irregular or incomplete and have a structure that may change rapidly or unpredictably.

Importance:

i. It may be desirable to treat web sources like a database, but we cannot constrain these sources with a schema;

24. Define J2ME, J2SE & J2EE

J2ME: Java 2 Platform, Micro edition. This is aimed at embedded & consumer electronics platform. This consists of only the API’s that are required for embedded application.

J2SE: Java 2 platform standard edition. This is aimed at typical desktop and workstation environment. This is equivalent to java 2 platform (JDK 1.2)

J2EE: Java 2 platform, Enterprise Edition. This is aimed at robust, scalable, multi-user and secure enterprise applications.

25. Name the EJB (Enterprise Java Beans) components

The EJB components are:

* EJB session Beans
* EJB Entity Beans

The entity Beans are of two types

Bean- Managed persistence (BMP) entity beans
Container- Managed persistence (CMP) Entity beans

26. Define SQLJ

A consortium of organization oracle, IBM & Jan dim specified proposed a specification for java with static embedded SQL and this is called SQLJ

27. Servlets

Servlets are programs that run on a java-enables web server and built web pages. These are java servelts. Their extensive features are-

- Improved Performance
- Portability
Extensibility
Simpler session management
Improved security & reliability

28. Java server pages
   JSP is a java based server side scripting language that allows static HTML to be mixed with dynamically generated HTML.

16 Marks

1. Discuss briefly about THE WEB?
   i). Definition of WWW.
   ii). Hyper Text Transfer Protocol
      - Definition of HTTP
      - Stages of HTTP
        - Connection
        - Request
        - Response
        - Close
      - Multi Purpose Internet Mail Extension
      - Definition of MIME
      - Use of MIME
      - Type
      - Subtype
      - Some useful MIME types
      - HTTP request
        - The main HTTP Request Types
          - GET
          - POST
          - HEAD
          - PUT(HTTP/1.1)
          - DELETE(HTTP/1.1)
          - OPTIONS(HTTP/1.1)
        - HTTP response
   iii). Hyper Text Mark Up Language
      - Definition of HTML
      - SGML
   iv). Uniform Resource Locator
      - Definition of URL
      - URN(Uniform Resource Name)
      - Syntax of URL
        - Protocol
        - The Host Name
        - The Path Name
   v). Static and Dynamic WEB Page
2. Web-DBMS Architecture
   i) Traditional two-tier client-server architecture
   ii) Three-tier architecture
   iii) Advantages
       - Simplicity
       - Platform independence
       - Graphical User Interface
       - Standardization
       - Cross-platform support
       - Transparent network access
       - Scalable deployment
       - Innovation
   iv) Disadvantages
       - Reliability
       - Security
       - Cost
       - Scalability
       - Limited functionality of HTML
       - Statelessness
       - Bandwidth
       - Performance
       - Immaturity of development tools

3. Explain about scripting languages, CGI and HTTP cookies:
   Javascript and Jscript:
   - VBScript
   - Perl and PHP
   CGI:
   Introduction
   - Passing information to a CGI script
   - Passing parameters on command line
   - Passing environment variables to CGI program
   - Passing data to CGI program via standard input
   - Using extra path information
   - Advantage
     - Simplicity
     - Language independence
     - Webserver independence
     - Wide acceptance
   - Disadvantage
     - Bottleneck
     - Lack of efficiency & transaction processing
     - Server has to generate new process or thread for each CGI
   HTTP Cookies:
4. Define Oracle internet Platform and explain briefly the various services provide by it?

