✆ STD - 0431 - 2670142
FAX - 0431 - 2670143

DHANALAKSHMI SRINIVASAN
INSTITUTE OF TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Anna University)
NH - 45, Trichy - Chennai Trunk Road,
SAMAYAPURAM,TRICHY - 621 112.
E.mail:dsit2011@gmail.com Website:www.dsit.ac.in

DEPARMENT OF COMUTER SCIENCE AND ENGINEERING
QUESTION BANK

Subject Code: CS8492	**Branch/Year/Sem: CSE/ II / IV**
Subject Name : DATABASE MANAGEMENT SYSTEMS	**Batch:2018-2022**
Staff Name: Ms.R.PADMAVATHI	**Academic Year:2019-2020(EVEN)**

	UNIT-1 RELATIONAL DATABASES
S.NO	**PART-A**
1	Define Database Management System and its apllications. Database management system (DBMS) is a collection of interrelated data and a set of programs to access those data. Applications:Banking, Airlines,Universities, Credit card transactions,Tele communication,Finance, Sales, Manufacturing and Human resource.
2	List the advantages of DBMS The advantages of using a DBMS are a) Controlling redundancy b) Restricting unauthorized access c) Providing multiple user interfaces d) Enforcing integrity constraints. e) Providing backup and recovery
3	Classify the levels of abstraction. a)Physical level-Describes how the data are actually stored. b) Logical level-Describes what data are stored and what relationships exist among those data c) View level – Describes the part of entire database
4	What are the types of data models? a) Entity relationship model b) Relational model c) Hierarchical model d) Network model e) Object based data model(Object Oriented & Object relational)
5	Define storage manager. A storage manager is a program module that provides the interface between the low level data stored in a database and the application programs and queries submitted to the system.
6	List out the functions of database administrator. a) Schema definition

b) Storage structure and access method definition
c) Schema and physical organization modification
d) Granting user authority to access the database
e) Specifying integrity constraints
f) Monitoring performance and responding to changes in requirements

7 What is data dictionary?
A data dictionary is a data structure which stores meta data about the structure of the database i.e. the schema of the database.

8 What is a primary key?
A primary key is an attribute that allows us to identify uniquely an entity in the entity set.

9 Define relational algebra.
The relational algebra is a procedural query language. It consists of a set of operations that take one or two relation as input and produce a new relation as output.

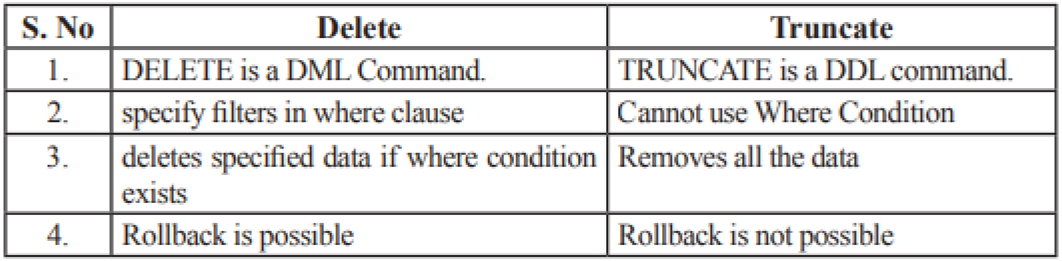
10 Define foreign and candidate key
A key in which primary key attribute of one table and same attribute be used in another table, then that key is known as foreign key.

11 What are referential integrity constraints?
A value that appears in one relation for a given set of attributes also appears for a certain set of attributes in another relation.

12 Define SQL .
Structured Query Language (SQL) is the standard common set used to communicate with the relational database management systems. All tasks related to relational data management— creating tables, querying the database for information, modifying the data in the database, deleting them, granting access to users, and so on — can be done using SQL.

13 List the DML Commands used in SQL.
SELECT: This command is used to retrieve rows from a table.
The select syntax:
SELECT [column name(s)] from [table name] where [conditions].
UPDATE: This command modifies data of one or more records.
The update command syntax:
UPDATE table name SET column name = value where [condition].
INSERT: This command adds one or more records to a database table.
The insert command syntax:
INSERT INTO table name [column(s)] VALUES [value(s)].
DELETE: This command removes one or more records from a table according to specified conditions.
Delete command syntax:
DELETE FROM table name where [condition].

