

## Normalization

Date 12/10/2019

Expt. No. \_\_\_\_\_

Page No. 14

- ① It is used to reduce data redundancy (having the same data but at different places).
- ② There are basically two problems arises due to redundancy  
1st one is increases the size of the database, as the same data is repeated in many places.  
2nd is inconsistency problem arises during insert, delete and update operations.

1NF - If a database contains composite or multivalued attributes, it violates 1NF. A relation is in 1NF if every attribute in that relation is single valued attri.

2NF - A relation must be in 1NF and relation must not contains any partial dependency.

A relation is in 2NF iff it has no partial dependency. i.e. no nonprime attribute (attribute which are not part of any candidate key) is dependent on any proper subset of any candidate key of the table.

3NF - A relation is in 3NF there is no transitive dependency, for nonprime attribute is it is 2NF.

A relation is in 3NF iff at least one of the following condition holds in every nontrivial functional depen.

$$x \rightarrow y$$

(i)  $x$  is superkey

(ii)  $y$  is prime attribute (i.e. each element of  $y$  is part of some candidate key).

Teacher's Signature \_\_\_\_\_

Example: Consider the relation  $R(A, B, C, D, E)$

$$A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A$$

All possible candidate keys in the above relation are  
 $\{A, E, AD, BC\}$  all attribute are in right side  
of functional dependency are prime.

Transitive dependency : Suppose we have a relation

Table : student (stud\_no, stud\_name, stud\_phone, stud\_state,  
stud\_country, stud\_age).

FD are - ①  $\{stud\_no \rightarrow stud\_name\}$

②  $\{stud\_no \rightarrow stud\_state\}$

③  $\{stud\_state \rightarrow stud\_country\}$

④  $\{stud\_no \rightarrow stud\_age\}$

Candidate Key:  $\{stud\_no\}$

For the above relation -

$\{stud\_no \rightarrow stud\_state\}$  and  $\{stud\_state \rightarrow stud\_country\}$

are true. So  $\{stud\_country\}$  is transitively dependent on  $\{stud\_no\}$ . It violates the 3NF.

To convert it in 3NF, we will decompose the relation

Table1 student {stud\_no, stud\_name, stud\_phone, stud\_age}

Table2 State\_Country {state, country}

## Boyce-Codd Normal Form (BCNF)

A relation is in BCNF if R is in 3NF and for every functional dependency, LHS is superkey.

A relation is in BCNF iff in every nontrivial FD,  $x \rightarrow y$ ; x is a superkey.

Example : Find the highest normal form of a R(A,B,C,D,E) with FD set as  $\{ BC \rightarrow D, AC \rightarrow BE, B \rightarrow E \}$

Step 1 - As we can see  $(AC)^+ = \{ A, C, B, E, D \}$  but none of its subset can determine all the attribute of relation. So AC will be the candidate key. A and C can't be derived from any of the attribute of the relation. So there will be only one candidate key  $\{ AC \}$ .

Step 2 - Prime attribute are those attribute which are part of candidate key  $\{ A, C \}$  in this example and other are nonprime attribute  $\{ B, D, E \}$ .

Step 3 - The relation R is in 1NF as it does not allow multivalued or composite attribute.

The relation R is 2NF because  $BC \rightarrow D$  is in 2NF ( $BC$  is not a proper subset of candidate key  $AC$ )

$AC \rightarrow BE$  is in 2NF ( $AC$  is candidate key).

$B \rightarrow E$  is in 2NF ( $B$  is not a proper subset of  $AC$ )

The relation is not in 3NF because in  $BC \rightarrow D$   
(neither BC is superkey nor D is a prime attribute)  
in  $B \rightarrow E$  (neither B is a Superkey nor E is prime attr.)  
But to satisfy 3NF, either LHS of a FD should be  
superkey or RHS should be prime attribute).

So, the highest N.F. of above relation will be 2NF.

- ① BCNF is free from redundancy.
- ② If a relation is in BCNF, then 3NF also satisfied.
- ③ If all attribute of a relation are prime attribute  
then the relation is always in 3NF.
- ④ A relation in a RDBMS is always atleast in 1NF.
- ⑤ ~~A relation~~ Every Binary relation (A relation with only  
two attributes) is always in BCNF.
- ⑥ If a relation has only singleton candidate keys  
(i.e. every Candidate key consists of only one attribute)  
then the relation is always in 2NF (because no partial  
dependency possible).
- ⑦ sometimes going for BCNF form may not preserve FD,  
in that case go for BCNF only if the lost FD may not  
required, else normalize till 3NF only.

