Date
Page

DBMS

1) introduction : Read from gfg.

- before DBMS we use filesystem in os but it has lots of DisAdvantage then DBMS comes into picture.

2)

OLAP	OLTP
online analytical	online transaction.
Historical data	current Data
subject oriented	Application oriented
decision making	Day to Day operation
TB PB	mB HB

we have to devide we All Data or we conside All Data available on internet is devided into Historical Data & current data

2) ER Diagram - Read from Wfg.

introduced by D & Peter Chen in 1976.

- Entity: identify uniquely in world
- tangible - Physically present - person
- intangible - virtually - Account.

5) Functional dependency

A functional dependency $A \rightarrow B$ a relation holds if two tuples having same value of attribute A then B Attribute must have same value.

$$X \rightarrow Y$$

X — determinant, Y — dependent

X	Y
1	P
2	P

X	Y
1	P
1	P

incorrect in FD

X	Y
1	P
1	B

X	Y
1	P
2	B

Same value of X we can not get different values of Y. Not Allowed

FD

Trivial — $XY \rightarrow Y$ — subset of left part

Non trivial — $X \rightarrow Y$ — not subset of left part.

Date
Page

Attribute closee / closee on attribute set / closee of attribute set.

Attribute closee of on attribute set A can be defined as a set of attribute which can be functionally determined from it denoted by f^+.

means

$A \rightarrow B$
$C \rightarrow DE$
$AC \rightarrow F$
$D \rightarrow AF$
$E \rightarrow CF$

Compute $(D)^+$

$= (D)^+ = Af$
= Now we have A
= so we can write B.f

$(D)^+ = ABDF$

$(DE)^+ = DE$
$= ADEF$
$= ACDEF$
$(DE)^+ = ABCDEF$

} All Possible Dependency means closee.

Armstrong axiom rules:

axiom is a statement that is taken to be true and serve as a premise or starting point for function Arguments.

Armstrong axiom hold on every relational database and can be used to generate closer set.

Rules Devide in two part.

1) Primary Rules

1) Reflexity

if $Y \subseteq X$ then $X \rightarrow Y$

2) Augumentation

if $X \rightarrow Y$ then $XZ \rightarrow YZ$

we can Apply Z to both side.

3) Transitivity

if $X \rightarrow Y$ && $Y \rightarrow Z$

then $X \rightarrow Z$.

2) Secondary Rules

1) union

if $X \rightarrow Y$ && $X \rightarrow Z$ then $X \rightarrow XY$.

2) Decomposition

$X \rightarrow YZ$

then $X \rightarrow Y$ && $X \rightarrow Z$

3) Pseudo transitivity

$X \rightarrow Y$ && $WY \rightarrow Z$

then $WX \rightarrow Z$

4) composition

$X \rightarrow Y$ && $Z \rightarrow W$

then

$XZ \rightarrow YW$

* Ireducible set of fd (canonical form)

यो dependency redundent है उसलाई remove गर्नु.

Check out following Example:

if you calculat dependecy then in that dependency there is some redudent dependecy posible thats why we calculate canonical cover.

R(wxyz)

$y \to w$
$z \to xy$
$y \to wxy$

we can also write like this from Above by using Decomposition Rule.

$x \to w$
~~$wz \to x$~~
$wz \to y$
~~$y \to w$~~
$y \to x$
$y \to z$

① $(x)^+ = xw$ } both diff
$(x)^+ = x$ }

$x \to w$ is essential.

② $(wz)^+ = wzxy$ }
$(wz)^+ = wzxy$ } means $wz \to x$ is Redundent.
$(wz)^+ = wz$ }

③ $(y)^+ = ywxz$ } both same means
$(y)^+ = yxzw$ } $y \to w$ is Redundent.
$(y)^+ = yz$
$(y)^+ = ywx$

so $x \to w$
$wz \to y$
$y \to x$
$y \to z$
} = $x \to w$
$wz \to y$
$y \to xz$
} final.

in this there is no Redudant values or dependencys.

Ex: R(ABCD)

$A \to B$
$C \to B$
$D \to ABC$
$AC \to D$

Decompose

$A \to B$
$C \to B$
$D \to A$
$D \to B$
$D \to C$
$AC \to D$

how To check which is redundant.

① find closeof that dependency

$(A)^+ = AB$

② ignore first dependency then Again find Dependency closee

$(A)^+ = A$ means we can not calculate. Dependency mens same calculate नहीं कर पा रहे है. so $A \to B$ is not Redundant.

अगर $(A)^+ = AB$ After ignoring.
Again $(A)^+ = AB$ मिलता then $A \to B$ is Redundant. होता.

