Page No.
Date

College Name:

DELHI GLOBAL INSTITUTE OF TECHNOLOGY

Name : BAZGHA RAZI

Course Code : PCC-CSE-201G

Subject : Database Management System

Session : 2019-2023

Page No.
Date

Ans: a) ~~Exp~~ TRC and DRC

Tuple Relational Calculus	Domain Relational Calculus
→ The TRC specified to select the tuple in a relation.	→ The second form of relation is known as DRC. In this filtering variable ~~with~~ uses the domain of attributes.
→ In TRC, filtering variable uses the tuples of a relation.	→ DRC, uses the same ~~operands~~ operators as the tuple calculus. It uses logical connectivities i.e., and ($\wedge$), or ($\vee$), not ($\neg$)
→ The result of the relation can have one or more tuples.	→ It uses existential ($\exists$) and universal ($\forall$) quantifiers to bind the variables.
→ Notation: $\{T \mid P(T)\}$ or $\{T \mid \text{Condition}(T)\}$ where: T ⇒ resulting tuples P(T) ⇒ the condition used to fetch T.	→ Notation: $\{a_1, a_2, \ldots, a_n \mid P(a_1, a_2, \ldots a_n)\}$ where: $a_1, a_2, a_3, \ldots a_n$ ⇒ attributes P ⇒ formula built by inner attributes.
→ Example: $\{T.name \mid Author(T)$ AND T.article = 'database'$\}$	→ Example: $\{<article, page, subject> \mid \in javapoint \wedge subject = 'database'\}$

Page No.
Date

Ans 1 b) Oracle

Oracle database is a relational database management system. It is known as Oracle database, OracleDB or simply oracle. The oracle relational database management system provides an open comprehensive, integrated approach to information approach.

Oracle is not a DBMS, but we can say that it is RDBMS. i.e., it give the flexibility to create a table, multiple tables and create relations between multiple tab. It is widely used in enterprise applications.

Server parameter file, control files, datafiles, password file, log files, etc are the components of oracle database.

Application of Oracle.
→ Supply chain management
→ Customer relationship management
→ Enterprise resource planning
→ Manufacturing (discrete fields)
→ Product life-cycle management.

Advantages of Oracle
→ Backup and recovery
→ Multiple database support
→ Performance

Disadvantages of Oracle
→ Expensive
→ Complexity
→ For every system, we have to buy individually.

Page No.
Date

Ans 1(c) Integrity Constraints

a) NOT NULL Constraint : Specifies that values of a column can not be null

b) DEFAULT <value> Constraint : Define a default value for an attribute

c) UNIQUE Constraint : Specifies that values of two rows for a specified column must be unique & different.

d) CHECK <condition> constraint : utilized to check the validity of data entered into particular table columns.

e) Primary Key Constraint : It specifies one or more attributes that make up the primary key of the table.

f) Foreign Key Constraint : It is used to enforce referential integrity b/w tables in a relational db. A column defined as a foreign key is used to reference a column defined as a foreign key is used to reference a column defined as a primary key in another table.

Date

Referential Integrity Constraints

It is based on the concept of foreign keys. A foreign key is an important attribute of relation which should be referred to in other relationships. Referential integrity constraint state happens where relation refers to a key attribute of a different or same relation. However, that key element must exist in the table.

Example:

Coustomer_ID	Coustomer_Name	Status
1	Ram	Active
2	Sita	Inactive
3	Meenu	Active

:- Table Coustomer

Invoice_No	Coustomer_ID	Amount
1	1	$100
2	1	$200
3	2	$150

:- Table Billing

In the above example, we have two relations, coustomer and billing.

Tuple for Coustomer_ID = 1 is referenced twice in the relation billing. So we know

Coustomer_Name = Ram has ~~referenced~~ billing amount $ 300

Page No.
Date

Ans 1.d)

Open Source	Commercial Source
i) Which do not requires the purchase of several additional software licenses.	i) Which requires the purchase of several additional software licenses.
ii) Supports many common operating systems.	ii) Supports windows operating system; can support Linux and Mac OS using Docket.
iii) Supports only english language.	iii) Supports variety of languages.
iv) Can't stop a query execution.	iv) Can stop a query execution.
v) Capable to SQL injection attacks.	v) Highly secure.
vi) It provide limited technical support.	vi) It provide guaranteed technical support.
vii) In opensource database anyone can easily view source code of it.	vii) Commercial source database are that which has been created for commercial purpose.
viii) Example: MySQL, MongoDB, etc.	viii) Example: Oracle, DB2, etc.

Page No.
Date

Ans) Keys are the set of attributes which helps you to identify a row (tuple) in a relation.

There are mainly ~~seven~~ five different types of Keys in DBMS:

→ Superkey: It is a group of single or multiple keys which identifies rows in a table. A Superkey may have additional attributes that are not needed for unique identification.

S.No	Roll_no.	Name
98124	01	Ellen
98139	02	Rose
97335	03	James

So, in this example "S.No" and "Roll_no." are super keys.

→ Primarykey: It is a column or group of columns in a table that uniquely identify every row in that table. The primary key can't be duplicate meaning same value can't appear more than once in the table. A table can not have more than one primary key.