B. Find the highest NF in relation under  $R(A, B, C, D, E)$  following FDs -  $ABC \rightarrow D$ ,  $CD \rightarrow AE$ .

- SOL:
- ① Always checking from BCNF then 3NF and so on.
  - ② If any FD satisfies a NF then there is no need to check for lower NF.  $ABC \rightarrow D$  is in BCNF (Because  $ABC$  is superkey) so no need to check this dependency for lower normal form.

### \* Functional Dependency

FD is a constraints between two sets of attributes in a relation from a DB. A FD is denoted by ( $\rightarrow$ ) if an attribute  $A$  functionally determines  $B$ , then it is written as  $A \rightarrow B$ .

For example - empid  $\rightarrow$  name means empid functionally determines the name of employee.

$A \rightarrow B$  means for all instances of a particular value of  $A$  there is the same value of  $B$ .

### \* Trivial functional dependency

$X \rightarrow Y$  is trivial only when  $Y$  is subset of  $X$ .

$$ABC \rightarrow AB; \quad ABC \rightarrow A; \quad ABC \rightarrow BC.$$

Non Trivial FD -  $X \rightarrow Y$  is a nontrivial FD when  $Y$  is not subset of  $X$ .

## \* Some Important points on FD.

- ① For FD we find attribute closure using FD set F.
- ② If B is subset of A+, then  $A \rightarrow B$  is true.
- ③ Candidate key is minimal superkey.
- ④ Attributes which are part of any candidate key of relation are called prime attribute.

Q1 Consider a relation schema  $R = (A, B, C, D, E, H)$  on which the following FD holds -

$$\{ A \rightarrow B, BC \rightarrow D, E \rightarrow C, D \rightarrow A \}$$

What are the candidate key of R?

SOL  $(AE)^+ = \{ A, B, C, D, E \}$  which is not a set of all attributes  
so AE is not a candidate key.

$$(AEH)^+ = \{ A, B, C, D, E, H \} \quad (BEH)^+ = \{ A, B, C, D, E, H \}$$

$$(BCH)^+ = \{ B, C, H, D, A \}$$
 which is not set of all attributes

so, Ans is  $\{ AEH, BEH, DEH \}$

## Dependency Preserving in DBMS Decomposition

The dependency preserving decomposition is another property of decomposed relational DB schema D in which each functional dependency  $x \rightarrow y$  specified in F either appeared directly in one of the relational schemas  $R_i$  in the decomposed D or could

be inferred from the dependencies that appear in some  $R_i$ .

A decomposition  $D = \{R_1, R_2, R_3, \dots, R_n\}$  of  $R$  is dependency preserving w.r.t. a set  $F$  of FD  $(F_1 \cup F_2 \cup F_3 \cup \dots \cup F_m)^+ = F^+$ .

Consider a Relation  $R$ ,  $R \rightarrow F$  {with some FD?}

$R$  is decomposed of  $R_1$  with FD  $\{f_1\}$  and  $R_2$  with FD  $\{f_2\}$ ; there can be three cases:

- $f_1 \cup f_2 = f$ ; decomposition is dependency preserving.
- $f_1 \cup f_2$  is a subset of  $f$ ; Not dependency preserving.
- $f_1 \cup f_2$  is a superset of  $f$ ; this case is not possible.

Q Let a relation  $R(A, B, C)$  and a functional D.  $\{AB \rightarrow C, C \rightarrow D, D \rightarrow A\}$ . Suppose relation  $R$  is decomposed into  $R_1(A, B, C)$  &  $R_2(C, D)$ . Check whether a decomposition is dependency preserving or not?

Sol) Let us find the closure of  $f_1$  and  $f_2$ .

Closure of  $f_1$

Closure of  $(A)$  =  $\{A\}$  Trivial

Closure of  $(B)$  =  $\{B\}$  Trivial

Closure of  $(C)$  =  $\{C, A\}$

But  $D$  can't be closure as  $D$  is not present in  $R_1$ .

=  $\{C, A\}$

$C \rightarrow A$  Removing  $C$  from right side as it is trivial attribute.

