

Q1.

A schedule is the representation of execution sequence for all the instructions of the transactions. They are categorized in two types:

1. > Serial Schedule.
2. > Concurrent Schedule.

A schedule is said to be serial if and only if all the instructions of all the transactions get executed non-preemptively as a unit. Thus for a set of  $n$  transactions, there exists  $n!$  different valid serial schedules.

Eg for 3 transactions there are  $3! = 6$  serial schedules.

$T_1 \rightarrow T_2 \rightarrow T_3$   
 $T_1 \rightarrow T_3 \rightarrow T_2$   
 $T_2 \rightarrow T_1 \rightarrow T_3$   
 $T_2 \rightarrow T_3 \rightarrow T_1$   
 $T_3 \rightarrow T_1 \rightarrow T_2$   
 $T_3 \rightarrow T_2 \rightarrow T_1$

Also serial schedules gives guarantee for data consistency.

A schedule is said to be concurrent in case the instructions of the transactions get executed preemptively.

When the database system executes several transactions concurrently the corresponding schedule no longer needs to be serial.

Say for eg if two transactions are running concurrently, the OS may execute one transaction for a little while; then perform a context switch, execute the second transaction for some time, and the switch back, soon and so forth. With multiple transactions the CPU time is shared among all the transactions.

In general it is not possible to predict how many instructions of a transaction will be executed before CPU switches to another transaction i.e. it is much larger than  $n!$

Concurrent schedules might get affected with conflicting operations and hence does not give guaranty for data consistency.

**Conflicting operations:** a pair of operations is said to be conflicting iff  $\rightarrow$  they belong to diff. transaction.

- $\rightarrow$  Both operat<sup>n</sup> are accessing same data item.
- $\rightarrow$  Atleast one operation is write operation.

**Advantages of Concurrent schedule:** 1.) less waiting time.

2.) Improved response time & throughput.

**Disadvantages**  $\rightarrow$  ① Possibly data inconsistency ② some time too much context switching



Q2. Functional dependency determines the relation of one attribute to another attribute in a database system. It helps you maintain the quality of data in the database. A functional dependency is denoted by an arrow  $\rightarrow$ .

The Armstrong's axioms are the basic inference rule which are used to conclude functional dependencies on a relational database. They can apply on a functional dependency<sup>(FD)</sup> to derive another FD.

There are 6 rules:

1. Reflexive rule ( $IR_1$ ): In the reflexive rule, if  $Y$  is a subset of  $X$ , the  $X$  determines  $Y$ .  
If  $Y \subseteq X$  then  $X \rightarrow Y$ .

2. Augmentation Rule ( $IR_2$ ): It is also known as partial dependency. If  $X$  determines  $Y$ , then  $XZ$  determines  $YZ$  for any  $Z$ .

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

3. Transitive Rule ( $IR_3$ ): If  $X$  determines  $Y$  and  $Y$  determines  $Z$  then  $X$  must also determine  $Z$ .

If  $X \rightarrow Y$  and  $Y \rightarrow Z$  then  $X \rightarrow Z$ .

4.) Union Rule ( $IR_4$ ):

If  $X$  determines  $Y$  and  $X$  determines  $Z$  then  $X$  must also determine  $Y$  and  $Z$ .

Proof:  $X \rightarrow Y$  and  $X \rightarrow Z$  then  $X \rightarrow YZ$

$$X \rightarrow Y$$

$$X \rightarrow Z$$

$$X \rightarrow XY \text{ (using augmentation)}$$

$$XY \rightarrow YZ \text{ (using } IR_2)$$

$$X \rightarrow YZ \text{ (using } IR_3 \text{ on above 2).}$$

5.) Decomposition Rule ( $IR_5$ ): Also known as project rule, it is reverse of union rule.