Oracle Internet Platform:
Oracle’s web-centric approach different from Microsoft’s approach
Comprises Oracle Internet Application Server(iAS) and Oracle DBMS

N-tier architecture based on standards such as:
- HTTP and HTML/XML for web enablement
- The Object Management Group’s CORBA technology
- Internet Inter-Object Protocol (IIOP) for object interoperability and
- Remote Method Invocation (RMI)
- Java, enterprise javabeans (EJB), JDBC and SQLJ, java servlets and java server pages (JSP)

Oracle Internet Application Server (iAS):
It is a reliable, scalable, secure, middle-tier application server that is
designed to support e_business. It provides a set of services for assembling a
complete, scalable middle-tier infrastructure.
iAS is available in three versions:
- Standard Edition
- Enterprise Edition
- Wireless Edition

Communication Services:
- Handle all incoming requests received by iAS
- Some requests are processed by HTTP server and
  some are routed to other areas of iAS

HTTP server modules:
In addition to the combined Apache HTTP server, Oracle has enhanced
Oracle specific ones:
They are
- Mod_ssl
- Mod_plsql
- Mod_perl
- Mod_jserv
- Mod_ose

Business logic services:
Business logic support the application logic and include:
- Oracle BC4J (Business Components for Java)
- Oracle JVM
- Oracle PLSQL
- Oracle Forms

Presentation Services:
These services deliver dynamic content to client browsers, supporting
servlets, JSP, Perl/CGI Script, PL/SQL pages, forms and business intelligence:
- Apache Jserv
- Oracle JSP
- Oracle PSP (PL/SQL Server pages)
- Perl Interpreter

Caching Services:
Oracle Database cache:
   It is a middle-tier service that improves performance and scalability of applications by caching frequently used data
   Supports running stateful servlets, JSPs, EJBs and CORBA objects
Oracle web cache:
   Stores frequently accessed URLs in virtual memory

5. MICROSOFT’S WEB SOLUTION PLATFORM
Microsoft’s web solution platform
   * Previously called window’s DNA
   * Various services like BizTalk Server etc
   * Includes OLE, Com & DCOM
OLE
   * Use of DDE protocol earlier
   * OLE is a OO technology
COM
   Component Object model
   * Component objects
   * COM is a object based model
   * Establishes connection between client application & object & its services
   * Provides binary interoperability standard
DCOM
   Distributed component object model
   * DCOM is extension of COM

UNIVERSAL DATA ACCESS
   * Using ODBC
   * (DAO) Data Access Objects
fig OLE DB Architecture
Active server pages & ActiveX Data objects
   * ASP
   * ActiveX Data Objects (ADO)
   * Remote Data services
Web Page generation using MS Access
   * Using of static pages
   * Dynamic pages using Active server pages
   * Dynamic pages using Data Access Pages

UNIT 4

2 Marks

1. What are the components of a rule in ECA model?
   There are 3 components:
   1. The event that triggers the rule.
2. The condition that determines whether the rule action should be executed.
3. The action to be taken.

2. What is immediate consideration?
   The condition is evaluated as part of the same transaction as the triggering event and is evaluated immediately. The case can be further categorised into:
   - Evaluate the condition before executing the trigger event.
   - Evaluate the condition after executing the trigger event.
   - Evaluate the condition instead of executing the trigger event.

3. What is deferred consideration?
   The condition is evaluated at the end of the transaction that included the triggering event. In this case, there could be many triggered rules waiting to have their conditions evaluated.

4. What is deactivated rule and active command?
   A deactivated rule will not be triggered by the triggering event. This allows users to selectively deactivate rules for certain periods of time when they are not needed.
   The active command will make the rule active again.

5. What is row-level trigger and statement-level trigger?
   If the rule is triggered separately for each tuple, then it is known as a row level trigger.
   If the rule is triggered only once for each triggering statement, then it is known as statement level trigger.

6. What is meant by bitemporal database?
   There are two time dimensions namely valid time dimension and transaction time dimension. In some applications, only one of the dimensions is needed and in other cases both the time dimensions are required, in which case the temporal database is called a bitemporal database.

7. What is proactive, retroactive and simultaneous update?
   If the update was applied to the database before it becomes effective in the real world, it is proactive update. If the update was applied to the database after it becomes effective in the real world, it is retroactive update. An update that is applied at the same time when it becomes effective is called a simultaneous update.