Closure of  $(AB)$  =  $\{A, B, C, D\}$

$AB \rightarrow C$  =  $\{A, B, C\}$  (Incomplete)

Q.1

Let  $R(A, B, C, D)$  be a relational schema with the following FDs:  
 $A \rightarrow B$ ,  $B \rightarrow C$ ,  $C \rightarrow D$  &  $D \rightarrow B$ . The decomposition of  $R$  into  
 $(A, B)$ ,  $(B, C)$ ,  $(B, D)$ . Gives lossless join & is dependency preserving.

Ans:

Q.2

Let a relation  $R(A, B, C, D)$  and FDs  $\{AB \rightarrow C, C \rightarrow D, D \rightarrow A\}$   
Let  $R$  is decomposed into  $R_1(A, B, C)$  &  $R_2(C, D)$ . Check whether,  
the decomposition is dependency preserving or not?

Soln.

since  $f_1 \cup f_2 \subset f$  so, it is not. Canonical Cover

Whenever a user updates the DB, the system must check whether any of the FD are getting violated in this process. If there is a violation of dependencies in the new DB state the system must roll-back.

Working with a huge set of FD can cause unnecessary added computational time. This is where the Canonical cover comes into play.

A Canonical cover of a set of FD  $f$  is a simplified set of FD that has a same closure as the original set  $F$ .

Extraneous attribute  $\rightarrow$  An attribute of a FD is said to be extraneous, if we can remove it without changing the closure of the set of FDs.

A canonical cover  $F_C$  of a set of FDs  $F$  such that all the following properties are satisfied -

- (i)  $F$  logically implies all dependencies in  $F_C$ .
- (ii)  $F_C$  logically implies all dependencies in  $F$ .
- (iii) No FD in  $F_C$  contains an extraneous attribute.
- (iv) Each left side of FD in  $F_C$  is unique i.e. there are no two dependencies.

$\alpha_1 \rightarrow \beta_1$  &  $\alpha_2 \rightarrow \beta_2$  in such that  $\alpha_1 \rightarrow \alpha_2$ .

Example - consider the following set of FD  $F$  -

$$f = \{ A \rightarrow BC, AB \rightarrow C, B \rightarrow C, A \rightarrow B \}$$

steps to find the canonical cover.

Soln - 1. There are two FD with the same set of attribute on the left hand side i.e.  $A \rightarrow BC$ ,  $A \rightarrow B$  these two can be combined to get  $A \rightarrow BC$ .  
Now, the combined set becomes

$$f = \{ A \rightarrow BC, AB \rightarrow C, B \rightarrow C \}$$

2. There is an extraneous attribute in  $AB \rightarrow C$  because even after removing  $AB \rightarrow C$  from the set we get the same closure. This is because  $B \rightarrow C$  is already a part of  $f$ .

Now  $f$  becomes

$$f = \{ A \rightarrow BC, B \rightarrow C \}$$

Teacher's Signature \_\_\_\_\_

3. c is an extraneous attribute in  $A \rightarrow BC$   
also,  $A \rightarrow B$  is logically implied by  $A \rightarrow B$  &  $B \rightarrow C$   
(by transitivity) so,  $F = \{A \rightarrow B, B \rightarrow C\}$

4. After this step, f does not change anymore  
hence the required Canonical Cover is

$$f_C = \{A \rightarrow B, B \rightarrow C\}$$

Ans:

### Equivalence of FD

For understanding equivalence of FD sets basic idea about attribute closure is -

Given a relation with different FD sets for that relation, we have to find out whether one FD set is subset of other or both are equal

→ How to find relationship between two FD sets ↗

Let  $FD_1$  &  $FD_2$  are two FD sets in a relation R,

(i) If all FDs of  $FD_1$  can be derived from FD present in  $FD_2$ , we can say that  $FD_1 \supseteq FD_2$

(ii) If all the FDs of  $FD_2$  can be derived from FD present in  $FD_1$ , we can say that  $FD_2 \supseteq FD_1$

(iii) If both are true, then  $FD_1 = FD_2$

# Concurrency Control

Date 15/10/2019

Expt. No. \_\_\_\_\_

Page No. \_\_\_\_\_

## Serializability

If any transaction (non serial) produce an outcome which is equal to the outcome of that schedule's transaction executed serially, then we can call that transaction schedule, a serializable schedule.

