MUTHAYAMMAL ENGINEERING COLLEGE

(An Autonomous Institution)

(Approved by AICTE, New Delhi, Accredited by NAAC & Affiliated to Anna University)
Rasipuram - 637 408, Namakkal Dist., Tamil Nadu

Estd. 2000

IQAC

MUST KNOW CONCEPTS

MKC

CSE

2020-2021

Subject		19CSC03 Database Management Systems		
S. No.	**Term**	**Notation (Symbol)**	**Concept/Definition/Meaning/Units/Equation/ Expression**	**Units**
UNIT I INTRODUCTION AND CONCEPTUAL MODELING				
1	Database administrator	DBA	A person responsible for the design, implementation, maintenance and repair of an organization's database.	
2	Types Of Data Model		• Hierarchical model • Network model • Relational model • Entity-relationship • Object-relational model	
3	Database Management System	DBMS	Database management system is a collection of interrelated data and a set of programs to access those data.	
4	Applications of DBMS		• Banking • Airlines • Universities • Credit card transactions • Tele communication • Finance	
5	Disadvantages of File Processing System		• Data redundancy and inconsistency • Difficulty in accessing data • Data isolation • Integrity problems • Atomicity problems • Concurrent access anomalies	
6	Advantages of DBMS		• Controlling redundancy • Restricting unauthorized access • Providing multiple user interfaces • Enforcing integrity constraints. • Providing backup and recovery	
7	Levels Of Data Abstraction		**a)** Physical level **b)** Logical level **c)** View level	
8	Instance		Collection of data stored in the data base at a particular moment is called an Instance of the database.	
9	Schema		The overall design of the data base is called the data base schema.	

10	Conceptual Schema		The schemas at the view level are called subschema's that describe different views of the database.	
11	Storage Manager		The low level data stored in a database and the application programs and queries submitted to the system.	
12	Components of Storage Manager		**a)** Authorization and integrity manager **b)** Transaction manager **c)** File manager **d)** Buffer manager	
13	Disk Storage		**a)** Data files **b)** Data dictionary **c)** Indices	
14	Data Dictionary		A data dictionary is a data structure which stores meta data about the structure of the database ie. The schema of the database.	
15	Entity Relationship Model		The entity relationship model is a collection of basic objects called entities and relationship among those objects.	
16	Attributes		An entity is represented by a set of attributes. Attributes are descriptive properties possessed by each member of an entity set.	
17	Relationship		A relationship is an association among several entities.	
18	Entity set		The set of all entities of the same type is termed as an entity set.	
19	Relationship set		The set of all relationships of the same type is termed as a relationship set.	
20	Types of Attributes		• Single valued attributes • Multi valued attributes • Stored attributes • Derived attributes • Composite attributes	
21	Degree of Relationship set		The degree of relationship type is the number of participating entity types	
22	Mapping cardinalities		• One to one • One to many • Many to one • Many to many	
23	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.	
24	Types of Calculus		1. Tuple relational calculus. 2. Domain relational calculus	
25	Basic operations in relational algebra		Selection, Projection, Cartesian product, Union, and Set Difference	
UNIT II RELATIONAL MODEL				
26	Structured Query Language	SQL	(SQL) is a standard computer language for relational database management and data manipulation.	
27	Query Language		A query is a statement requesting the retrieval of information.	
28	Data Definition	DDL	Data base schema is specified by a set of definitions	

	Language		expressed by a special language	
29	Data Manipulation Language	DML	A language that enables users to access or manipulate data as organized by the appropriate data model	
30	Different types of DML		• Procedural DML • Non-procedural DML	
31	DDL commands		CREATE DROP ALTER TRUNCATE RENAME	
32	DML commands		INSERT UPDATE DELETE	
33	Data Control Language	DCL	DCL includes commands such as GRANT and REVOKE which mainly deals with the rights, permissions and other controls of the database system.	
34	Transaction Control Language	TCL	TCL commands deals with the transaction within the database.	
35	TCL commands		COMMIT ROLLBACK SAVEPOINT	
36	Primary key		A primary is a column or set of columns in a table that uniquely identifies tuples (rows) in that table.	
37	Foreign key		Foreign keys are the columns of a table that points to the primary key of another table. They act as a cross-reference between tables.	
38	Candidate key		A super key with no redundant attribute is known as candidate key	
39	Super key		A super key is a set of one of more columns (attributes) to uniquely identify rows in a table.	
40	Built in domains available in SQL		1. Char(n) 2. Varchar(n) 3. Int 4. Small int, 5. Numeric (p,d) 6. Neat double precision 7. Float 8. Date 9. Time.	
41	Aggregate Functions		Aggregate functions are functions that take a collection of values as input and return a single value.	
42	View		A view in SQL is a single table ie., derived from other tables like base tables.	
43	Functional dependency		It is a constraint between two sets of attributes from the database. It is denoted by x -> y	
44	Armstrong rule		Reflexive rule, Augmentation rule, Transitive Rule	
45	Properties of Decomposition		1. Lossless join 2. Dependency preservation 3. No repletion of information	
46	Normalization		• Minimizing redundancy • Minimizing the insertion deletion update anomalies.	
47	First normal form	1NF	Each column is unique in 1NF.	
48	Second normal form	2NF	The entity should be considered already in 1NF, and all attributes within the entity should depend solely on the unique identifier of the entity.	