14 Compare delete command and truncate command.



S. No	Delete	Truncate
1.	DELETE is a DML Command.	TRUNCATE is a DDL command.
2.	specify filters in where clause	Cannot use Where Condition
3.	deletes specified data if where condition exists	Removes all the data
4.	Rollback is possible	Rollback is not possible

15 What is Data Control Language?
create privileges to allow users access to, and manipulation of, the database.
There are two main commands:

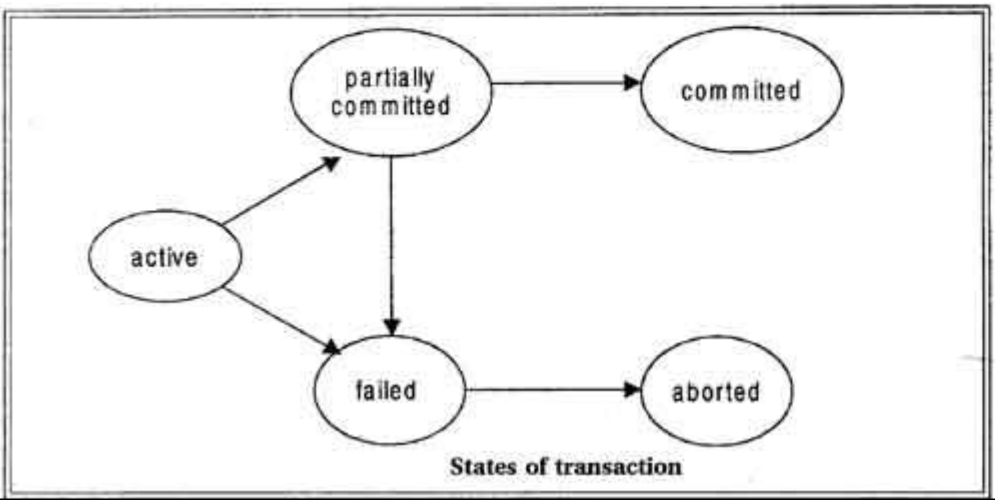
	GRANT to grant a privilege to a user REVOKE to revoke (remove) a privilege from a user
16	Define Dynamic SQL. Programs that contain embedded dynamic SQL statements must be precompiled like those that contain static SQL, but unlike static SQL, the dynamic statements are constructed and prepared at run time. The source form of a dynamic statement is a character string that is passed to DB2 by the program using the static SQL statement PREPARE or EXECUTE IMMEDIATE.
17	Define instance and schemas. April/May 2012 - Database change over times as information is inserted and deleted. - The collection of information stored in the database at a particular moment called an instance of the database. - The overall design of the database is called the database schema
18	List the aggregation functions in SQL Aggregate functions are functions that take a collection of values as input and return a single value as output. SQL offers 5 built in aggregate funtions: avg - average value min - minimum value max - maximum value sum - sum of values count - number of values.
19	Define views. A view is an object that gives the user a logical view of data from an underlying table or tables (relation or relations). ™ It is not desirable for all users to see the entire logical model.
20	What is Embedded SQL? An embedded SQL statement is distinguished from the host language statements by enclosing it between EXEC SQL or EXEC SQL BEGIN and a matching END-EXEC or EXEC SQL END (or semicolon) ™ Syntax may vary with language ™ Shared variables (used in both languages) usually prefixed with a colon (:) in SQL
	PART-B
1	Explain the basic architecture of a database management systems.
2	Discuss about the different types of model in DBMS.
3	With relevant example to discuss about the various operations of Relation Algebra.
4	Explain the different types of languages used in SQL with an example.
5	Discuss about the types of keys with an example.
6	Explain in detail about Advanced SQL and Embedded SQL.

7	Consider the employee database , where the primary keys are Underlined.
	employee(empname,street,city),works(empname,companyname,salary)
	company(companyname,city), manages(empname,management)
	Give an expression in the relational algebra for each request.