8. What is attribute versioning?
   An alternative approach can be used in database systems that support complex structured objects, such as object databases or object relational systems. This approach is called attribute versioning.

9. What is coalescing?
   If the two time periods [T1, T2] and [T3, T4] are adjacent, they are combined into a
single time period $[T_1, T_4]$. This is called coalescing of time periods. Coalescing also combines intersecting time periods.

10. Why do we go for time series management systems?
   Because of the specialized nature of time series data, and the lack of support in older DBMSs, it has been common to use specialized time series management systems rather than general purpose DBMSs for managing such information.

11. What is deductive database system?
   Some database systems provide capabilities for defining deduction rules for inferencing new information from the stored database facts. Such systems are called deductive databases.

12. Define prolog and datalog:
   The deductive database work based on logic has used prolog as a starting point. A variation of prolog called datalog is used to define rules declaratively in conjunction with an existing set of relations, which are themselves treated as literals in the language.

13. Define rules and facts:
   Facts are specified in a manner similar to the way relations are specified, except that it is not necessary to include the attribute name. Rules specify virtual relations that are not actually stored but that can be formed from the facts by applying inference mechanisms based on the rule specifications.

14. Write about clausal form and horn clauses:
   In clausal form, the formula is made up of a number of clauses, where each clause is composed of a number of literals connected by 'OR' logical connectives only. Hence the clausal form of a formula is a disjunction of literals, i.e., conjunction of clauses, e.g.:\(-\text{not}(p_1) \text{ or not}(p_2) \text{ or...not}(p_n) \text{ or } q_1 \text{ or } q_2 \text{ or...qm.}
   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, e.g.: \(-\text{or not}(p_2) \text{ or...or not}(p_n) \text{ or } q.

15. Define fact defined predicates and rule defined predicates:
   Fact defined predicates (or relations) are defined by listing all the combinations of values (the tuples) that make the predicate true. They correspond to base relations whose contents are stored in a database system. Rule defined predicates (or views) are defined by being the head (LHS) of one or more datalog rules. They correspond to virtual relations whose contents can be inferred by the inference engine.

16. What is knowledge base?
   The knowledge base is also called the international database. It consists of rules, as opposed to the base data, which constitutes the extensional database. The knowledge base means the combination of both the intensional and extensional databases. It often includes complex objects as well.
17. What is knowledge base management system (KBMS)?

The software that manages the knowledge base is known as the knowledge base management system. It is similar to deductive DBMS.

18. What is valid time and valid time database?

The most natural interpretation is that the associated time is the time that the event occurred or the period during which the fact was considered to be true in the real world. If this interpretation is used, the associated time is called as valid time. A temporal database using this interpretation is called a valid time database.

19. What is detached consideration?

The condition is evaluated as a separate transaction, spawned from the triggering transaction. This is known as detached consideration.

20. What is a predicate?

A predicate has an implicit meaning which is suggested by the predicate name and a fixed number of arguments. If the arguments are all constant values, the predicate states that a certain fact is true. If the predicate has variables as arguments, it is either considered a query or as a part of a rule or constraint.

16 Marks

1. Explain in detail about active database concepts and triggers:
   - Generalised model for active databases and Oracle triggers.
   - Design and implementation issues for active databases.
   - Potential Applications for active databases.
   - Triggers in SQL

2. What are temporal databases? Explain in detail.
   - Temporal database concepts
   - Time representation, calendars and time dimensions
   - Incorporating time in relational databases using tuple versioning.
   - Incorporating time in object-oriented databases using attribute versioning.
   - Temporal querying and the TSQL2 language.
   - Time series data.

3. Write notes on deductive databases:
   - Definition of deductive databases.
   - Prolog / Datalog notation.
   - Clausal form and horn clauses.
   - Interpretations of rules.
   - Datalog programs and their safety.
   - Relational operations.
   - Non-recursive datalog queries.
UNIT 5

2 Marks

1. Define mobile databases?
   A database that is portable and physically separated from a centralized DB server, but capable of communicating with that server from remote sites, by allowing the sharing of corporate data is known as mobile database.
   Eg : Electronic valets, news reporting, brokerage services and automated sales forces.