- (i) serial schedule have less resource utilization and low throughput. To improve this we use concurrency.
- (ii) Concurrency <sup>of transaction</sup> may lead to inconsistency in database.
- (iii) To avoid this, we need to check whether these concurrent schedules are serializable or not.

These are two types

view serializable  
conflict serializable

### conflict serializable

A schedule is called conflict, if it can be transformed into a serial schedule by swapping non conflicting operation.

Two operations are said to be conflicting if all conditions satisfy :-

- (i) They belongs to different transactions.
- (ii) They operate on same data.
- (iii) Atleast one of them is a write operation.

Ex: conflicting operations pair [ R<sub>1</sub>(A), W<sub>1</sub>(A) ]

because they belongs to two different transaction on same data A and one of them is write.

Teacher's Signature

## Some Important points

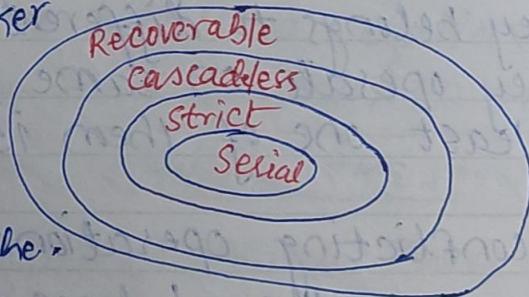
- To check for conflict serializable, we need to make precedence Graph. If the graph contains the cycle, then its not conflict serializable. Hence their is no cycle the schedule is conflict serializable.
- Now to check for recoverable, we need to check for a dirty read operations ( write by  $T_i$  transactions followed by  $T_j$  read operation before  $T_i$  Commit ), between any pair of operations. If no dirty read operation the recoverable schedule.

→ If there is no interleaving of operation the conflict.

Q. Which of the following scenarios may lead to an irrecoverable error in a database system.

Ans A transaction read a data item after it is written by an uncommitted transaction.

- ① Cascadele schedule is stricter than recoverable schedule.
- ② Strict schedule is stricter or subset of cascadeless sche.
- ③ Serial schedule Satisfy Constraints of all schedule recoverable, cascadeless and strict schedule & hence it is a subset of strict schedule.



- There are basically two reasons for allowing concurrency
- Improved throughput and resource utilization.
  - Reduced waiting time also reduce avg. response time.
- The motivation for concurrent execution in DB is same as using multiprogramming in OS.

Example 1: Consider the following schedule

T <sub>1</sub>	T <sub>2</sub>	T <sub>3</sub>
R(A)		
	w(A) commit.	
w(A)		w(A) commit.

Ans. (i) First of all it is view serializable as view equal to serial schedule  $T_1 \rightarrow T_2 \rightarrow T_3$ .

which satisfies the initial read & updated read and final write on variable A.

(ii) Now there is a Write-Write pair done by transaction T<sub>1</sub> & T<sub>3</sub>, which is violating the above mention conditions of strict schedule. So, The given schedule is serializable but not strict recoverable.

## Two-phase Locking Protocol (2PL)

A simple lock based protocol does not guarantee serializability (Also known as binary lock).

To guarantee serializability, 2-phase locking is used that concern the positioning of locking and unlocking operations in every transactions.

" 2-Phase locking ensures serializability "

A transaction is said to follow 2PL protocol if locking & unlocking can be done in 2-Phase -

- (i) **Growing Phase** - New locks on data items may be acquired, but none can be released.
- (ii) **shrinking phase** - Existing locks may be released but no new locks can be acquired.

$[S(a)] \rightarrow$  shared lock ;  $X(a) \rightarrow$  Exclusive locks.

\* If lock conversion is allowed then upgrading of lock from  $S(a)$  to  $X(a)$  is allowed in growing phase.

And

\* Downgrading of lock from  $X(a)$  to  $S(a)$  must be done in shrinking phase.

\* **strict 2PL** - It ensures that the schedule is recoverable and cascadeless but deadlock is possible.

\* **Rigorous 2PL** - Lock being 2Phase, all exclusive( $X$ ) and shared( $S$ ) LOCKS. It is more restricted than strict 2PL.