Date
Page

☆ Keys in DBMS

student table

student No	Name	phone	state	AGE	country
1	ABC	1234	XYZ	20	india
2	PKB	5678	PSD	30	india
3	SDK	2822	DSD	46	india
4	CDK	2484	CSD	60	india

course table

stud NO	course NO	course Name
1	C1	PSDF
2	C2	KDPB
1	C3	CSDB

primary key :

A primary key is A column that is used to uniquely identify the row in table. studen NO in above table, phone number student id.

candidate key

minimum set of Attribut that can uniquly identify tuple in table.

Ex : student no is Also candidate key
student no + phone no this is Also candidate key.

③ Alternate key:

The candidate key other than the primary key is called Alternate key.

Ex student no + Phone no

④ Foreign key:

The referenced attribute of the relation should be the primary key.

foreign key means given reference of primary key to other table.

student-No in course table is foreign key. to student_No in student table.

foreign key is the columns of a table that points to the primary key of another table.

Date
Page

super key : super key is basic key one or more columns which can be used for identify data in table.

① super key ② candidate key ③ primary ④ foreign.

- Every super key is not candidate key.
- minimum set of Attribut which can uniqly identify the row.

student-no is candidate key
sundeep-no + phone = candidate but.

we use student-no = minimum.

Date
Page

insertion & deletion & modification Anomalies

insertion: when certain data can not be inserted into A database without the presence of Another data.

उदाहरण table is college table. And यामे चाहते है insert branch details but you can not insert ~~because~~ without student.

deletion:

if we want to delete certain unwanted data ~~but~~ but with that deleted some that that is important data that is called Deletion Anomies.

modification:

we want to delete or update single peace of Data it must be done on all of its copyes.

4. Normalization

Normalization is the process which is used to minimize redundency because redudency is cause of All problems.

Normalization is Decomposition of table which based on functional dependency.

★ 1NF

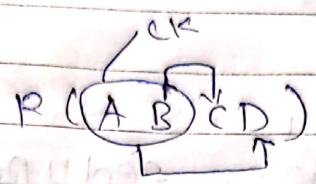
- by default every table is in 1NF
- you can not have multivalued Attribute
- if you have the devide that into multiple rows.
- column All values is from same domain.
- every column name is must diff.

Roll-NO	name	couese.
101	P K B	ABC, 05 - multivalued
102	B K P	CBA

Roll NO	name	couese.
101	P K B	ABC
101	P K D	05
102	B K P	CBA.

Date
Page

2NF:



- Table must be in 1NF
- ~~in table there is~~ .
- partial dependency is not Allowed.

- prime Attribute: Attribut part of CK = AB
- NON prime Attribute: Not part of CK = CD
- partial dependency: B → C

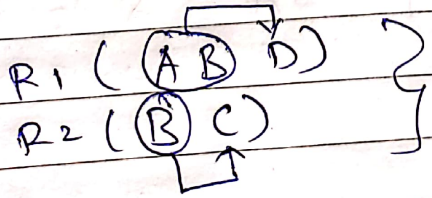
- partial dependency means the NON prime Attribut depend on part of candidate key like B → C that is called partial.

how to identify 2NF:

① must be in 1NF
② NO partial dependency
③ A B both Not Allowed to be null.
if one value is null that Allowed.

how to convert in 2NF

① R (A B C D) Decompose:



} 2NF form.

(2) R(A B C D E)

(A B D) is CK. Because D → E Partially Dependency

Decompose.

R1(A B C)

R2(D E)

R3(A B D) = This is must because full key पे table होनी चाहिए.

★ 3NF

- Table must be in 2NF
- NO transitive Dependency.

R(A B C) ← transitive hai

Decompose.

R1(A → B)

R2(B → C)

Date
Page

★ BCNF (Boyce codd Normal form)

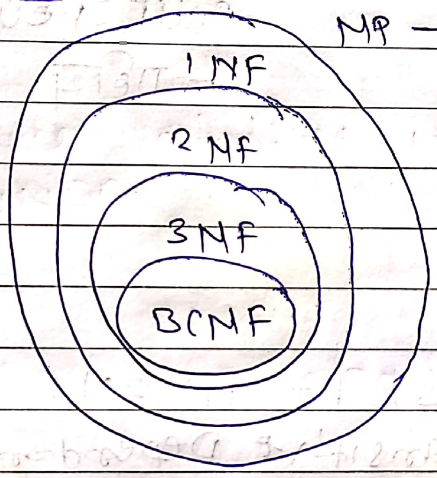
- it should be in 3NF
- following is not Allowed

$A \rightarrow B$ (P / NP → P)

P = Prime
NP = Non-Prime.