Roll_no.	Name	Marks
1	Ellen	45
2	Rose	48
3	James	45

So in this example "Roll_no." is a primary key.

Page No.
Date

→ Candidate key : It is a set of attributes that uniquely identify tuples in a table. A table can have multiple candidate keys but only a single primary key.

Roll_No.	Name	Email
1.	Ellen	abc@gmail.com
2.	Rose	uvw@gmail.com
3.	James	xyz@gmail.com

So, in this example "Roll_No." and "Email" are candidate key which help us to uniquely identify the student record in table.

→ Foreign key : It is a column that creates a relationship between two tables. The main purpose of foreign key is to maintain data integrity. It act as a cross-reference b/w two tables as it references the primary key of another table.

Subject_Code	Subject
001	Science
002	English
003	Computer

Subject Code is a foreign key. By this we create a relationship b/w these 2 tables.

ID	Subject_Code	Name
B002	002	Ellen
B015	003	Rose
B014	002	James

Page No.
Date

→ Compound key : It has two or more attributes that allow you to uniquely recognize a specific record. The purpose of the compound key in database is to uniquely identify each record in the table.

Order_No	Product_ID	Product_Name	Quantity
B005	DKT32173	Maggi	10
B005	OM37683	Biscuits	5
B002	DKT32173	Chocolate	20

In this example, Order_No and Product_ID can't be a primary key as it does not uniquely identify a record. However, a compound key of order No and Product_ID could be used as it uniquely identified each record.

→ Alternate key : It is a column or group of columns in a table that uniquely identify every row in that table. A table can have multiple choices for a primary key but only one can be set as the primary key. All the keys which are not primary key are called an Alternate key.

Roll_no.	Name	Email
1.	Ellen	abc@gmail.com
2.	Rose	xyz@gmail.com
3.	James	uvw@gmail.com

So, in this example "Roll_no" is primary key and "Email" is alternate key.

Page No.
Date

Ans 2.(a) DBMS architecture will help us understand the components of database system and the relation among them. The architecture of DBMS depends on the computer system on which it runs.

The architecture is divided into 3 general levels:

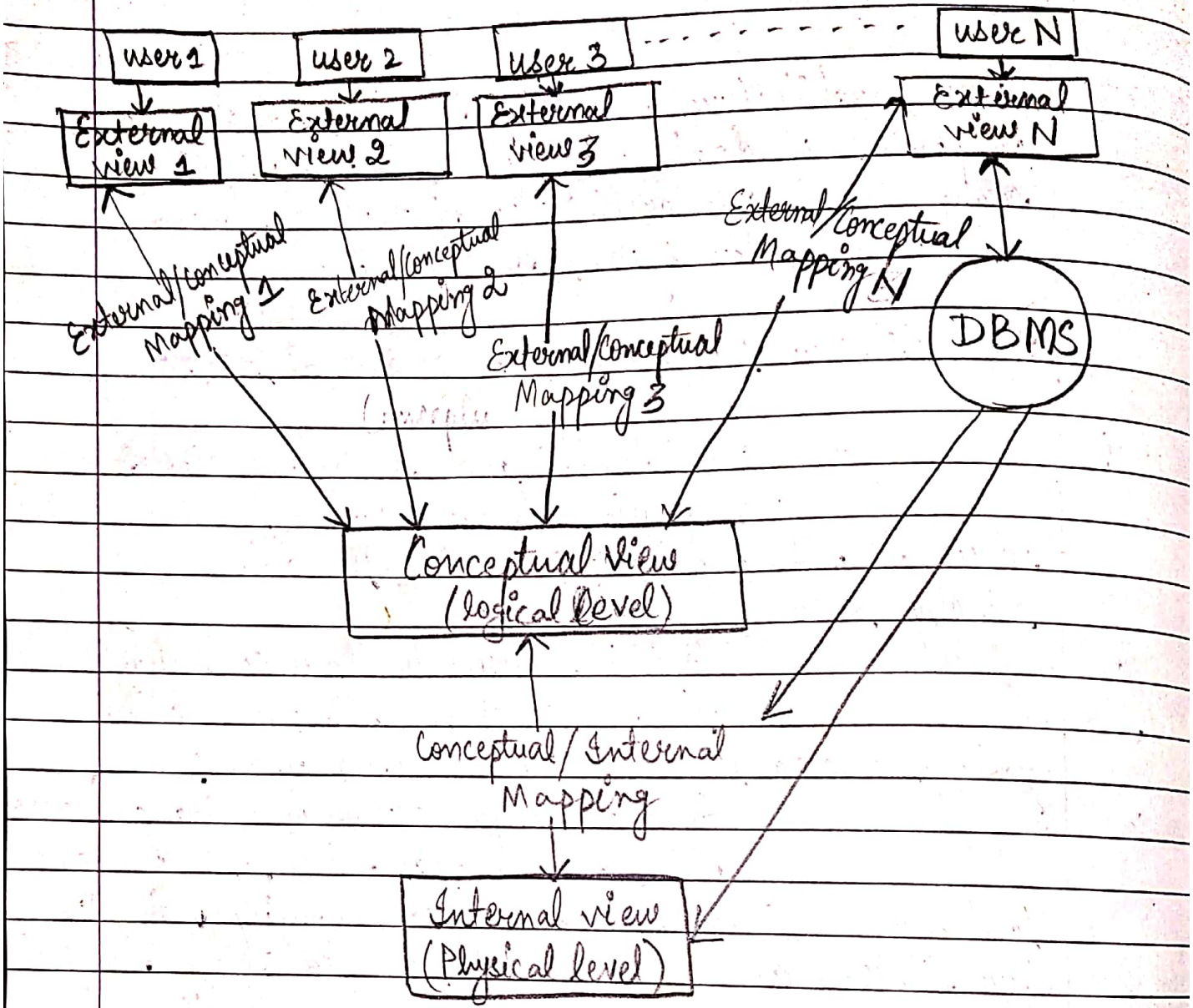
1) Internal Level: The internal level is the one closest to physical storage that is, the one concerned with the way in which the data is actually stored.