If  $X \rightarrow YZ$  then  $X \rightarrow Y$  and  $X \rightarrow Z$

Proof:  $X \rightarrow YZ$ .

$$YZ \rightarrow Y \text{ (using } IR_1)$$

$$X \rightarrow Y \text{ (using } IR_3 \text{ on above 2).}$$

6.) Pseudo transitive rule ( $IR_6$ ): If  $X \rightarrow Y$  and  $YZ \rightarrow W$  then  $XZ \rightarrow W$ .

Proof:  $X \rightarrow Y$ .

$$WY \rightarrow Z$$

$$WX \rightarrow WY \text{ (using } IR_2)$$

$$WX \rightarrow Z \text{ (using } IR_3 \text{ on above 2)}$$



Q3  $R = (A, B, C, D, E)$

FD =  $(A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A)$

Now,

$A \rightarrow BC, B \rightarrow D$  so it means  $A \rightarrow D \Rightarrow A \rightarrow DC \rightarrow E$

~~∴~~  $A \rightarrow ABCDE$ .

$E \rightarrow A, A \rightarrow ABCDE$ , so  $E \rightarrow ABCDE$

$CD \rightarrow E$ , so  $CD \rightarrow ABCDE$ .

$B \rightarrow D, BC \rightarrow CD$ , so  $BC \rightarrow ABCDE$

∴ the candidate keys are  $A, E, CD$  and  $BC$  and any combination of attributes that includes those is a superkey.

~~To check~~ Now, prime attributes  $\rightarrow A, C, D, E$   
non prime "  $\rightarrow B$ .

$R$  is in first normal form  $\rightarrow$  for each FD, LHS must be PK or SK or RHS must be prime attribute.

$A \rightarrow BC$  ✓ valid.

CK  $CD \rightarrow E$  ✓ valid.

CK

$B \rightarrow D$  (prime attribute) ✓ valid.

CK  $E \rightarrow A$  ✓ valid.

$\Rightarrow R$  is In 1<sup>st</sup> normal form (at least)

Q4

Deadlock refers to a specific condition when two or more processes are each waiting for another to release a resource, or more than two processes are waiting for resources in a circular chain. It is common in multiprocessing where many processes share a specific type of mutually exclusive resource known as a software lock. Computers intended for the time sharing and/or real-time markets are often equipped with a hardware lock. Deadlocks are troubling because there is no general solution to avoid deadlocks.

These are the following conditions for deadlock:

1. Mutual exclusion condition: A resource cannot be used by more than one process at a time.
2. Hold and wait condition: processes already holding resources may request new resources.
3. No preemption condition: only a process holding a resource may release it.
4. Circular wait condition: Two or more processes from a circular chain where each process waits for a resource that the next process in chain holds.



Q5

(18EJ1C169) Vaibhav Saxena 7

Serializability is a concept that helps us to check which schedules are serializable. It is the major correctness criterion for concurrent transactions executions. It is considered the highest level of isolation b/w transactions and plays an essential role in concurrency control. Its theory provides the formal framework to reason about and analyze serializability & its techniques. As such it is supported in all general purpose database systems.

Conflict serializability	View Serializability
1. > <del>conflict</del> It is easy to be achieved.	1. > It is hard to be achieved.
2. > Every conflict serializable is view serializable.	2. > not vice versa.
3. > It is easy to test and cheap	3. > It is not easy & expensive to test
4. > Most of concurrency control schemes used in practice are based on conflict serializability	4. > not so for view serializability.

Contd.  
Q3

★ - Check for II normal form: (18EJ1CS169) Vaibhav Saran. (8)

for each FD, LHS must be proper subset of CK and.  
RHS must be NPA.

$$\begin{array}{c|c|c|c} A \rightarrow BC & \times & CD \rightarrow E & \times \\ T & F & T & F \end{array} \quad \begin{array}{c|c|c|c} B \rightarrow D & \times & E \rightarrow D & \\ F & F & F & F \end{array}$$

∴ the highest normal form of given functional dependency is 2<sup>nd</sup> normal form.