	UNIT II DATABASE DESIGN
S.NO	**PART-A**
1	What is an entity-relationship model? The entity relationship model is a collection of basic objects called entities and relationship among those objects. An entity is a thing or object in the real world that is distinguishable from other objects.
2	What is an entity and entity set? An entity is an object that exists and is distinguishable from other objects. Example: specific person, company, event, plant Entity set: The set of all entities of the same type is termed as an entity set.
3	Define single valued and multivalued attributes. Single valued attributes : Attributes with a single value for a particular entity Multivalued attributes : Attributes with a set of value for a particular entity
4	Define null values. In some cases a particular entity may not have an applicable value for an attribute or if we do not know the value of an attribute for a particular entity. In these cases null value is used
5	Define Mapping cardinalities. Mapping cardinalities or cardinality ratios express the number of entities to which another entity can be associated. Mapping cardinalities must be one of the following: • One to one • One to many • Many to one • Many to many
6	Define the terms Generalization and Aggregation Generalization is a containment relationship that exists between a high-level Entity set and one or more low-level entity set. Aggregation is an abstraction through which relationships are treated as higher-level entities
7	Compare weak and strong entity sets. Weak entity set : entity set that do not have key attribute of their own are called weak entity sets. Strong entity set : Entity set that has a primary key is termed a strong entity set
8	What is ER diagrams? An entity-relationship diagram is a data modeling technique that creates a graphical representation of the entities, and the relationships between entities, within an information system.
9	What are the steps involved in creating ERD? a) Identify the entities. b) Find the relationships c) Identify key attributes for every entity d) Draw the ERD.
10	What is normalization? Database normalization is the process of organizing the fields and tables of a relational database to minimize redundancy and eliminate dependency. Normalization usually involves dividing large tables into smaller (and less redundant) tables and defining relationships between them.
11	Define 1NF. A relation said to be first normal form if and if only all attributes are atomic in nature
12	Define 2NF. A relation said to be second normal form if and if only relation should be in first normal form and to

	eliminate partial dependency
13	Define 3NF. A relation said to be third normal form if and if only relation should be in second normal form and to eliminate transitive dependency.
14	Define BCNF. A relational schema R is in Boyce–Codd normal form if and only if for every one of its dependencies X → Y, at least one of the following conditions hold ™X → Y is a trivial functional dependency (Y ⊆ X) ™X is a superkey for schema R
15	What is 4NF and 5NF? A Table is in 4NF if and only if, for every one of its non-trivial multivalued dependencies X Y, X is a superkey that is, X is either a candidate key or a superset A table is said to be in the 5NF if and only if every non-trivial join dependency in it is implied by the candidate keys.
16	Define De-normalization. De-normalization is the process of attempting to optimize the read performance of a database by adding redundant data or by grouping data.
17	What is sub-class and super class? For example, members of entity Employee can be grouped further into Secretary, Engineer, Manager, Technician, Salaried_Employee. The set listed is a subset of the entities that belong to the Employee entity, which means that every entity that belongs to one of the sub sets is also an Employee. Each of these sub-groupings is called a subclass, and the Employee entity is called the super-class.
18	List out the constraints on Specialization and Generalization. Several specializations can be defined on an entity type. ™ Entities may belong to subclasses in each of the specializations. ™ The specialization may also consist of a single subclass, such as the manager specialization, in this case we don't use the circle notation.
	PART-B
1	Explain about Entity relationship model with ER diagram and example.
2	Draw an ER diagram for Banking System and Payroll system.
3	What is functional dependency? Explain the types of it with an example.
4	What is normalization? Explain the different types of normal forms with an example.
5	Draw an ER diagram for Hospital Management System.
6	Construct ER-Relation mapping diagram for College management system.
7	Define generalization and aggregation. Demonstrate it using ER diagram.
8	Compare the different types of normal form in detail.