## Example

Expt. No. \_\_\_\_\_

strict 2PL  
Rigorous 2PL

Page No. \_\_\_\_\_

T1

T2

1 LOCK-S(A)

2 -

3 LOCK-X(B)

4 -

5 unlock(A)

6 -

7 unlock(B)

8 -

9 -

10

LOCK-S(A)

-

LOCK-X(C)

-

unlock(A)

unlock(C)

Transaction T1 → Growing Phase from 1-3.  
Shrinking Phase from 5-7.  
Lockpoint at 3.

Transaction T2 → Growing Phase from 2-6.  
Shrinking Phase from 8-9.  
Lockpoint at 6.

Lockpoint → The point at which growing phase ends.

Drawbacks of 2-Phase Locking Protocol

- Cascading rollback is possible under 2PL.
- Deadlocks and starvation is possible.

Teacher's Signature \_\_\_\_\_

## Timestamp Ordering Protocol

- (1) It is used to order the transaction based on their timestamps. The order of transaction is nothing but the ascending order of the transaction creation.
- (2) The priority of the older transaction is higher than why it executes first. To determine the timestamp of the transaction, this protocol uses system time or logical counter.
- (3) The lock based protocol is used to manage the order between conflicting pairs among transactions at the execution time.
- (4) Timestamp based protocol starts working as soon as transaction is created.
- (5) It ensures conflict serializable & free from deadlock.
- (6) One Drawback of basic TO is that it cascading rollback is possible. Also may not be recoverable.

Granularity :  
It is the size of the data items allowed to lock. Now multiple granularity means hierarchical breaking up the database into ~~database~~ blocks which can be locked and track.

## multiple choice Questions

Date \_\_\_\_\_

16/10/19

Expt. No. \_\_\_\_\_

Page No. \_\_\_\_\_

- ① Rigorous 2PL protocol permits releasing all locks at the end of transaction.
- ② A scheme that creates a new version of a data item for each transaction is defined by multiversion concurrency control scheme.
- ③ A 2Phase locking protocol variant that requires all locks to be held until transaction commit is called Rigorous 2-phase locking protocol.
- ④ Deadlock prevention scheme that requires each transactions to locks all data items before it begins execution.
- ⑤ Deadlock prevention scheme named wound wait is a preemptive technique.
- ⑥ A transaction that is inserting a new tuple into DB is given an Exclusive Lock.
- ⑦ multiple granularity locking protocol uses multiple lock modes to ensure serializability.
- ⑧ cascading rollback can be avoided by strict 2PL Prop.
- ⑨ A process known to receive msgs from transaction and in response send msg is - Lock manager.
- ⑩ In a Preemptive technique, transactions are assigned a unique timestamp to decide whether it should wait or rollback.
- ⑪ snapshot isolation is a scheme of multiversion concurrency control.
- ⑫ Two modes of locking data items, are termed as shared & exclusive.

Teacher's Signature \_\_\_\_\_

- (13) A set of rules applied over a transaction that may lock & unlock each of data items in DB, is known to be **Locking protocol**.
- (14) A protocol that permits a transaction to lock a new data items only if it has not yet unlocked any data items is called **2PL protocol**.
- (15) Exclusive locks are released at the end of transaction to ensure - **Recoverability & cascadelessness**.
- (16) For dealing with **deadlocks**, system construct **wait-for graph**.
- (17) Snapshot isolation scheme does not support **Serializability**.
- (18) No. of actions need to be taken in order to recover from deadlock is/are - **3**.
- (19) A way of guarding data items against **deadlock prevention** is to - **order data items**.
- (20) To ensure deadlock prevention, we can use preemptive approach and **Rollbacks**.
- (21) A write operation in multiversion timestamp ordering may result into the - **Rollback of transaction**.
- (22) A protocol that requires each transactions to be executed in 2 or 3 diff. phases in its lifetime is called **validation protocol**.
- (23) multiversion 2PL protocol requires transactions to be declared as **Readonly & updates**.

(2) Deadlock prevention scheme named **wait-Die** is a non preemptive scheme.

Date \_\_\_\_\_

Expt. No. \_\_\_\_\_

## SQL commands

Page No. \_\_\_\_\_

**DDL**

- create
- Drop
- Alter
- Truncate
- drop

**DML**

- insert
- update
- delete

**DCL**

- Grant
- Revoke

**TCL**

- commit
- Rollback
- Savepoint

**DDL**

select

### (1) DDL (Data Definition Language)

- It changes the structure of the table like creating a table, deleting a table, altering a table etc.
- It is auto-committed i.e. permanently save all the changes in the database.