2. Define wireless communication?
   The wireless medium on which mobile units and base stations communicate have bandwidth significantly lower than those of a wired network. The current generation of wireless technology has data rates that range from the tens to hundreds of kilobits per second (2G cellular telephony) to tens of megabits per second (wireless Ethernet, properly known as WiFi).

3. What are the applications of MANET?
   • Peer-to-peer, meaning that a mobile unit is simultaneously a client and a server.
   • Transaction processing and data consistency control become more difficult.
   Resource discovery and data routing by mobile units make computing in a MANET even more complicated.

4. What are the different types of data management issues?
   Data distribution and replication.
   Transaction models
   Query processing
   • Recovery and fault tolerance
   • Mobile db design
   Location based service
   Division of labor
   Security

5. Define “dozing”.
   A client is unreachable because it is dozing - an energy conserving state in which many subsystems or shut down or because out of range of a base station.

6. Define GIS?
   GIS are used to collect, model, store and analyze information describing physical properties of the geographical world. The scope of GIS broadly encompasses two types of data, they are:
   1. Spatial data
   2. Non-spatial data.
   Normally GIS is a rapidly developing domain that offers highly innovative approaches to meet
some challenging technical demands.

7. List out the applications of GIS?
   The applications of GIS are divided into three categories they are
   Cartographic applications
   Digital terrain modeling applications
   Geographic objects applications.
   The first two categories of GIS applications require a field-based representation, whereas the third category requires an object-based one.

8. What are the Data Management Requirements of GIS?
   The functional requirements of the GIS applications are
   Data Modeling and Representation
   Data Analysis
   Data Integration
   Data Capture

9. Mention the Specific Data operations of GIS?
   GIS applications are conducted through the use of special operators such as the following
   Interpolation
   Interpretation
   Proximity analysis
   Raster image processing
   Analysis of Networks
   Visualization

10. What are the two formats for the Representation of GIS?
    GIS data can be broadly represented in two formats they are:
    1. Vector
    2. Raster.
    Vector data represents geometric objects such as points, lines, and polygons. Raster data is characterized as an array of points. Raster images are n-dimensional arrays where each entry is a unit of the image and represents an attribute. Two-dimensional units are called pixels, while three-dimensional units are called voxels.

11. What is GDB?
    GDB stands for Genome Data Base. It is a catalog of human gene mapping data, a process that associates a piece of information with a particular location on the human genome.

12. What is Genotype?
    Mendel realized that genes can exist in numerous different forms known as alleles. A genotype refers to the actual allelic composition of an individual.

13. What is Gene Ontology?
It is a consortium, which was formed in 1998 as collaboration among three modal organism databases:
- Fly Base
- Mouse Genome Informatics (MGI)
- Saccharomyces or yeast Genome Database (SGD)

14. What are the ontologies developed by GO consortium?
GO consortium has developed three ontologies, to describe the attributes of genes, gene products, or gene product groups, they are:
- Molecular Functions
- Biological Process
- Cellular Component

15. What are the nature of Multimedia applications?
Repository Applications
- Presentation Applications
- Collaborative work using multimedia information

16. What are the various types of multimedia data?
The various types of multimedia data are text, graphics, images, animations, video, structured audio, audio, composite or mixed multimedia data.

17. What are the characteristics of hyperlinks (hypermedia links)?
Links can be specified with or without associated information, and they may have large descriptions associated with them.
Links can start from a specific point within a node or from the whole node.
Links can be directional or non-directional when they can be traversed in either direction.

18. What is an Ontology?
An ontology necessarily entails or embodies some sort of world view with respect to a given domain. The world view is often conceived as a set of concepts, their definitions and their inter-relationships which describe a target world.

19. Define DBA (Data Base Administrator)?
A DBA is an individual person or group of persons with an overview of one or more databases who controls the design and use of these databases.