	UNIT III TRANSACTIONS
	PART-A
1	What are the ACID properties? Atomicity, Consistency, Isolation, Durability is a set of properties that guarantee database transactions are processed reliably. In the context of databases, a single logical operation on the data is called a transaction. For example, a transfer of funds from one bank account to another, even though that might involve multiple changes (such as debiting one account and crediting another), is a single transaction.
2	What is transaction? Collections of operations that form a single logical unit of work are called transactions.
3	What are the two statements regarding transaction? The two statements regarding transaction of the form: Begin transaction and End transaction
4	What are the two types of serializability? A (possibly concurrent) schedule is serializable if it is equivalent to a serial schedule. Different forms of schedule equivalence gives rise to the notions of: • Conflict serializability • View serializability
5	List the SQL statements used for transaction control. *Commit :* Saves all transactions *Rollback :* Used to undo transactions *Savepoint :* Establishes a point back to which you may roll *Set Transaction :* Establishes properties for current transaction.
6	What is Conflict-Serializability? Conflict-Serializability is defined by equivalence to a serial schedule (no overlapping transactions) with the same transactions, such that both schedules have the same sets of respective chronologically ordered pairs of conflicting operations (same precedence relations of respective conflicting operations).
7	What is View-Serializability? View-Serializability of a schedule is defined by equivalence to a serial schedule (no overlapping transactions) with the same transactions, such that respective transactions in the two schedules read and write the same data values ("view" the same data values)
8	What are two pitfalls (problem) of lock-based protocols? *Shared Lock :* If a Transaction Ti has obtained a shared mode lock on data item Q, then Ti can read, but cannot write Q. *Exclusive Lock :* If the transaction Ti has obtained a shared mode Lock on item Q, then Ti can read and also write Q.
9	What is meant by deadlock? A system is in a deadlock state if there exists a set of transaction such that every transaction in the set is waiting for another transaction in the set.
10	Define the phases of two phase locking protocol *Growing phase:* a transaction may obtain locks but not release any lock. *Shrinking phase:* a transaction may release locks but may not obtain any new locks.
11	What is Time-stamp based protocol? Timestamp based protocol ensures Serializability. It selects an ordering among transactions in advance using time stamps. With each Transaction in the system, a unique fixed timestamp is associated. It is denoted by TS(Ti). This

	timestamp is assigned by the database system before the transaction Ti status execution. If a transaction Ti has been assigned time- stamp TS(Ti) and new transaction Tj enters the system, then TS(Ti)<TS(Tj).
12	What is Concurrency Control? Process of managing simultaneous execution of transactions in a shared database, to ensure the Serializability of transactions, is known as concurrency control.
13	List out the statements associated with a database transaction? 6. • Commit work . • Rollback work
14	Draw the states of Transaction.  partially committed committed active failed aborted States of transaction
15	Difference between deadlock prevention and deadlock avoidance. Deadlock Prevention: • Preventing deadlocks by constraining how requests for resources can be made in the system and how they are handled (system design). • The goal is to ensure that at least one of the necessary conditions for deadlock can never hold. Deadlock Avoidance: • The system dynamically considers every request and decides whether it is safe to grant it at this point, • The system requires additional apriori information regarding the overall poten- tial use of each resource for each process. • Allows more concurrency.
	PART-B
1	Explain about ACID properties with suitable example.
2	Describe briefly about Serializability and its types with relevant example.
3	Discuss the following transaction with relevant example: I. Read Only Transaction II. Read Write Transaction III. Aborted Transaction
4	Discuss in detail about two phase commit protocol.
5	Explain the concepts of Concurrency control mechanism.
6	Describe briefly about deadlock handling.
7	Why concurrency control is needed? Explain the problems that would arise when concurrency control is not provided by the database system.

UNIT IV –IMPLEMENTATION TECHNIQUES

PART-A

1 List out the levels of RAID.

RAID level 0 –Block striping non-redundant
RAID level 1 –Mirroring
RAID level 2 –Memory style Error correcting code
RAID level 3 –Bit interleaved parity
RAID level 4 –Block interleaved parity
RAID level 5 – Block interleaved distributed parity
RAID level 6 – P+Q redundancy disk

2 List the merits and demerits of B+ tree index structure.

Mertis
Insertion in B+tree is easy.
Deletion in B+ tree is simple than B tree.
Demeris :
Requires redundant storage for search-key values every search key appears in some leaf node.
Several are repeated in non-leaf node.
Lookup on B+-tree requires traversal of a path from root of tree to some leaf node.

3 Differentiate Static hashing and Dynamic hashing.

S. No	Static hashing	Dynamic hashing
1.	Numbers of buckets are fixed	Numbers of buckets are not fixed
2.	As the file grows, performance decreases.	As the file grows, performances do Not degrade
3.	Space overhead is more	Space overhead is less
4.	Do not use bucket address table	Bucket address table is used
5.	Open hashing and closed hashing are forms of it.	Extendable hashing and linear hashing are forms of it.
6.	Implementation is simple	Implementation is complex
7.	It is less attractive techniques	It is more attractive techniques
8.	Overflow chaining is not used	Overflow chaining is not used