49	Third normal form	3NF	The entity should be considered already in 2NF, and no column entry should be dependent on any other entry (value) other than the key for the table.	
50	Boyce codd normal form	BCNF	3NF and all tables in the database should be only one primary key.	
UNIT III DATA STORAGE AND QUERY PROCESSING				
51	Types of Data Storage		Cache, Main Memory, Flash Memory, Magnetic Disk Storage, Optical Storage and Tape Storage	
52	Cache		The cache is the fastest and memory is small; its use is managed by the operating system.	
53	Disk Controller		It is an interface between the computer system and the actual hardware of the disk drive.	
54	Data-Transfer Rate		The data-transfer rate is the rate at which data can be retrieved from or stored to the disk.	
55	File Organization		A file is organized logically as a sequence of records. These records are mapped on to disk blocks.	
56	Index-Sequential Files		The files that are ordered sequentially with a primary index on the search key.	
57	Types of indices		a) Ordered indices b) Primary Indices c) Secondary Indices	
58	Primary Index		An ordered index whose search key is also the sort key used for the sequential file	
59	Dense index		An index record appears for every search-key value in the file.	
60	Sparse Index		An index record is created for only some of the Values.	
61	Secondary Index		An ordered index whose search key is NOT the sort key used for the sequential file	
62	B+ tree		This is a Balanced Tree with intermediate Nodes and Leaf Nodes	
63	B tree		Each node will have only two branches and each node will have some records.	
64	Hashing		Hash File organization method is the one where data is stored at the data blocks whose address is generated by using Hash Function.	
65	Two Types of Hash File Organization		Static Hashing Dynamic Hashing	
66	Static Hashing		The resultant data bucket address will be always same.	
67	Dynamic Hashing		The data buckets grow or shrink as the records increase or decrease.	
68	Linear Probing		Linear probing is a type of open hashing. If a bucket is full the system inserts records in to the next bucket that has space.	
69	Fixed length records		Every record has the same fields and field lengths are fixed.	
70	Variable length records		File records are of same type but one or more of the fields are of varying size.	

71	Query Optimization		It is a process of selecting the most efficient query evolution plan for a query.	
72	Query Processing		The entire process of translating a query into low level instructions in which the DBMS can easily work with.	
73	Basic Steps in Query Processing		Parsing and Translation Query Optimization Evaluation or query code generation	
74	Minimize the Cost Function		I/O Cost + CPU Cost + Communication Cost	
75	Two Methods of Query Optimization		Cost based Optimization (Physical) Heuristic Optimization (Logical)	
UNIT IV TRANSACTION MANAGEMENT				
76	Transaction		Transaction is a unit of program execution that accesses and possibly updated various data items.	
77	Operation Transaction		Read(x)- transfer data item x from database. Write(x)- transfer data item x from the local buffer.	
78	ACID Properties		1. Atomicity 2. Consistency 3. Isolation 4. Durability.	
79	Atomicity		This property ensures that either the transaction occurs completely or it does not occur at all.	
80	Consistency		It ensures that the database remains consistent before and after the transaction.	
81	Isolation		This property ensures that multiple transactions can occur simultaneously without causing any inconsistency.	
82	Durability		This property ensures that all the changes made by a transaction after its successful execution are written successfully to the disk.	
83	Transaction States		**i)** active **ii)** Partially Committed **iii)** Failed **iv)** Abort **v)** committed.	
84	Schedule		A sequences of instructions that specify the order in which instructions of concurrent transactions are executed	
85	Serializability		Set of transaction will have an effect equivalent to a schedule in some order of guarantee.	
86	Types of Serializability		Conflict Serializability View Serializability	
87	Conflict Serializability		A schedule is called conflict serializable if we can convert it into a serial schedule after swapping its non-conflicting operations.	
88	View Serializability		A serial schedule with the same transactions, such that respective transactions in the two schedules read and write the same data values	
89	Concurrency Control		It deals with interleaved execution of more than one transaction	
90	Two Concurrency control protocols		Lock based protocols Time stamp based protocols	
91	Lock		Lock are used as a means of synchronizing the access by concurrent transaction to the database item.	
92	Modes of lock		• Shared • Exclusive	