2) External Level: It is one closest to the users, that is the one concerned with the way in which the data is viewed by individual users.

3) Conceptual Level: It is the representation of entire contents with individual user views, the conceptual level may be though as defining a community user view. There will be many external views, each consisting of more or less abstract representation of source portion of the database and there will be a single conceptual view. It is also known as the logical level.

4) Conceptual/Internal mapping: Defines the correspondance b/w the conceptual view and the stored database. It specifies how conceptual records and fields are represented at the internal level.

5) External/Conceptual mapping: It define the correspondance relation between a particular external view and the conceptual view.

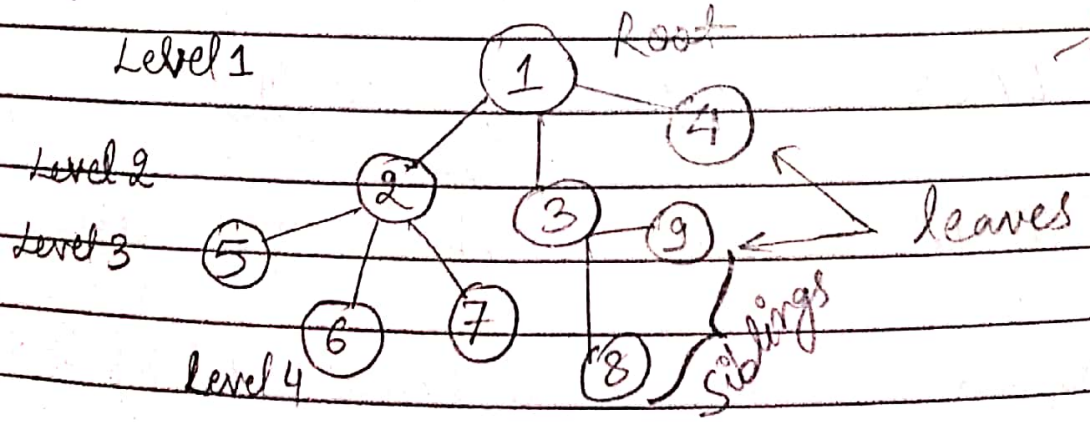


↳ Architecture of Database management system.

Ans 2b) A data model is a set of concepts that can be used to describe the structure of a database. By structure of a database mean the data types, relationships & constraints that should hold for the data.

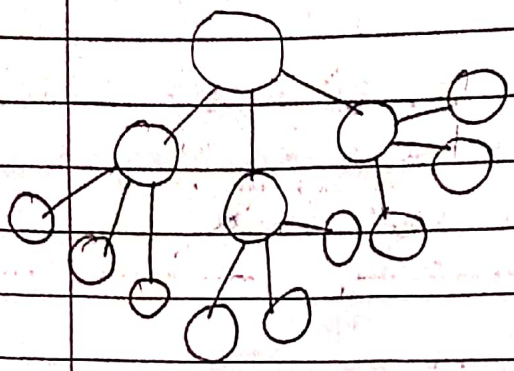
HIERARCHICAL MODEL

This is the oldest of the 3 record data models. This uses the tree as its basic structure. A tree is a data structure that consists of a hierarchy of nodes, with a single node, called the root, at the highest level. A node may have any number of children, but each child node may have only one parent node on which it is dependent. The parent-to-child relationship in a tree is thus a one-to-many relationship but the child-to-parent is one-to one. Parent child relationship are shown by drawing a line or edge between the parent and child nodes.