Truncate - It is used to delete all the rows from the table & free the space containing the table.

Ex: **Truncate table employee;**

### (2) DML (Data manipulation language).

- It is used to modify the DB.
- It is not auto-committed that means it can't permanently save all the changes in DB. They can be rollback.

### (3) TCL (transaction control language)

It is only used with DML like insert, delete etc.

Savepoint — It is used to roll the transactions back to a certain points without rolling back the entire transactions.

Example ↴ `SAVEPOINT savepoint_name;`

### SQL operators

Arithmetic operator

Comparison operator

Logical operator

### SQL SET operations

union

union all

intersection

minus

→ minus operators are used to display the rows which are present in the 1st query but absent in the second query.

### SQL Aggregation function

Count

Sum

Avg

Max

Min

→ Count function is used to count no. of rows in a DB table. It works on both numeric & non numeric datatypes.  
→ Count(\*) returns all the rows also considers duplicated Null.

\* Data Stripping → data can place across multiple disks.  
If one disk fails, then all data is lost.

Expt. No. \_\_\_\_\_

RAID

Page No. \_\_\_\_\_

- It refers to Redundancy array of the independent Disk.
- It is a technology which is used to connect multiple secondary storage devices for increased performance, data redundancy or both.
- There are seven levels of RAID schemes –  
RAID 0, RAID 1, ..., RAID 6.

RAID 0 It provides data stripping.

RAID 1 Mirroring of data

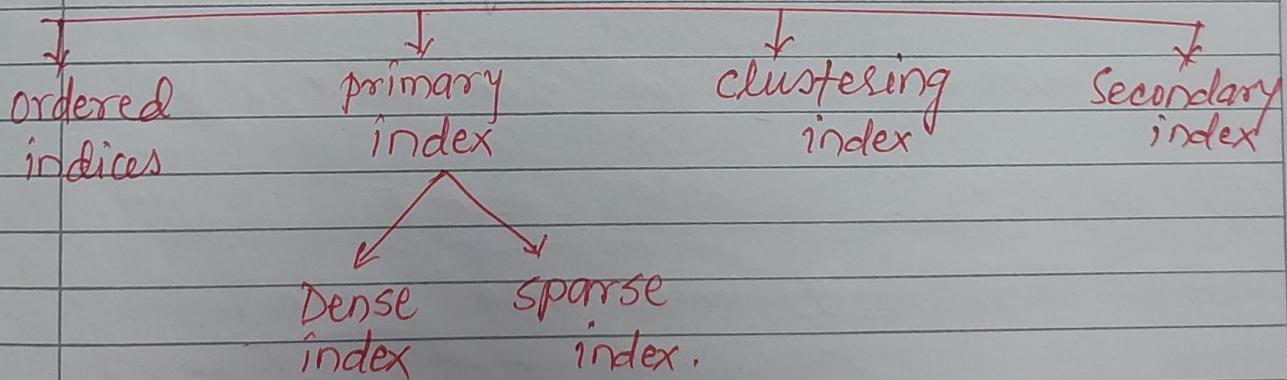
RAID 2 bit level stripping using hamming code parity.

RAID 3 Byte level stripping with dedicated parity.

RAID 4 Block level stripping with parity disk.

RAID 5 Slight modification of RAID 4, distributed parity.

RAID 6

Indexing

- Indexing is used to optimize the performance of the DB by minimizing the access of the data in the DB table quickly.

Dense index - It is primary index based on primary key. The dense index contains an index record for every search key value in the data file. It makes searching faster.

The no. of records in the index table is same as the no. of records in the main table. Therefore it needs more space to store index record itself.

The index record have searchkey and a data reference/pointer to the actual record in the disk.

Index record

Search key | data reference/pointer

Sparse index

- In the data file index record appears only for a few items. each item points to a block. The index points to the records in the main table in gap.

- ① The B+ tree is a balanced binary search tree.  
It follows a multilevel index format.
- ② In this, leaf node denote actual data pointers.  
B+ tree ensures that all leaf node remain at the same height.
- ③ Leaf nodes are linked using link list. Therefore  
B+ tree can support random access as well as sequential access.
- ④ An internal node of B+ tree can contain at least  $n/2$  record pointers except the root node.  
At most, an internal node of the tree contains  $n$  pointers
- ⑤ The leaf node of the B+ tree can contains at least  $n/2$  record pointers and  $n/2$  key values.