93	Two-Phase Locking		Growing phase Shrinking phase	
94	Deadlock		Neither of the transaction can ever proceed with its normal execution. This situation is called deadlock.	
95	Recovery from deadlock		1.Selection of a victim 2. Roll back. 3. Starvation	
96	Recoverability		If transaction fails we need to undo the effect of this transaction to ensure the atomicity property	
97	Types of Recoverability		Recoverable schedule Cascadeless schedule	
98	Types of failure		Transaction failure • Logical error • System error System crash Disk failure	
99	Shadow Paging		In this recovery technique, database is considered as made up of fixed size of logical units of storage which are referred as pages.	
100	Two Page tables		Current page table Shadow page table	
UNIT V CURRENT TRENDS				
101	Object Oriented Data Model	OODM	Encapsulation of Data and Code related to an object	
102	Object Query Language	OQL	It has been designed for use in Network Manager.	
103	Need for Complex Data Types		Traditional database applications in data processing had conceptually simple data types	
104	Data classification		The process of organizing data by relevant categories so that it may be used and protected more efficiently.	
105	Components of Information System		Data, Procedures, Hardware, Software, Network, People	
106	Threats in a Database		• Availability loss • Integrity loss • Confidentiality loss	
107	Vulnerability		A vulnerability refers to a known weakness of an asset (resource) that can be exploited by one or more attackers.	
108	Risk		Risk = Threat X Vulnerability	
109	Measures of Control		Access Control, Flow Control and Data Encryption	
110	Two types of intruders		Passive Eavesdroppers Active Attackers	
111	Access Control		Access control is responsible for control of rules determined by security policies for all direct accesses to the system.	

112	Privileges		Privileges defines the access rights provided to a user on a database object.	
113	Two Types of Privileges		1)System privileges 2)Object privileges	
114	Functions of Distributed D/B		1. Keeping track of data 2. Distributed query processing 3. Distributed transaction management 4. Replicated data management 5. Security	
115	Statistical Database		A database used for statistical analysis purposes, It is an OLAP (online analytical processing)	
116	Parallel Database		A parallel database system seeks to improve performance through parallelization of various operations.	
117	Spatial Database		A database that is optimized for storing and querying data that represents objects defined in a geometric space.	
118	Multimedia Database		A collection of related for multimedia data. Eg. Audio, Video, Image, Text etc..,	
119	Multimedia Applications		Repository applications Presentation applications Collaborative work using multimedia information	
120	Mobile Database		Portable and physically separate or detached from the corporate database server.	
121	Web Database		Data regarding E-commerce, E-business, Blog, e-mail and other online applications.	
122	XML Database		The data stored in the database can be queried using XQuery, serialized, and exported into a desired format.	
123	XML Documents		Data Centric , Document Centric, Hybrid XML Document	
124	Data warehouse	DW	A data warehouse is a large collection of business data used to help an organization make decisions.	
125	Data mining	DM	Data mining is defined as a process used to extract usable data from a larger set of any raw data.	
GATE QUESTIONS				
126	Relational Database Management System	RDBMS	A relational database is organized into tables, records, and column and there is a well-defined relationship between database tables.	
127	Index Hunting		A database index is a data structure that improves the speed of data retrieval operations on a database.	
128	Database Partitioning		Database partitioning is a process where a logical database is divided into distinct independent parts.	
129	Weak Entity Sets.		Entity set that do not have key attribute of their own are called weak entity sets.	
130	Strong Entity Sets.		Entity set that has a primary key is termed a strong entity set.	
131	Use of Sub Queries		A sub query is a select-from-where expression that is nested with in another query.	
132	Trigger		Triggers are statements that are executed automatically by the system as the side effect of a modification to the database	

133	Assertion		An assertion is a predicate expressing a condition that we wish the database always to satisfy.	
134	Authorization Graph		Passing of authorization from one user to another can be represented by an authorization graph.	
135	Levels in Security Measures		1. Database system 2. Operating system 3. Network 4. Physical 5. Human	
136	Encryption Techniques		a) DES b) AES c) Public key encryption	
137	Method for delay deadlock		**(i)** Deadlock prevention **(ii)** Deadlock detection and deadlock recovery	
138	Types Failures in a System		1. Transaction failure 2. System crash 3. Disk failure	
139	Page		The database is partitioned into some number of fixed-length blocks, which are referred to as pages.	
140	Operation of Files		OPEN, FIND, FINDNEXT, READ, INSERT, DELETE, MODIFY, CLOSE, REORGANIZE	
141	Hot Swapping		Hot swapping permits the removal of faulty disks and replaces it by new ones without turning power off.	
142	Search Key		An attribute or set of attributes used to look up records in a file is called a search key	
143	Fudge Factor		The number of partitions is increased by a small value called the fudge factor	
144	Data warehousing		It is the process that is used to integrate and combine data from multiple sources and format into a single unified schema.	
145	Inference Engine		An inference engine within the system can deduce new facts from the database by interpreting these rules.	
146	Persistence		Persistence is the property of an object through which its existence transcends time and/or space	
147	Object		An object has two components, one is state (value) and another is behavior (operations).	
148	Back Flushing		The process of returning cleaned data to the source is called back flushing.	
149	Data mining		Data mining refers to the mining of discovery of new information in term of patterns or rules from cast amounts of data	
150	Applications of Data Mining		1. Advertising 2.Store location 3. Targeted mailing 4. Segmentation of customer 5. Design of catalogs	

Faculty Team Prepared	1. Dr.J.Preetha, ASP/CSE/II/CSE/A 2. Mr.S.R.Sridhar, AP/CSE/II/CSE/B	Signature:

Subject Expert **HoD**