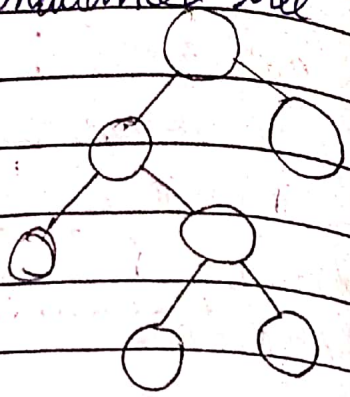


Leaf: which has no children

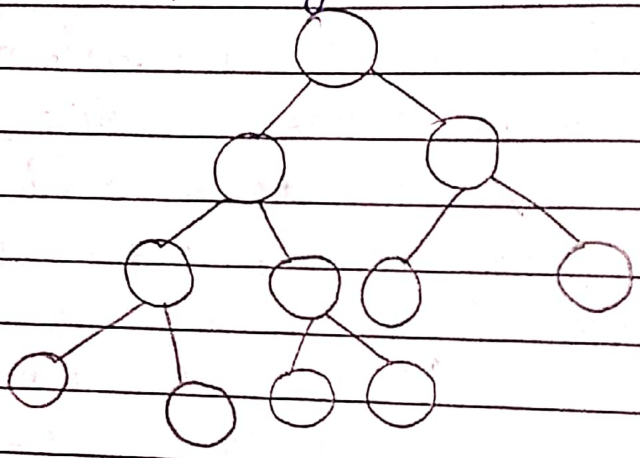
Balanced tree



Unbalanced tree



Binary Tree



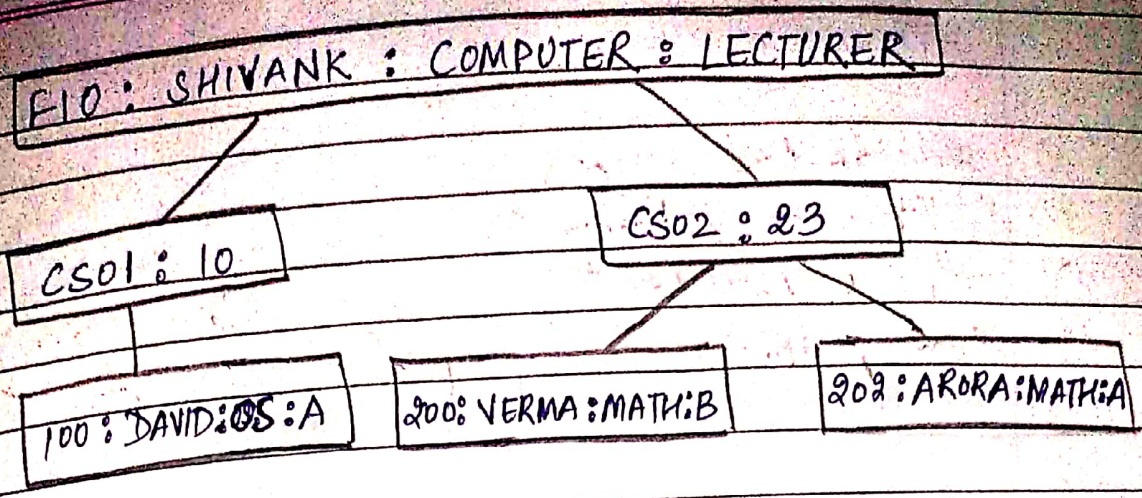
Each node has not more than 2 children

Ex: FACULTY | FID : F_Name : DEPT : RANK

CLASS | COURSE : ROOM

STUDENT | STUID : STUNAME : MAJOR : GRADE

FIG shows the occurence of the FACULTY-CLASS-STUDENT



Three operation on this Model

1) Deletion : If we ~~hav~~ want to remove the class say CS01, then student with STUID equals to 100 will also have to be removed.
So deletion is very difficult.

2) Insertion : A new class, say CS03 cannot be introduced unless some faculty is available at root level. So insertion is also difficult.

3) Updation : Suppose ROOM has changed, for this change, search problem will occur which is time consuming, hence we can say that this model has problem in updation

Hence we can say that this model has problem in insertion, deletion & updation.

Page No.
Date

Advantage

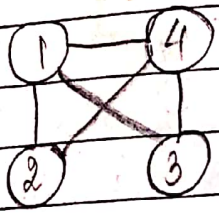
1) Easy to understand.
2) Performance is better then relational data model.
3) ~~Confluenced~~ Continues to be widely use
4) Some problems lend themselves to this model.

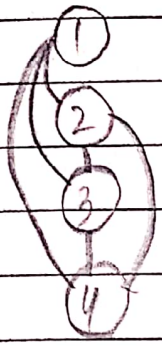
Disadvantage

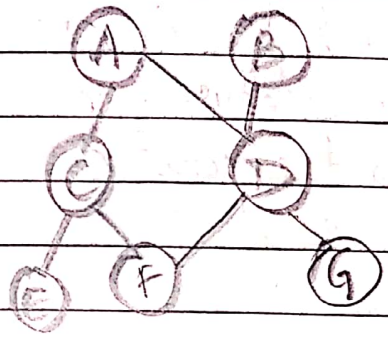
1) Difficult to access values at lower levels.
2) Insertion is difficult
3) Deletion is difficult
4) Updation is difficult

NETWORK MODEL

It uses the network or ple structure as its basic data structure. A network is a directed graph consisting of nodes connected by links or directed arcs. The nodes correspond to record types & the link to pointers.
The network data structure looks like a tree structure except that a dependent node, called a child or member, may have more than one parent or owner node.







eg:

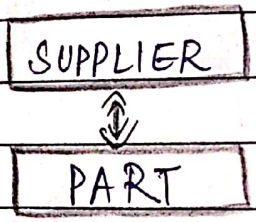
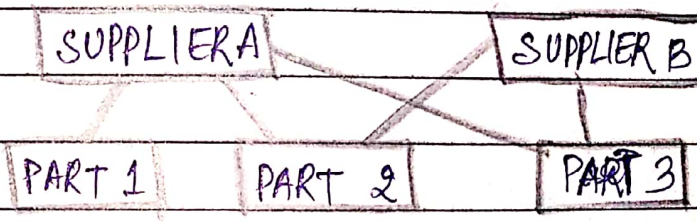


Fig:



1. Insertion: can be easily inserted
2. Deletion: very simple
3. Updation: Suppose Supplier B supplies PART 1 in place PART 2, we can make this change simply changing the link of SUPPLIER B from PART 2 to PART 1. Hence, updation is also very easy in this model.