# View Serializability

Date  
08/01/2020

## view equivalence

reads get the same view

Let  $t_1$  &  $t_2$  have two transactions

✓ 1.  $x \quad r_1(x) \quad r_2(x)$  initial read

✓ 2.  $x \quad w_1(x) \quad w_2(x)$  final write

3.  $x \quad T_1 \leftarrow T_2$

$T_1 \leftarrow T_2$

two schedules

view equivalence to some serial schedule.

Ex: ①  $S: r_1(a) \quad w_1(a) \quad r_2(a) \quad w_2(a) \quad r_1(b) \quad w_1(b)$

$r_2(b) \quad w_2(b)$ .

$T_1 \quad T_2 : \quad r_1(a) \quad w_1(a) \quad \overset{\text{initial}}{r_1(b)} \quad \overset{\text{initial}}{w_1(b)} \quad \overset{\text{initial}}{r_2(a)} \quad \overset{\text{final}}{w_2(a)}$

$r_2(b) \quad w_2(b)$ .

that's why it is view serializable.

Ex: ②  $S: r_1(a) \quad w_2(a) \quad w_1(a) \quad w_3(a)$

$T_1 \quad T_2 \quad T_3 : \quad r_1(a) \quad (\overset{\circ}{w_1(a) \quad w_2(a)}) \quad \overset{\circ}{w_3(a)}$

view serializable.

$\times T_2 T_1 : r_2(a) \cancel{w_2(a)} r_2(b) w_2(b) \cancel{r_1(a)} w_1(a) r_1(b) w_1(b)$

Date \_\_\_\_\_

Expt. No. \_\_\_\_\_

Page No. \_\_\_\_\_

Ex. ③

$S: r_1(a) w_2(a) w_1(a)$

$T_1 T_2 : \cancel{r_1(a)} w_1(a) w_2(a) \times$

$T_2 T_1 : w_2(a) \cancel{r_1(a)} w_1(a) \times$

not view serializable

conflict vs view serializable

Note: ① Every conflict serial. is also view serializable.  
i.e. conflict serial. is stronger than view serial.

② But Every view S. is not conflict serial.

③ Constrained write assumption x: C.S.  $\equiv V.S.$

$w_i(\delta(r(x)))$

unconstrained writers / blind writers

Conflict Serializability.

$i_i \neq i_j$  conflict  $t_i \neq t_j$  logical temporal order.

1. Access same data item

2. at least one of them is a write operation.

- conflict serializability

$S'$ : conflict equivalence  $S'$

Ex:  $S: r_1(a) w_1(a) r_2(a) w_2(a) r_1(b) w_1(b) r_2(b) w_2(b)$

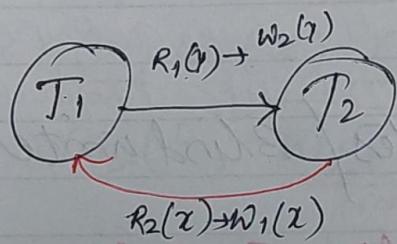
$T_1: T_2: r_1(a) w_1(a) r_1(b) w_1(b) r_2(a) w_2(a) r_2(b) w_2(b)$

→ Any view serializable schedule that is not conflict serializable, must contains a Blind write. " Blind write is simply when a transaction writes without reading".

→ Blind write happens in view serializability.

Ex: ①  $R_1(x) R_1(y) R_2(z) R_2(y) \underline{w_2(y)} \underline{w_1(x)}$

precedence  
Graph



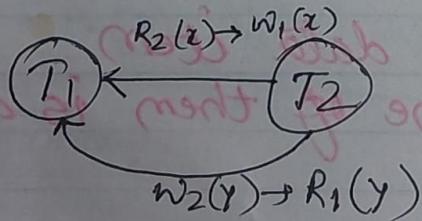
cyclic

Hence

Non conflict  
serializable

Ex: ② -  $R_1(x) R_2(x) \underline{w_2(y)} R_1(y) w_1(x)$

precedence  
Graph



acyclic

conflict serializable  
order  $T_2 : T_1$

(Q) Which normal form is considered as adequate for Expt. No. usual DB design - 3NF  
Page No.