4 What are the two types of ordered indices?

- Primary index
- Secondary index

5 What are different types of file organization?

- Sequential file organization
- Heap file organization
- Hashing file organization

6 Difference between dense index and parse index.

S.No	Dense index	Parse index
1	Index entry for every search key value (and hence every record) in the data file.	Index entries for only some of the search values.
2.	Dense indices are faster than sparse	Sparse indices are slower than dense
3	Require more space and impose more	Require less space and impose less

	maintenance for insertion and deletion	maintenance for insertion and deletion

7 What are the factors to evaluate the indexing technique?

- Access types
- Access time
- Insertion time
- Deletion time
- Space overhead

8 What is B+ tree?

A B+ tree is an n-ary tree with a variable but often large number of children per node. The root may be either a leaf or a node with two or more children.A B+ tree can be viewed as a B-tree in which each node contains only keys (not key-value pairs), and to which an additional level is added at the bottom with linked leaves.

9 Difference between B+ tree and B tree.

S. No	B+ tree	B- tree
1.	B+ tree leaf node data are ordered	B tree the leaf node cannot store
	in a sequential linked list.	using linked list.
2.	B +tree store redundant search key	B tree store non-redundant search Key
3.	B+-tree data store only leaf nodes.	B tree search keys and data stored in
		internal or leaf nodes
4.	Insertion of a B+ tree is not complicated	Insertion of a B tree is more complicated

10 Define Seek time.

Seek time: Time it takes to reposition the arm over the correct track. The seek time ranges from 2 to 30 milliseconds. The average seek time is one-third the worst case seek time and one half the maximum seek time. Average seek time currently ranges between 4 to 10 milliseconds.

11 Define File organization.

The database is stored as a collection of files. Each file is a sequence of records. A record is a sequence of fields. We have a file with 2 types of records.

Fixed length records.

Variable length records.

12 What are the two main goals of parallelism?
Load –balance multiple small accesses, so that the throughput of such accesses increases.
Parallelize large accesses so that the response time of large accesses is reduced

13 What are the factors to evaluate the indexing technique?

- Access types
- Access time
- Insertion time
- Deletion time
- Space overhead

14 What are the two types of blocks in the fixed-length representation?
Anchor block: Contains the first record of a chain.
Overflow block: Contains the records other than those that are the first record o f a chain.

15 List the five methods of file organization

- Sequential organization.
- Indexed-sequential organization.
- Direct or Hash organization.
- Heap organization.
- Clustered Organization

16 What are the two indices in files?

There are 2 basic kinds of indices.

Ordered indices - Based on a stored ordering of the values

Hash indices - Based on a uniform distribution of values across a range of buckets.

17 List the advantages and disadvantages of B+ tree.

Advantage of B+-tree index files

- Automatically reorganizes itself with small, local changes, in the face of insertions and deletions. Reorganization of entire file is not required to maintain performance.

Disadvantage of B+-tree

- Extra insertion and deletion overhead, space overhead.

18 Difference between Primary index and Secondary index.6.

S. No	Primary index	Secondary index
1.	An index on a set of fields that includes the unique primary key for the field	An index that is not a primary key
2.	Guaranteed not to contain duplicates	May have duplicates
3.	Also Called a Clustered index.	Also Called a Non-Clustered index.
4.	Eg: Employee ID	Eg: Employee name

PART-B

1 What are the various ways of organizing records in files and explain any one file organization in detail.

2 Describe the structure of B+ tree and list the characteristics of B+ tree.

3 What is RAID? Briefly explain different level of RAID.

4 Give brief notes on overviews of physical storage media.

5 Describe briefly about indexing and hashing.

6 Explain about query processing in detail.

7 Discuss about query optimization in detail.

UNIT V – ADVANCED TOPICS

PART-A

1 What are the two approaches to store a relation in the distributed database?

Replication: System maintains several identical replicas (copies) of the relation and stores each replica at a different site.

Fragmentation: System Partitions the relation into several fragments and stores each fragment at a different site.

2 Define Distribute Database.

The computers in a distributed system communicate with one another through various communication media, such as high-speed networks or telephone lines. They do not share main memory or disk. The computers in distributed system are referred by names such as sites or nodes.

3 What are the types of Transactions?

Distributed database system supports two types of transactions.

Local transaction : It is one that accesses data only from site where that transac- tion was initiated.