Page No.
Date

Advantages

1) Easy access to data
2) Flexible
3) Efficient

Disadvantages

1) Complex to design & develop.
2) Extra memory is required for the storage of the pointers.
3) For large db, operation & maintenance of the network model are time consuming and expensive.

RELATIONAL MODEL

It is physically represented as a table. Tables are used to hold information about the object to be represented in the database.

A relation is represented as a 2-D table
Rows → called tuple
Columns → called Attributes

Ex: STUDENT

STUID	SNAME	SUBJECT	MARKS
1015	Mary	Maths	42
1005	Jones	History	13
1001	Smith	History	90

Operations on Relational Model

Insertion : A new student, in fig can be Inserted quite easily

Deletion : delete a tuple from STUDENT relation is also easy.

Updation : is also easy.

Advantages

1) Easy to use & understand
2) Very flexible
3) Widely used
4) Provides excellent support for adhoc queries.

6) Structure & access strategy

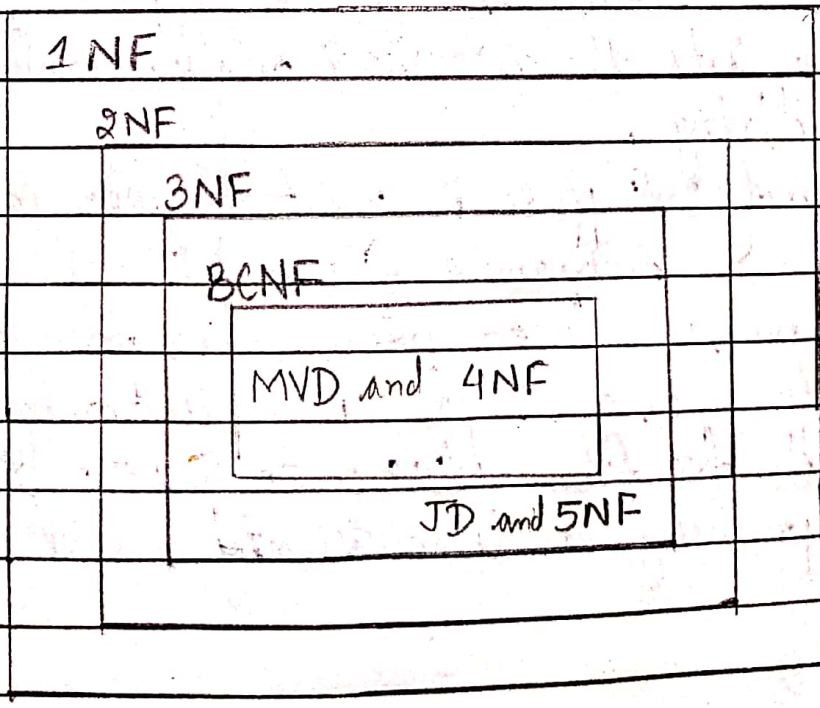
Disadvantages

1) For large db, the performance in responding to queries is definitely degraded.

Ans 3(a) It involves dividing a database into two or more tables and defining relationship between the tables. The objective is to isolate data so that addition, deletion and modification of a field can be made in just one.

Normal Forms

A relation is send to be in particular normal form if it satisfies or certain specified set of contraints.
"It is a process of decomposing a set of relations to produce smaller and well structured relations that contain minimum or no redundancy."



Page No.
Date

Modification Anomalies

SID	Activity	Fee
100	SKING	200
150	SWIMMING	150
175	SQUASH	50
200	SWIMMING	150

1. Deletion Anomaly : Suppose that activity has a fixed fee that is the same for all students. If we delete the tuple for the student 100, we loose not out the fact that student 100 is skier, but also the fact that skiing costs $200. This is called deletion anomaly.
2. Insertion Anomaly : Suppose we want to store the fact that scuba diving costs $175. We can't enter this data into the ACTIVITY relation until a student takes scuba diving.
3. Updation Anomaly : Suppose we want to change the fee of swimming from $150 to $250. For this change, we are faced with either problems of searching to find every tuple with swimming activity and change the fee from $150 to $250 So possibility of ~~productivity~~ producing an inconsistenty result.

First NF : A relation is, in first normal form if and only if all underlying domains contain atomic values only.

The relation Activity in figure 1.