→ R(A B C). Not in BCNF

NP → P



- if $A \rightarrow B$ A must be super key

How to remove convert into BCNF

→

A	B	C
a	1	X
b	2	Y
c	2	Z
c	3	W
d	3	W
e	3	W

⟹

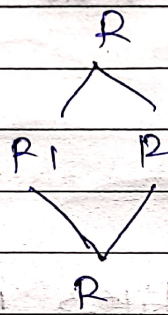
A	C
a	X
b	Y
c	Z
c	W
d	W
e	W

C	B
X	1
Y	2
Z	2
W	3

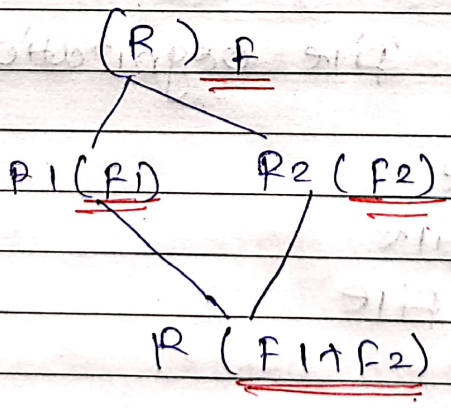
Date
Page

lossi & lossless Decomposition.

if we devide table in more table and After combile All tables we should have to get same table As previous that is called lossless if previous and After table is not same the it will be lossi



★ Dependency Preserving Decomposition in DBMS



if we devide table into multiple tables table must be same After we combine Again but if dependencies Also same then that dependency is is Dependecy Preserving.

5. File Structures // Read from gfg.

A data base consist of a huge amount of data. The data is grouped within a table in RDBMS and each have record.

A user can see that the data is stored in form of tables but in actual this huge amount of data is stored in physical memory in the form of file.

File organization:

mangement of files is called file organization.

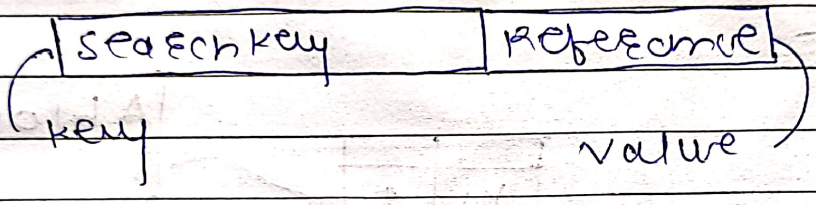
Types of file organization

* sequential file organization
Heap file
Hash file
B+ Tree file
clustered file

Read here & Read on gfg.

6. indexing of B and B+ trees.

indexing is the way to optimize performance of database by minimizing the number of disk accesses required when a query is processed.



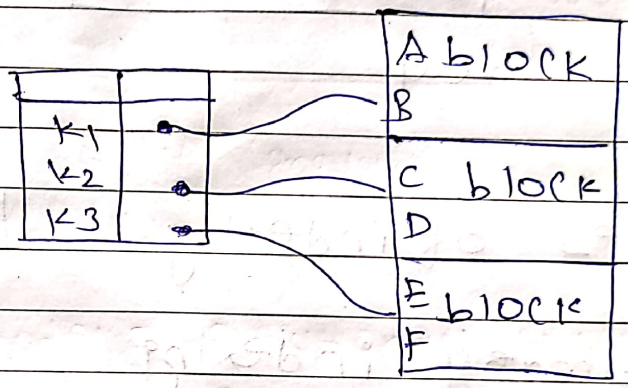
Types of indexing : // Read from gfg.

1. primary indexing
2. clustered
3. secondary
4. multivalued
5. dense
6. pase.

① primary indexing

- main file sorted must
- primary key of file is used to make indexing.
- Example of Spark indexing.

- no of entries in index file ↑ = no of blocks acquired by the main file ↑
- no of Access (block पर पहुंचने में) = log n. by using Binary Search.

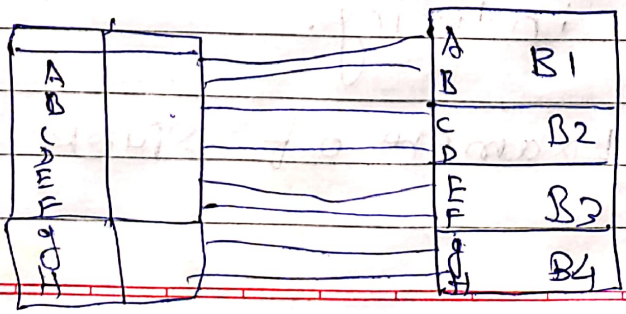


(1) What is Spare index?