20. What are the functions of DBA?
- A DBA provides the necessary technical support for implementing policy decisions of databases.
- DBA is responsible for the overall control of the system.
- DBA is the central controller of the database system who oversees and manages all the resources.
- DBA is responsible for authorizing access to the database for coordinating and monitoring its use and for acquiring software and hardware resources as needed.
21. What are responsibilities of DBA?
   Defining conceptual schema and database creation.
   - Storage structure and access-method definition.
   - Granting authorization to the users.
   - Physical organization modification.
   - Routine maintenance.
   - Job monitoring.

22. What is parallel database?
   In some of the architectures multiple CPUs are working in parallel and are physically located in a close environment in the same building and communicating at a very high speed. The databases operating in such environment is known as parallel databases.

23. What are the advantages of parallel databases?
   - Increased throughput
   - Improved response time
   - To query large databases
   - Substantial performance improvements
   - Increased availability
   - Greater flexibility
   - Large no. of users

24. What are the disadvantages of parallel databases?
   - More start up costs
   - Interference problem
   - Skew problem

25. What is spatial databases?
   Spatial databases keep track of objects in a multidimensional space. Spatial data support in databases is important for efficiently storing, indexing and querying of data based on spatial locations.

26. Define spatial db query and mention its types?
   Spatial query is the process of selecting features based on their geographic or spatial relationship with other features. The three types are:
   - Range query
   - Nearest neighbor query or adjacency
   - Spatial joins or overlays

27. Define spatial data model?
   The spatial data model is a hierarchical structure consisting of the following to correspond to representations of spatial data:
   - Elements
   - Geometrics
   - Layers
28. Define Data Warehouse?
A data warehouse is a collection of information as well as a supporting system. It provide access to data for complex analysis, knowledge discovery and decision making. They support high performance data and information.

29. What are the applications of Data warehouse?
OLAP (Online Analytical Processing)  
DSS (Decision Support System)  
OLTP (On Line Transaction Processing)

30. Define Data mining?
Data mining refers to the mining or discovery of new information in terms of patterns or rules from vast amount of data.

31. Define Association Rule?
A association rule is of the form X => Y, when X = {x1,x2,…xn} and Y = {y1,y2,….yn} are set of items, with xi and yi being distinct items for all i and j. This association rule states that if a customer buys X he or she is also likely to buy Y. In general any association rule has the form LHS => RHS, where LHS and RHS are set of items.

32. What are the classifications of Neural Networks?
Neural networks can be broadly classified in to two categories.
1. Supervised networks and  
Supervised Networks:
Adaptive methods that attempt to reduce the output error are supervised lemmings methods.

Unsupervised Networks:
Methods that develop internal representation without sample outputs are called unsupervised learning methods.

16 MARKS (Refer Book for Diagrams)

1. Explain mobile databases in detail.
1) Mobile databases:-
   ✦ Recent advances in portable and wireless technology have led to mobile computing.
   ✦ Some of the slow problems which may include data management, transaction management and database recovery.

*Mobile computing architecture:-
Mobile computing architecture is a distributed architecture
   ✦ A number of computer generally referred as fixed hosts and base stations.
Fixed hosts are general purpose computers.
Base stations functions as gateway to the fixed network for the mobile units.

*Wireless communication:*
- The wireless medium on which mobile units and base station communication.
- The data rates that ranges from the tens to hundreds of kilobytes per second to tens of megabits per second.
- Some wireless networks such as WIFI and bluetooth, use unlicensed areas of frequency spectrum.

*Client/network relationship:*
- Mobile units can move freely in a geographic mobility domain.
- MANET applications can be considered as peer-to-peer
- Sample MANET applications are multi-user games, shared whiteboards, distributed calendars, and battle information sharing.

**ii) Characteristics of mobile environment**
- The characteristics of mobile computing include high communication latency, intermittent wireless connectivity, etc.
- Mobile computing poses challenges for servers as well as clients.
- One way servers relieve this problem is by broadcasting data.
- Client mobility also allows new application that is location based.