STU ACT

SID	Activity
100	Skiing
150	Swimming
175	Squash
200	Swimming

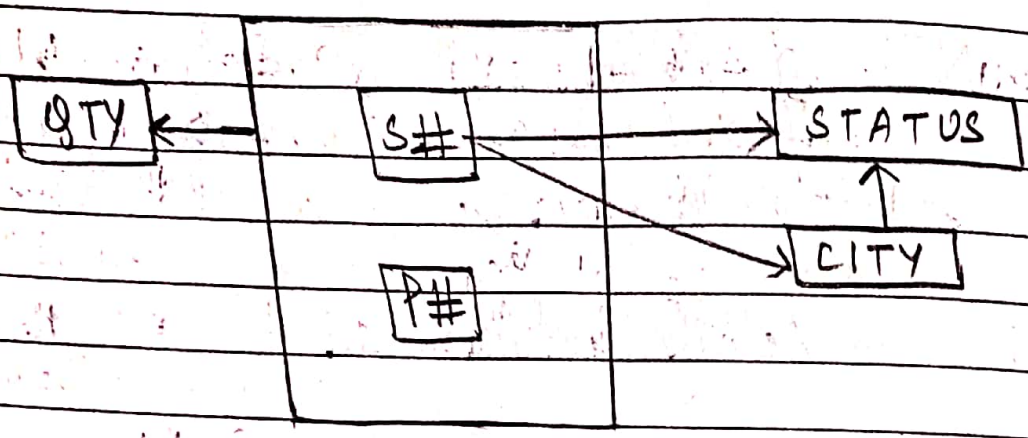
ACT-COST

Activity	Fee
Skiing	200
Swimming	150
Squash	50

Consider the relation FIRST (S#, STATUS, CITY, P#, QTY)

FIRST

S#	STATUS	CITY	P#	QTY
S1	30	Delhi	P1	100
S1	30	Delhi	P2	125
S1	30	Delhi	P3	130
S1	30	Delhi	P4	115
S2	10	Karnal	P1	200
S2	10	Karnal	P2	215
S3	40	Rohtak	P1	200
S4	30	Delhi	P4	200
S4	30	Delhi	P5	300



The key of FIRST relation is the combination (S#, P#)
It is clear that relation FIRST is in 1NF. It has following modification anomalies:

1) Deletion Anomaly : If we delete the first tuple with S# value S3 and P# value P1, we loose the the information that S3 is located in Rohtak. Hence the relation FIRST has a deletion anomaly.
2) Insertion Anomaly : We cannot enter the fact that a particular supplier is located in a particular supplies until that supplies is located in a particular supplies until that supplies supplies at least one part. So, we cannot enter the fact that supplier S5 is located in Rajasthan until S5 supplies at least one part.
3) Updation Anomaly : If supplies S1 moves from Delhi to Ambala, we are faced with either the problem of searching the FIRST relation to find every tuple connecting S1 and Delhi and change it or the possibility of producing an inconsistent result. The solution of these problems is to replace the relation FIRST by the two relation SECOND.

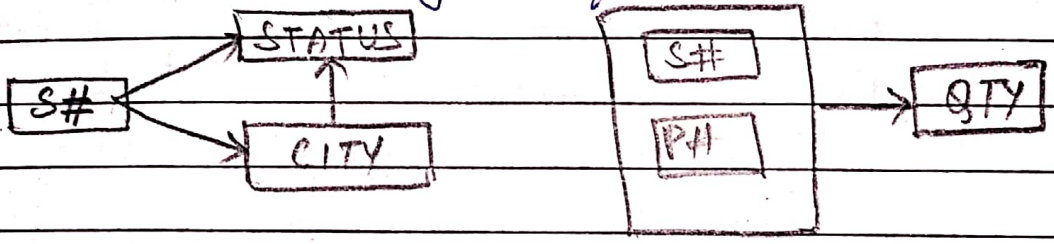
SECOND

S#	STATUS	CITY
S1	30	Delhi
S2	10	Karnal
S3	10	Rohtak
S4	30	Delhi

SP

S#	P#	QTY
S1	P1	1000
S1	P2	125
S1	P3	130
S1	P4	115
S2	P1	200
S2	P2	215
S3	P1	200
S4	P4	200
S4	P5	300

It shows the FD diagrams for the two relations.



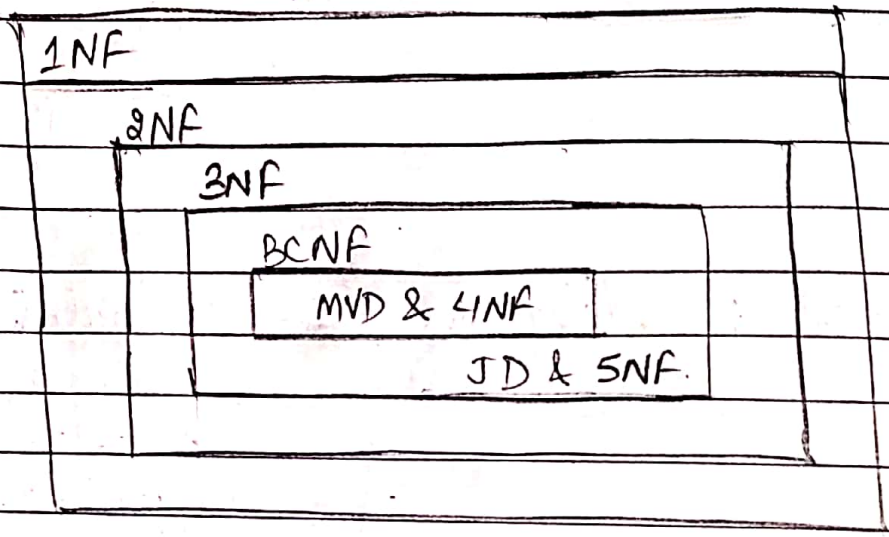
2NF: A relation is in (2NF) if it is in 1NF and every nonkey attribute is fully dependent on the key. If the key is a single attribute, then the relation is automatically in the 2NF. For eg: STU-ACT and ACT-COST, relations are in 2NF because they have simple attribute key. Relation SECOND and SP are also both is 2NF