Only block Addre are there Pointe are there As you see Above.

(2) Dense indexing?

for every Search key in data file. there is an index record.



Date
Page

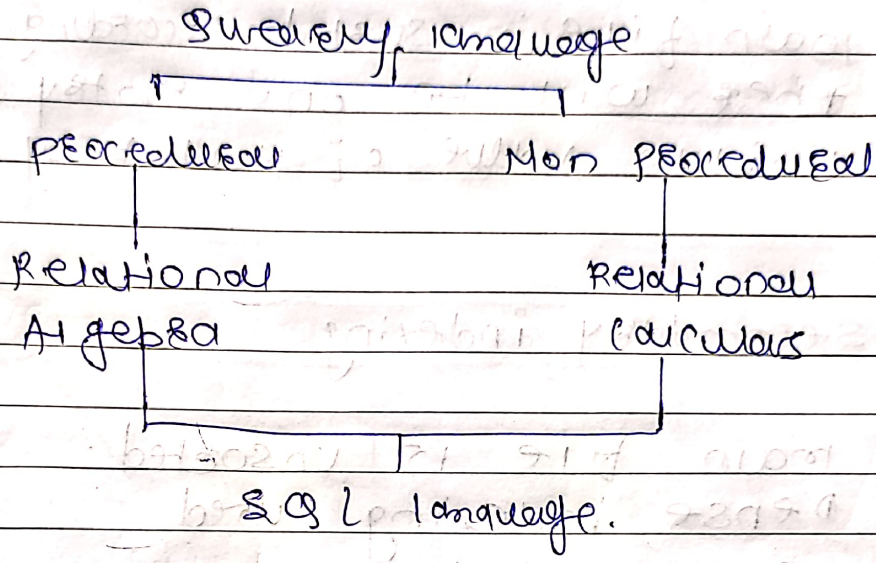
② clustred indexing:

- main file is sorted According to NON key
- there will be one entry for each unique value of NON key attribute.

③ Secondary indexing

- main file is unsorted.
- Dense indexing used.
- key or NON key any one used.

6. Relational Algebra.



- Relational model = Relational Algebra. } Algo
  Relational Calculus

- RDBMS = code (SQL)

- That's why we have to learn Relational model.

Date
Page

7. Transaction & concurrency control

Transaction is set of logical instructions. by default transaction is Atomic in nature means if that will complete or Not complete.

ACID Properties:

(1) Automacity: if transaction start it must be completed if that fails inbetween that must be Rollback अगर होगा तो complete होगा नहीं तो नहीं होगा.

(2) Isolation: if At a time 10 transaction then that must be Different Affect one transaction must not be on another transaction.

(3) Durability:
Transaction must be Durable if ones we perform transaction and Chan made them that must be constent.

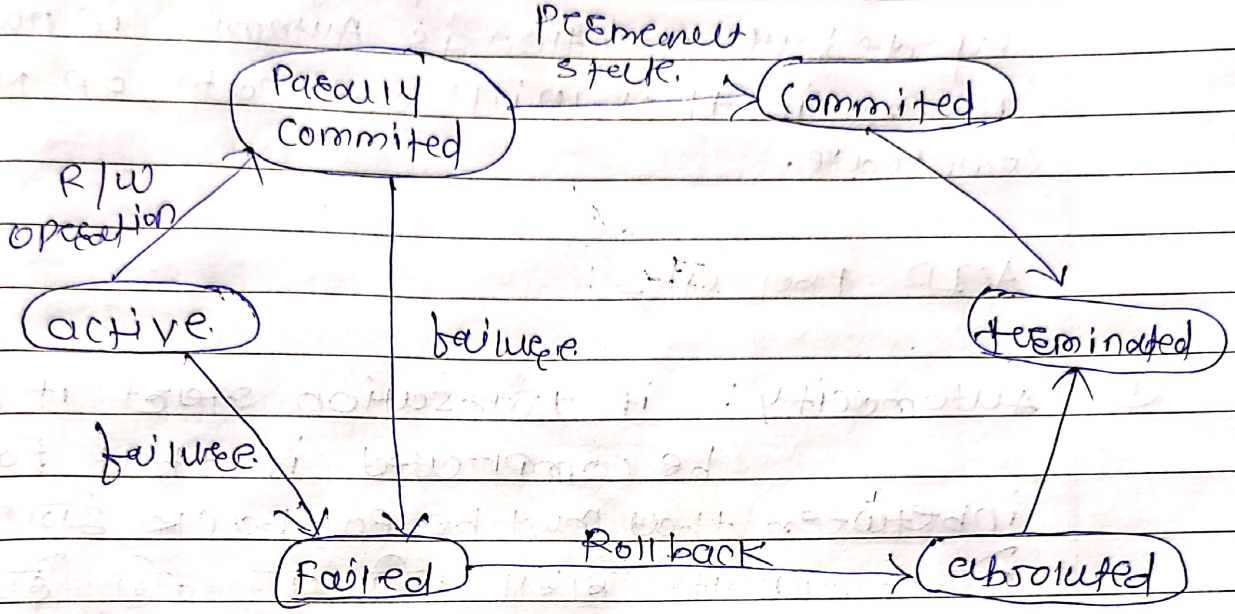
(4) consistent:
if Automacity + Isolation + Durability then it will Automatically consistent.