Global transaction : It is one that either accesses data from a site other than the site where that transaction was initiated or accesses data from several different sites.

4 Difference between homogeneous and heterogeneous database.

S. No	Homogeneous Database	Heterogeneous Database
1.	Different nodes may have same hardware & software	Database application used at each location must be same or compatible.
2.	Much easier to design and manage	Tough to design and manage
3.	Database application used at each location must be same or compatible	Database application used at each location must be incompatible.

5 What are the advantages & disadvantages of distributed databases?

Advantages :

- Sharing data
- Autonomy
- Availability

Disadvantages :

- Software development cost
- Greater potential for bugs
- Increased processing overhead

6 Define 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.

7 What is Object Oriented Database?

An object-oriented database is a database that subscribes to a model with information rep resented by objects. Object-oriented databases are a niche offering in the relational database management system (RDBMS) field and are not as successful or well-known as mainstream database engines.

8 What is Object?

Object consists of entity and attributes which can describe the state of real world object and action associated with that object.

9 What are the three types of attributes in OODB?

- Simple Attributes
- Complex Attributes
- Reference Attribute

10 What is ODMG?

The Object Database Management Group (ODMG) is made up of the leading Object Database vendors plus a large number of companies that are interested in an ODBMS standard. The ODMG has produced a standard for object databases. The ODMG Standard is an interoper- ability standard which allows applications written to the standard to run on any com pliant system

11 List out the standard in ODMG.

- Object Model
- Object Definition Language (ODL)
- Object Query Language (OQL)
- C++ Language Binding
- Java Language Binding

12 What is XML Database?

XML Database is used to store huge amount of information in the XML format. As the use of XML is increasing in every field, it is required to have a secured place to store the XML docu- ments. The data stored in the database can be queried using XQuery, serialized, and exported into a desired format.

13 Difference between HTML and XML.

S.No.	HTML	XML
1)	HTML is used to display data and focuses on how data looks.	XML is a software and hardware indepen- dent tool used to transport and store data. It focuses on what data is.
2)	HTML is a markup language itself.	XML provides a framework to define markup languages.
3)	HTML is not case sensitive.	XML is case sensitive.
4)	HTML is a presentation language.	XML is neither a presentation language nor a programming language.
5)	HTML has its own predefined tags.	You can define tags according to your need.
6)	In HTML, it is not necessary to use a closing tag.	XML makes it mandatory to use a closing tag.
7)	HTML is static because it is used to display data.	XML is dynamic because it is used to transport data.
8)	HTML does not preserve whitespaces.	XML preserve whitespaces.

14 Define DTD.

DTD stands for Document Type Definition. It defines the legal building blocks of an XML document. It is used to define document structure with a list of legal elements and attributes.

15	What is Information Retrieval (IR)? - Information retrieval (IR) systems use a simpler data model than database sys- tems - Information organized as a collection of documents - Documents are unstructured, no schema - Information retrieval locates relevant documents, on the basis of user input such as keywords or example documents
16	What is Web Crawling? Crawling is done by multiple processes on multiple machines, running in parallel .Set of links to be crawled stored in a database.New links found in crawled pages added to this set, to be crawled later·Indexing process also runs on multiple machines.Creates a new copy of index instead of modifying old indexOld index is used to answer queries .After a crawl is “completed” new index becomes “old” index
17	Define Simple attributes. Attributes can be of primitive data type such as, integer, string, real etc. which can take literal value. *Example:* 'ID' is simple attribute and value is 07.
18	List out the features of OODB. - Complexity - Inheritance - Encapsulation - Persistency
19	What is XQuery? XQuery is a standardized language for combining documents, databases, Web pages and almost anything else. It is very widely implemented. It is powerful and easy to learn. XQuery is replacing proprietary middleware languages and Web Application development languages.
20	What is XML schema? XML schema is a language which is used for expressing constraint about XML documents. There are so many schema languages which are used now a days for example Relax- NG and XSD (XML schema definition).An XML schema is used to define the structure of an XML document. It is like DTD but provides more control on XML structure.
	PART-B
1	Discuss in detail about the various types of distributed database with suitable examples.
2	Explain about Object Oriented Databases with suitable example.
3	Describe brief notes on ODMG.
4	Discuss in detail about XML Database with relevant examples.
5	Explain the DTD with suitable example.
6	Describe briefly on Information Retrieval (IR) with suitable example.