It has following modification anomalies:

1) Deletion Anomaly: If we delete the second tuple for a particular city, we distroy not only the information for the supplier concerned but also the information that the second

tuple for S2, we losse the information that the status for Karnal is 10.

2) Insertion Anomaly: We cannot enter the fact that a particular city has a particular status value. For eg:



We cannot state that any supplier in Ambala must have a status until we have some supplier located in that city.

3) Updation Anomaly: If we want to change the status value for Delhi from 30 to 50; we are faced with either the problem of searching the Second relation to find every tuple for Delhi or the possibility of producing an inconsistent result. Again the solution of the problems is to replace the Second relation by two relations.

SC(S#, CITY) and CS(CITY, STATUS)

SC

S#	CITY
S1	Delhi
S2	Karnal
S3	Rohtak
S4	Delhi

CS

CITY	STATUS
Delhi	30
Karnal	10
Rohtak	40

Page No.
Date

3NF : A relation R is in third normal form iff it is in 2NF and every nonkey attribute is nontransitively dependent on the primary key.

For Example: Relation SC, CS and SP are in 3NF but the relation Second is not in 3NF.

BOYCE-CODD NORMAL FORM (BCNF)

Even relations in 3NF can have anomalies. Consider the ADVISOR (SID, Major, Fname) relation in figure. Suppose the key underlying the relation are that a student (SID) can have one or more major, a major can have several faculty members (Fname) as advisors and a faculty member (Fname) advises in only one.

Key (primary) : (SID, Major)
Key (Candidate) : (SID, Fname)
Functional Dependencies : Fname → Major

ADVISOR

SID	Major	Fname
100	Math	Cauchy
150	Physics	Jung
200	Math	Rieman
250	Math	Cauchy
200	Physics	Chetan

Page No.
Date

Relation in 3NF but not in BCNF.
A relation in BCNF if every determinant is a candidate key.
Advisor is not in BCNF, it has a determinant Fname, that is not a candidate key.
Advisor can be decomposed into two relations.

STU-ADV (SID, Fname)
Key : (SID, Fname)

SID	Fname
100	Cauchy
150	Jung
200	Riemann
250	Cauchy
200	Chetan

ADV-SUBJ (Fname, Subject)
Key : Fname
Fname Major.

Fname	Major
Cauchy	Maths
Jung	Physics
Riemann	Maths
Chetan	Physics

4NF: Fourth normal form is a level of database normalization where there are no non-trivial multivalued dependencies other than a candidate key. It builds on the first three normal forms i.e., 1NF, 2NF, 3NF and the Boyce-Codd Normal form. It states that, in addition to a database meeting the requirements of BCNF, it must not contain more than one multivalued dependency.

A table with a multivalued dependency violates the normalization standard of Fourth normal form because it creates unnecessary redundancies and can contribute to inconsistent data. To bring this up to 4NF, it is necessary to break this information into two tables.

Properties: A relation R is in 4NF if and only if the following conditions are satisfied:

i) It should be in the Boyce-Codd Normal Form.
ii) The table should not have any multivalued dependency.

Page No.
Date

5NF: A relation R is in 5NF if and only if every join dependency in R is implied by the candidate key of R. A relation decomposed into two relations must have loss-less join property, which ensures that no spurious or extra tuples are generated, when relations are reunited through a natural join.

Properties: A relation R is in 5NF if and only if it satisfies following conditions:

i) R should be already in 4NF
ii) It cannot be further non loss decomposed (join dependency)

Ans 3(b) Relational Algebra is a procedural language. It specifies the operations to be performed on existing relations to derive result relations.

The eight relational operators are:

1. Union (∪) – The union of two relation A & B, A∪B is the set of all the tuple 't' belongs to either A or B (or both)

2. Intersection (∩) – The intersection of two relations A & B, A∩B is the set of all tuples 't' belonging to both A & B

3. Difference (−) – The difference between two relations A & B, A−B is the set of all tuples 't' belonging to A & not to B

Example:

SNUM	NAME	MAJOR
128	JONES	HISTORY
158	PARKS	MATH
271	SMITH	HISTORY

A

NUMBER	NAME	INTEREST
105	RAM	MGT
123	JONES	HISTORY

B

→ A ∪ B

SNUM	NAME	MAJOR
123	JONES	HISTORY
158	PARKS	MATH
271	SMITH	HISTORY
105	RAM	MGT

→ A ∩ B

SNUM	NAME	MAJOR
123	JONES	HISTORY

→ A − B

SNUM	NAME	MAJOR
158	PARKS	MATH
271	SMITH	HISTORY

4. Product (×) – The product of two relations (cartesian product) is the concatenation of every tuple of one relation with every tuple of a second relation.