- ① A table on the many side of a one to many or many to many relationship must:

Ans: Have a composite key. The relation in 2NF is also in 1NF and no partial dependencies on any column in primary key.

- ② Table is 2NF - Eliminate all hidden dependencies.  
 ③ Which one of the following statements about normal forms is false : → ④ BCNF is stricter than 3NF  
 ⑤ Lossless, dependency preserving decomposition into 3NF is always possible.  
 ✓ ⑥ Lossless, dependency preserving decomposition into BCNF is always possible.  
 ⑦ Any relation with two attributes is BCNF.

- ⑧ Which is bottom-up approach to db design that designs by examining the relationship between attributes -  
 ⑨ FD ⑩ Database modeling ⑪ Normalization ⑫ Decomposition

- ⑬ A table is in 2NF if it is in 1NF and if:  
 no column that is not a part of the primary key is dependent on only a portion of the primary key.

- ⑭ A normal form on which every determinant is a key BCNF. A relation is in BCNF if every determinant of the relation is a candidate key.

A relation is in BCNF if it is not in 4NF.

## Precedence Graph

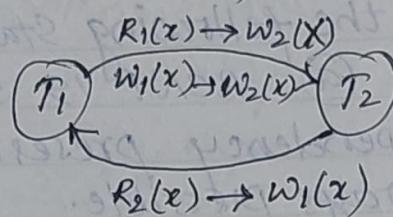
$R_1(A) \rightarrow w_2(A)$  Arrow from  $T_1 \rightarrow T_2$

$w_1(A) \rightarrow R_2(A)$  Arrow from  $T_1 \rightarrow T_2$

$w_1(A) \rightarrow w_2(A)$  Arrow from  $T_1 \rightarrow T_2$

Ex: ①.  $D = R_1(x) R_2(x) w_1(x) R_1(y) w_2(x) w_2(y)$

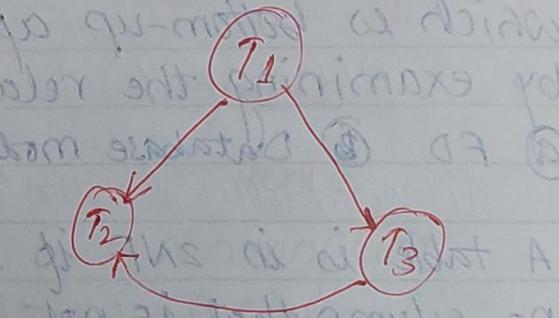
$T_1$	$T_2$
$R(x)$	
	$R(x)$
$w(x)$	
	$w(x)$
$R(y)$	
	$w(x)$
$w(y)$	



cyclic Hence Not Conflict  
Serializable.

Ex: ②

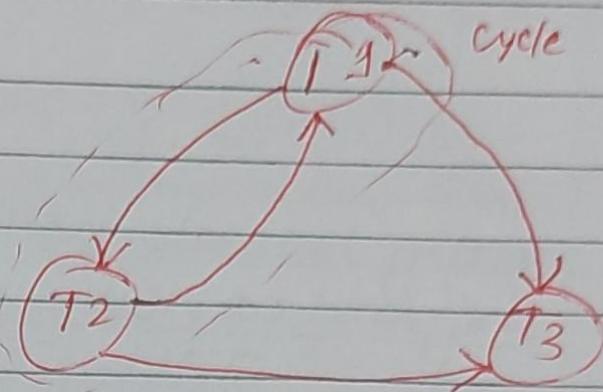
$T_1$	$T_2$	$T_3$
$R(A)$		
$R(B)$		
	$R(A)$	
	$R(C)$	
$w(B)$		
commit		
	$R(B)$	
	$R(C)$	
$w(B)$		
commit		
		$R(A)$
		$R(C)$
$w(A)$		
$w(C)$		
commit		



A cyclic then conflict  
Serializable.

Ex: ③

T1	T2	T3
R(A)		
	R(B)	
	W(A)	commit
w(A) commit		
		w(A) Commit



cycle: Hence not conflict serializable.

Q:

Consider the following two statements about transaction schedules:

- I Strict 2-Phase locking protocol generates conflict serializable schedules that are also recoverable.
- II Timestamp ordering concurrency control protocol with Thomas "Write rule can generate view serializable schedules that are not conflict serializable.

Ans: Both statements are true.