DB = DB

before Transaction After transaction.

Date
Page

Transaction states:



Concurrency

Running multiple transaction At A time is called concurrency.

Problems because of concurrency:

① Dirty read:

जब किसी uncommited transaction के value ~~इसरे~~ दूसरे transaction Read करके commit करके चला गया तो इसे dirty Read कहते है.

- Dirty Read से हमेशा problem नहीं होती.
- because of some Reason T1 is fail or After sometime T1 chang value then it cause problems.

☆ Unrepetable Read:

T1	T2
10 R(x)	
	R(x) 10
11 w(x)	
	R(x) 11

if you won't Read same value Again if you wants to Read then that cause.

☆ Phantom Read problem.

T1	T2
R(x)	
	R(x)
Delete(x)	
	R(x)

☆ lost update problem

T1	T2
10 R(A)	
11 w(A)	
	w(A)
	C
C.	

Date
Page

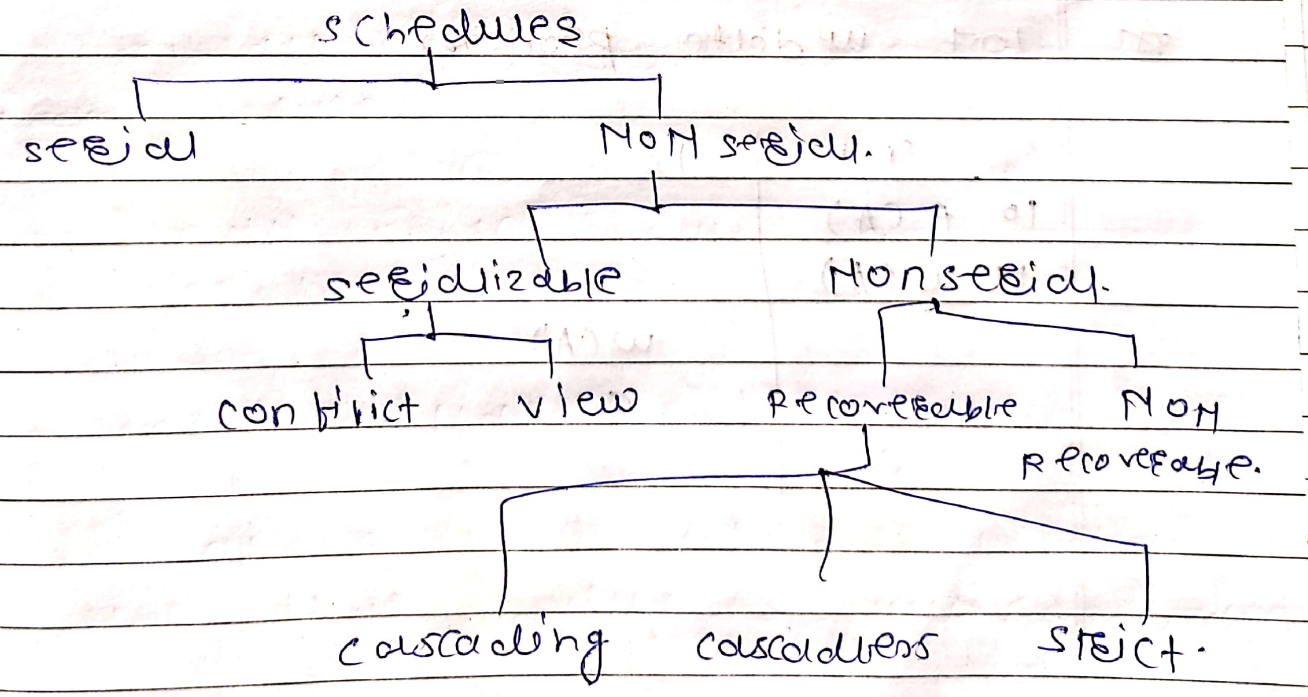
Shedules :

T1	T2
	R(B)
	W(B)
	R(A)
R(A)	
W(A)	
W(B)	

one by one execution of transaction is called seejdl shedules.