SID	NAME	MAJOR	AGE
123	Jones	HISTORY	21
158	Parks	MATH	26
105	Ram	MGT	27

STUDENT

SNUM	CNAME	PNUM
123	H350	1
105	BA490	3

ENROLLMENT

STUDENT X ENROLLMENT

SID	NAME	MAJOR	AGE	SNUM	CNAME	PNUM
123	JONES	HISTORY	21	123	H350	1
123	JONES	HISTORY	21	105	BA490	3
158	PARKS	MATH	26	123	H350	1
158	PARKS	MATH	26	105	BA490	3
105	RAM	MGT	27	123	H350	1
105	RAM	MGT	27	105	BA490	3

5. Project Operation (π) - It is an operation that selects specified attributes from a relation. The result of the projection is a new relation having the selected attributes

$\pi_{\text{<attribute list>}}(R)$

Example : STUDENT

SID	NAME	MAJOR	GRADE	AGE
123	Jones	History	JR	21
158	Parks	Math	GR	26
105	Ram	MGT	SN	27
271	Smith	History	JR	19

$\pi_{NAME, MAJOR}(STUDENT)$

$\pi_{MAJOR, GRADE}(STUDENT)$

NAME	MAJOR
JONES	History
PARKS	Math
RAM	Mgt
SMITH	History

MAJOR	GRADE
History	JR
Math	GR
Mgt	SN

6. Select Operation (σ) – Whereas the projection operator takes a vertical column of a relation, the selection operator takes a horizontal rows.

$\sigma_{\text{<selection condition>}}(R)$

Example :

SID	NAME	MAJOR	GRADE	AGE
123	Jones	History	JR	21
271	Smith	History	JR	19

Result of $\sigma_{AGE<25}(STUDENT)$

Page No.
Date

7. Join Operation (⋈): We can join two tables if both tables have a column (attributes) defined over some common domain.

The join of the relations, say A and B, operates as follows.

i) From the product of A, B i.e., A X B.
ii) Do a selection to eliminate some tuples.
iii) Then remove duplicate attributes with projection.

A ⋈ <join condition> B

Example: Consides the STUDENT relation shown in fig ENROLLMENT relation shown in fig.

ENROLLMENT

SNUM	CNAME	PNUM
128	H350	1
105	BA490	3
123	BA490	7

STUDENT ⋈ STUDENT.SID=ENROLLMENT.SNUM

we first take the product of STUDENT X ENROLLMENT

SID	NAME	MAJOR	GRADE	AGE	SNUM	CNAME	PNUM
123	Jones	History	JR	21	123	H350	1
123	Jones	History	JR	21	105	BA490	3
123	Jones	History	JR	21	123	BA490	7
158	Parks	Math	GR	26	123	H350	1
158	Parks	Math	GR	26	105	BA490	3

SID	NAME	MAJOR	GRADE	AGE	SNUM	CNAME	PNUM
158	Parks	Math	GR	26	123	BA490	7
105	Ram	Mgt	SN	27	123	H350	1
105	Ram	Mgt	SN	27	105	BA490	3
105	Ram	Mgt	SN	27	123	BA490	7
271	Smith	History	JR	19	123	H350	1
271	Smith	History	JR	19	105	BA490	3
271	Smith	History	JR	19	123	H350	7

STUDENT X ENROLLMENT

SID	NAME	MAJOR	GRADE	AGE	SNUM	CNAME	PNUM
123	Jones	History	JR	21	123	H350	1
123	Jones	History	JR	21	123	BA490	7
105	Ram	MGT	SN	27	105	BA490	3

EQUIJOIN

Observe that two attributes are identical ; SID and student number. One of these two is unnecessary. So we eliminate one of them with projection.

SID	NAME	MAJOR	GRADE	AGE	CNAME	PNUM
123	Jones	History	JR	21	H350	1
123	Jones	History	JR	21	BA490	7
105	Ram	MGT	SN	27	BA490	3

NATURAL JOIN

It is possible to join on condition other than equality.

for ex: $STUDENT \infty_{SID < SNUM} (ENROLLMENT)$

$STUDENT \infty_{SID > SNUM} (ENROLLMENT)$

8. Division ($\div$) – The division operator divides a divident relation A of degree m+n by a division relation B of degree n

DEND

S#	P#
S_1	P_1
S_1	P_2
S_1	P_3
S_1	P_4
S_1	P_5
S_1	P_6

S_2	P_1
S_2	P_2
S_3	P_2
S_4	P_2
S_4	P_4
S_4	P_5

DOR

P#
P_1

DOR

P#
P_2
P_4

DOR

P#
P_2
P_3
P_4
P_5
P_6

Page No.
Date

DEND & DOR

S#
S1
S2

S#
S1
S4

S#
S1

$P_2 = S_1, S_2, S_3, S_4$

$P_4 = S_1, S_4$