Data Base Management System

Assignment

Name : K. Soomanath Reddy
VTU No: 27832
Faculty: Dr. Thivya Lakshmi

5

1. Normalization and its various types of Normalization.

A) Normalization and its various types

Normalization is a process in database management system (DBMS) used to organise data in a database to reduce redundancy (duplicate data) and improve data integrity

i) First Normal Form (1NF)

→ Each cell must contain atomic (indivisible) values

→ No repeating groups or arrays allowed

Ex:

St_ID	St_Name	St_Phone	St_group.
123	Rakesh	9840723512	Physics
125	Smitha	9135462193 1341278963	stat
126	Archana	7365432107 8694532107	chemistry.

New table :-

St_ID	St_Name	St_Phone	St_group.
123	Rakesh	9840735l2	Physics
125	Smitha	9135462193	stats
125	Smitha	1341278963	stats.

Second Normal form (2NF).

→ In the 2NF, first table must be in 1NF
→ In the 2NF, all non key attributes are full functionally dependent on the primary key.
→ Every non key attribute should be FFD on key attribute.
If P→A holds, then those should not be any proper subset of Q of P
Q→A.

St-ID	St-Name	Prof-ID	Prof-Name	Grade
101	ABC	2	Sarees	4
102	xyz	3	Nirajan	6
103	PQR	1	Sushmitha	5

St-ID → Primary key and no multi valued are there so satistics

1NF

FD

St-ID → St-Name

St-ID	St-Name
101	ABC
102	xyz
103	PQR

FD

Prof-ID → Prof-Name

Prof-ID	Prof-Name
1	Sushmitha
2	Sarees
3	Nirajan.

Grades :-

St-ID	Prof-ID	Grade
101	2	4
102	3	6
103	1	5

3. Third Normal Form (3NF)

→ A relation will be in 3NF

→ If not contains any transitive dependency

The Non-key attributes should not have inter dependencies among them, and the non-key attributes should fully functionally depend on the key attribute.

Then it is called as 3NF principle

→ By using 3NF to achieve data integrity and data duplication.

Transitive dependency

If $A \Rightarrow B$ & B functionally dependent on A)

$B \Rightarrow C$

$A \Rightarrow C$ (C is indirectly dependent on A)

It is called T.D.

Ex:-

Stu_ID	Stu_Name	Dept Name	Dept_Head
101	Alice	CSE	Dr. Rao
102	Bob	ECE	Dr. Kumar
103	carol	CSE	Dr. Rao

Student table:

Student ID	Student_Name	Dept Name.
101	Alice	CSE
102	Bob	ECE
103	carol	CSE

Dept_table :-

Dept_Name	Dept_Head.
CSE	Dr RaO
ECE	Dr. Kumar
CSE	Dr. RaO.

4. BCNF (Boyce-Codd Normal form)

→ It is advanced version of 3NF
→ It is in 3NF
→ For every FD A→B A→ Super key
→ A should be super key of a table.

Ex :-

Student	course	Teacher.
Ramesh	Physics	Kishore
Kumar	chemistry	Ramu
Sruthi	maths	Sanjeev
Vinay	physics	Kishore

Keys→ {student, course} →
To eliminate redundant data we move to 2 tables.

Student	course
Ramesh	Physics
Kumar	chemistry
Sruthi	maths
Vinay	Physics

course	Teacher.
Physics	Kishore
chemistry	Ramu
Maths	Sanjeev.

4NF (Fourth Normal Form)

→ First relation in BCNF (or) 3.5NF

→ It may not contain more than one multivalued attributes.

For a dependency

$A \Rightarrow B$

→ If for a single value of "A" multiple values of B exists, then the relation will be multi-valued dependency.

Ex:-

St_ID	course	hobbies.
111	maths	Dancing
111	computer	singing
222	Chemistry	Dancing
444	sanskrit	cricket.

St-ID	course
111	maths
111	computers
222	chemistry
444	sanskrit

St-ID	hobby
111	Dancing
111	singing
222	Dancing
444	cricket.

2) Deadlock and its Handling?

A) A deadlock occurs in a database when two or more transactions are waiting for each other to release locks on resources, preventing further progress.

Example:

* Transaction T1 locks Resource A and waits for Resource B
* Transaction T2 locks Resource B and waits for Resource A

⇒ Both wait forever ⇒ deadlock

Conditions for Deadlock (coffman conditions)

1. Mutual Exclusion :- only one transaction can use a resource at a time.
2. Hold & wait :- A transaction holds one resource and waits for another.
3. No preemption :- Resources cannot be forcibly taken away.
4. Circular wait :- A circular chain of waiting transactions exists.

	Deadlock Handling Technique
Method	Description.
1. Deadlock Prevention	Avoids deadlocks by denying at least one coffman condition Example: Assign resource ordering or use time stamp based methods (wait-die, wound-wait).

2. Dead lock Avoidance	The system checks before granting a lock to ensure no deadlock will occur Ex: Banker's algorithm checks safe states.
3. Deadlock detection	The system allows deadlocks to occur but periodically checks for them using a wait-for graph (WFG)
4. Dead lock Recovery	once detected, resolve by rolling back one or more transaction (victim selection)

3. Explain about RAID storage and its types?

A) RAID (Redudant Array of independent Disks) is a data storage technology that combines multiple physical hard drives into a single logical unit.

- The main goals are:
  - → Improved performance
  - → Increased storage capacity
  - → Data redudancy (fault tolerance)

RAID is commonly used in servers, data centers, and storage systems to ensure data realiability and faster access.

2. Key features

Feature	Description:
Redundancy	Data is duplicated or spread to protect against drive failure
Performance	multiple disks can read write data simultaneosly
Fault Tolerance	System continues to function even if one drive fails.
striping	Data is split accross multiple disks to improve speed
Mirroring	Data is copied identically on two or more disks.
Parity	Error checking information is stored to recover data in case of disk failure.

Types of RAID:

RAID Level :- Raid 0 (striping)

Technique used :- Data divided into blocks and spread across disks.

Description :- Fast performance (no redundancy)

Advantages : High speed

Disadvantages : No fault tolerance. If one disk fails, all data lost.

RAID LEVEL :- RAID 1 (Mirroring)

Technique used :-
Some data copied on two disks.

Description :- Provides redundancy

Advantages :- High reliability, simple recovery

Disadvantages :- Storage cost doubles.

RAID Level :- Bit level striping with error correction (ECC)

Technique used :- Rarely used.

Description :- Rarely used

Advantages :- Error correction possible

Disadvantages :- Expensive, complex.

RAID Level :- RAID 3

Technique used :- Byte level striping with dedicated parity disk.

Description :- All drives work together.

Advantages : Good for large sequential data

Disadvantages :- Parity disk may be a bottleneck