T1	T2
R(A)	
W(A)	
	R(A)
R(B)	
	R(A)

context switching in between transaction is called Nonseejdl shedule.



★ Conflict Seriliizability

A shedule is caired conflict seri-lizable if it can be transformed into a serial shedule by swaping non conflicting operation. is known as.

shedule as the name suggests is a process of lining the transaction and executing them one by one.

when to savy conflicting operation:
if satisfi following 3 -
- They belong to different transaction
- They operate on the same data item
- At least one of them is write operation

The conflict होतोय तर you can not swap them.

S1

T1	T2
R(A)	
W(A)	
	R(A)
	W(A)
R(B)	
W(B)	
	R(B)
	W(B)

⟹

T1

T1
R(A)
W(A)
R(B)
W(B)

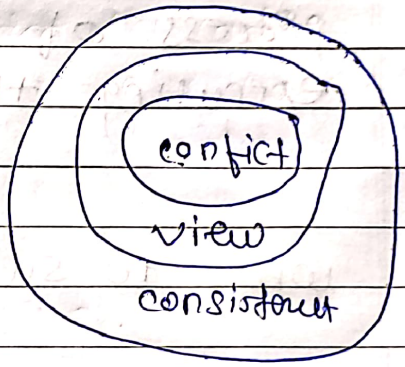
T2

T2
R(A)
W(A)
R(B)
W(B)

✱ View Serializability

A schedule is called view serializable if it is view equal to a serial schedule.

if schedule is not conflict schedule then it goes in zone of view schedule.



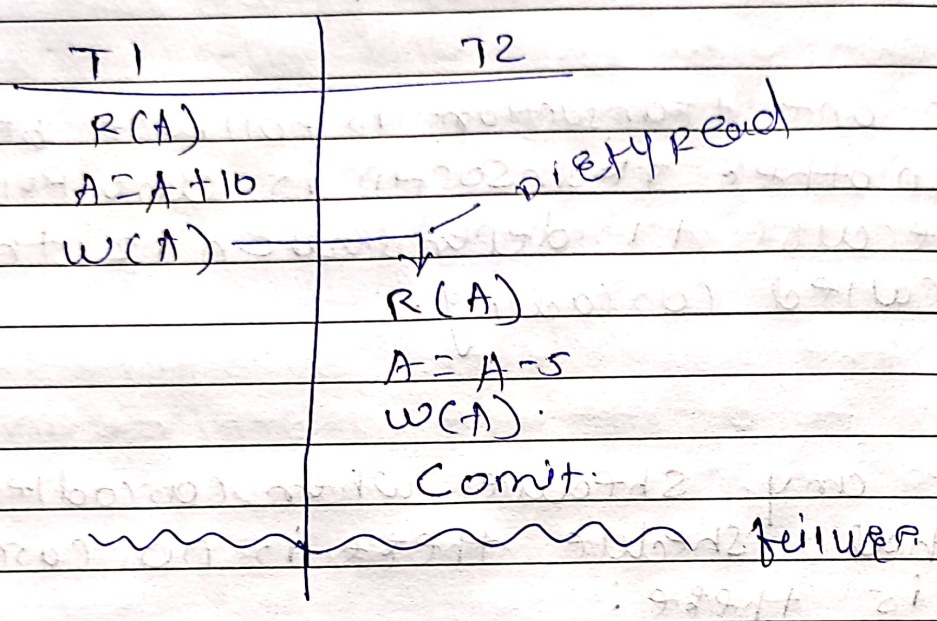
✱ When should we called view equivalence?

- initial Read should same in both ~~Tran~~ schedule.
- final write should same in both schedule.
- intermediat Read should be same

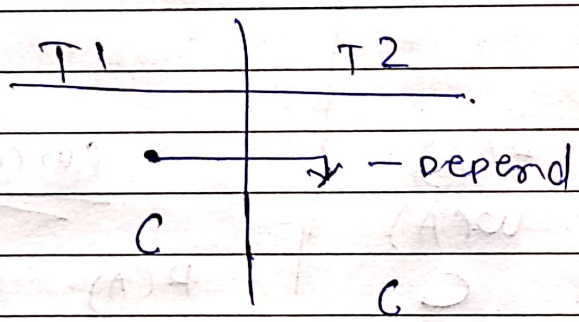
in above All condition is satisfy then that schedule is view serializable.

Date
Page

★ Recoverable shedule & No recoverable.



in above exame if any shedule comit before the dependent shedule then it is called NORecoverable.



in above example if any Shedule comit after dependent shedule then it is recoverable shedule.