**iii) Data management issues:**
- The entire database is distributed mainly among the wired components, possibly with full or partial replication.
- The database is distributed among wire and wireless components.
- The following are the additional considerations and variations of mobile DB:
  - Data distribution and replication.
  - Transaction models
  - Query processing
  - Recovery and fault tolerance
  - Mobile database design
  - Location-based service
  - Division of labour
  - Security

**iv) Application: Intermittently synchronized databases:**
- A client connects to the server when it wants to exchange updates. This communication maybe unicast.
- A server cannot connect to a client at will.
- Issues of wireless versus wired client connection and power conservation.
- A client is free to manage its own data and transaction while it is disconnected.
- A client has multiple ways of connecting to a server.

2. Explain the characteristics of biological data in GDB?
Characteristic 1:
Biological data is highly complex, when compared with most other domains or applications.

Characteristic 2:
The amount and range of variability in data is high.

Characteristic 3:
Schemas in biological databases change at a rapid pace.

Characteristic 4:
Representation of the same data by different biologists will likely be different.

Characteristic 5:
Most users of biological data do not require write access to the database; read only access is adequate.

Characteristic 6:
Most biologists are not likely to have any knowledge of the internal structure of the database or about schema design.

Characteristic 7:
The context of data gives added meaning for its use in biological applications.

Characteristic 8:
Defining and representing complex queries is extremely important to the biologist.

Characteristic 9:
Users of biological information often require access to “old” values of the data – particularly when previously reported results.

3. Write short notes on – Multimedia Databases.

i) Nature of Multimedia Data & Applications:

Multimedia Data:
- Text
- Graphics
- Images
- Animations
- Video
- Structured Audio
- Audio
- Composite or mixed multimedia data

Multimedia Applications:
- Repository Applications
- Presentation Applications
- Collaborative work using multimedia information

ii) Data Management Issues:
- Modeling
- Design
- Storage
- Queries and Retrieval
- Performance
iii) Open Research Problems:

- Information Retrieval Perspective in Querying Multimedia Databases
- Requirements of Multimedia/hypermedia Data Modeling & Retrieval
- Indexing of Images:
  - Three indexing schemes are
    - Classificatory Systems
    - Keyword-based Systems
    - Entity-attribute-relationship Systems

Problems in Text Retrieval:
- Phrase indexing
- Use of Thesaurus
- Resolving ambiguity

iv) Multimedia Database Applications:

- Some important applications are:
  - Documents and records management
  - Knowledge dissemination
  - Education and Training
  - Marketing, advertising, retailing, entertainment and travel
  - Real-time control and monitoring
- Commercial Systems for Multimedia Information Management

4. Explain in detail the Parallel & Spatial Database.

i) Parallel Data Base:

Architecture of parallel data bases:
- Shared-memory multiple CPU
- Shared-disk multiple CPU
- Shared-nothing multiple CPU

Key elements of parallel DB processing:
- Speed-up
- Scale-up
- Synchronization
- Locking

Query Parallelism:
- I/O parallelism
- Intra-query parallelism
- Inter-query parallelism
- Intra-operation parallelism
- Inter-operation parallelism

ii) Spatial Data Base:

- Spatial DB characteristics:
- Spatial Data Model:
  - Elements
  - Geometries
  - Layers
- Spatial database queries:
5. Define Data mining? Briefly explain the different types of knowledge discovered during data mining.

**Data Mining**:
Data mining refers to the mining or discovery of new information in terms of patterns or rules from vast amount of data.

*Types of knowledge discovered during Data mining*:

Association Rule:
- Apriori Algorithm
- Sampling algorithm
- Frequent pattern tree algorithm
- Partition algorithm

Classification

Clustering

Approaches to other Data mining problems:
- Discovery of sequential patterns
- Discovery of patterns in time series
- Regression
- Neural networks
- Genetic algorithms

Applications of Data Mining:
- Marketing
- Finance
- Manufacturing
- Health care