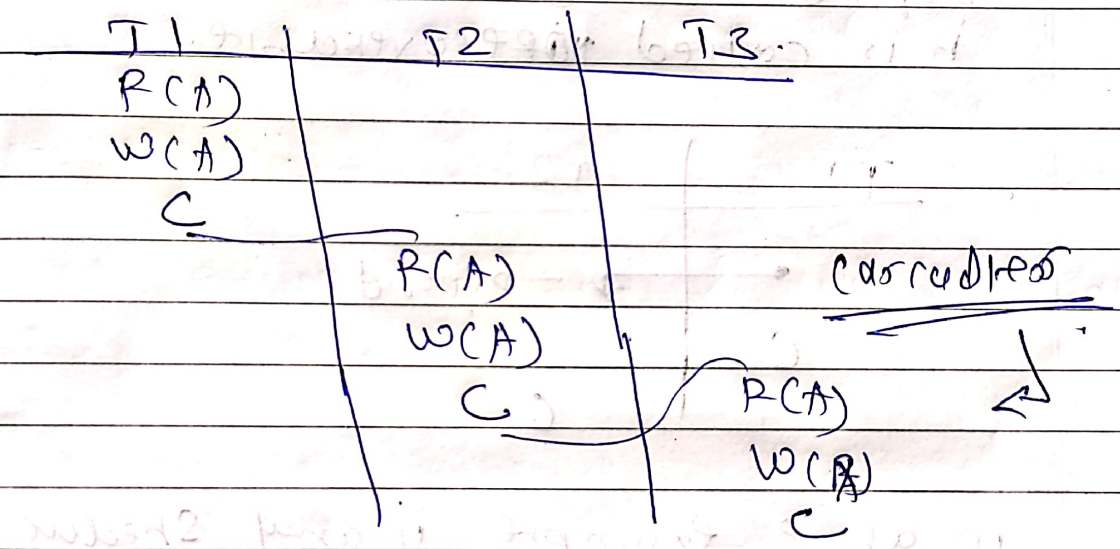
Date
Page

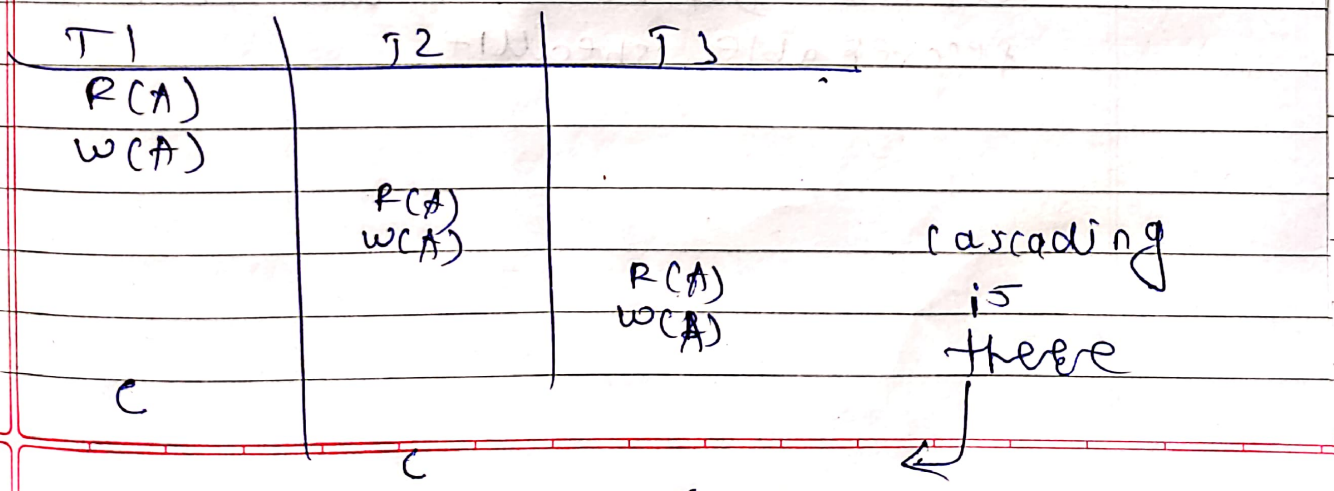
Cascadless shedule. & cascascading.

if one transaction is Rollback because on other transaction is also Rollbacking because t1 dependent on t2 that is called cascading.

if any shedule when cascadless in that shedule there is no cascading is there.

solution:



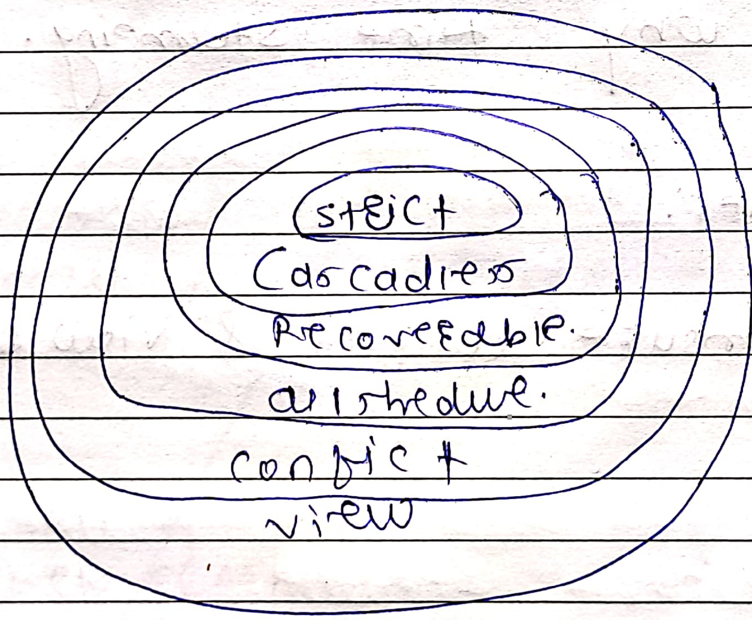


Date ____
Page ____

A Strict Shedule.

T1	T2
R(a)	
w(a)	
C	
	w(a)
	R(a)

જ્યાં સુધી first transaction write કરે છે commit નહીં કરતા ત્યાં સુધી બીજો transaction Read કે write નહીં કર શકતા is called strict shedule.



Date
Page

★ Concurrency control protocol

To manage concurrency we have to use these protocols.

① Time stamping protocol

The basic idea of time stamping is to decide the order of giving to each number to each transaction before enters into system.

We use time as stamp order token that's why time stamping.

Properties:

- It ensure conflict & view serializable.
-

locking:

we use locking for achive seriliazability how we make shedule recoverable and serializable. that is the aim of locking.

① Shared : means if you want to perform only Read then use shared.

② Exclusive: means Read and write then Exclusive.

	S	E
S	✓	✗
E	✗	✗

} compatability table.

↑ chek for concurrent transaction

Disadvantages:

① many not sufficient to produce only serializable shedule.

② may not free from recoverability.

③ may be deadlock occured

④ may not free from starvation.

2PL (2 phase locking)

growing : lock લેતો જાવ લેતો જાવ.

shrinking : unlock છોડતો જાવ છોડતો જાવ.

lock point : Last lock જ્યાં લિધાતવ.

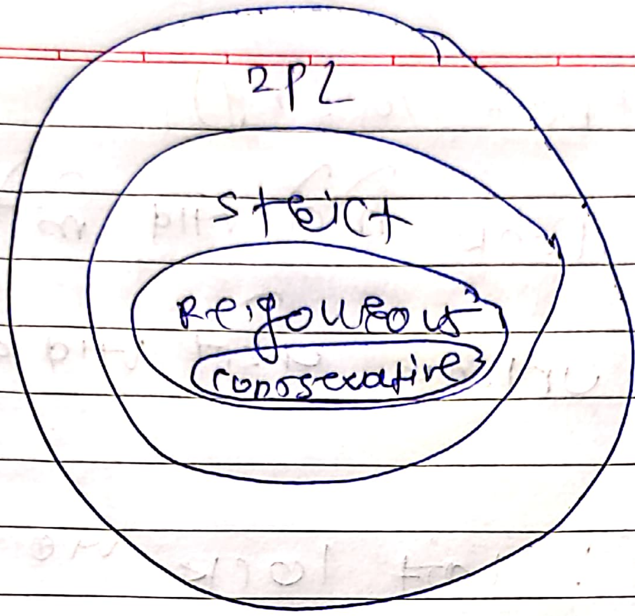
T1	T2
L(S(A))	
L(S(A))	L(E(A))
U(A)	L(A(D))
U(A)	L(B)

- it ensures srilizability
- but Deadlock may be occored
- may be irrecoverable.

strict 2PL

T1	T2
L(A)	
L(A)	
C	
U(A)	L(A)
	R(A)
	C

જ્યારે transaction commit થઈ જાતે જ પછી unlock કરે means strict 2PL



gate smashers देखो for confusion.

Dead lock : same as os in os on process होता है यहां पे Transaction पे है.

solution : - prevention
- Detection & Recovery