Database System Concept

CSE 3103

Last Sem ও Rupol sir-ই class নিয়েছেন।
(Nazmus Sakib)

Text books:

1. Database System Concepts, Siberschatz-Korth (6th Edition)
2. Database System: The complete Book, Molina-Ulman (3rd Edition).

Quiz — 4 (Best 3 will be counted)
always Sunday তে Quiz হবে (ক্লাসের আগে 2টি + ক্লাসের পর 2টি)

Nazmus Sakib
Asst. Professor
Room: 7A01/M, Block-A, AUST
Websites: sites.google.com/site/nazmussakib009/
Email: nazmussakib009@gmail.com
Cell: 01939900271

শনি ও মঙ্গল বিকাল ৫টা-৬টায় পাওয়া যাবে।

Scanned by CamScanner

What is a Database?

Database Management System (DBMS)

1টি sofware যা total data কে নিয়ন্ত্রণ করে। Backup করে রাখা যায়।

Drawbacks of using file systems to store data

24 April 2017

University Database Example

Levels of Abstraction

* physical level (how a record is stored) describes
* logical level (কী-কী ধরনের Data, তাদের Relation etc)
* view level (details of data types, information can be hidden)

Scanned by CamScanner

Instances and Schemas

- Logical scheme
- Physical scheme
- Scheme
- Sub Scheme
- Instance
- Physical Data Independence
- Logical Data Independence

Schema

— তিন প্রকার।

— total structure কে Schema বলে।

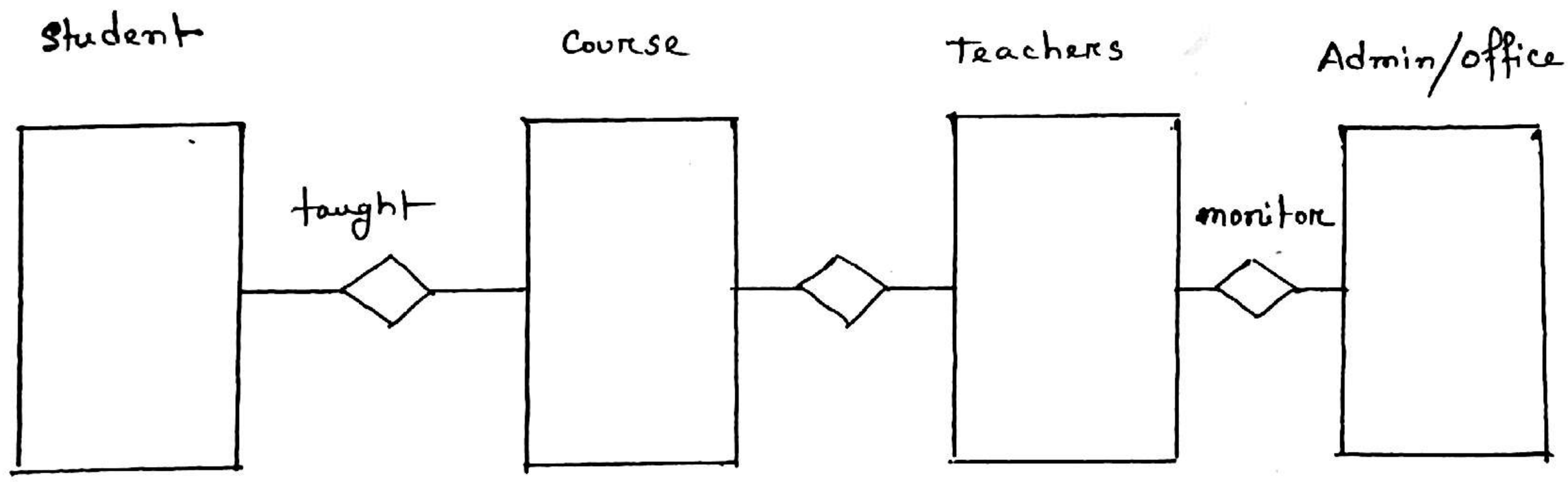


fig: Schema

Scanned by CamScanner

* Database এর "instance" create করা খুব জরুরী।

* Instance → নির্দিষ্ট moment এ Database এর snapshot.

* Industry level এ "Instance" কে "Snap Shot" বলা হয়।

* Physical level এ কিছু করলে যদি logical level এ কোনো impact না পড়ে, তাকে Physical data Independence বলে।

* logical data independence

** "Level of Database" - short Question আসতে পারে।

Database Users & DBA Activities

(self study)

DBA ⇒ Database Administrator

Scanned by CamScanner

Data Models (Just জেনে রাখব, xm ও আসবে না)

E-R Model ⇒ Entity-Relationship Model

* 850 - 1050 MHz frequency এর মধ্যে হতে হবে primary phone call dedicated path ও।

* ITU ⇒ International Telecommunication Union

* 6s এর মধ্যে (Max. 7s) Connect হতে হবে phone calls.

Data Definition Language (DDL)

Database ও Define করার কাজে ব্যবহৃত হয় through checking constraints.

Data Manipulation Language (DML)

Database ও নিজে ইচ্ছানুযায়ী search / query করার মাধ্যমে Data manipulate / control করা যায়, তাকে Data Manipulation Language (DML) বলে।

Scanned by CamScanner

30 April 2017

Database Engine

3rd part शुरू।

Storage Management

Student

ID	Name	Address	D.O.B.
1	A	abc	15 APR
2	B	xyz	12 Sep
3	C	mno	01 JAN
4	D	uvw	19 APR
5	E	stx	18 DEC

View/show

Select Name
FROM Student
where DOB = " ** APR" } Query

A 15 APR
D 19 APR } Showing Results

in Relational Algebra (Database এর language কোনো)

σ_{name} π student

(DOB = ** APR)

* "Optimization" এর main idea হচ্ছে "কম সময়ে বেশি কাজ করা"।

Query Processing

* দুই way তে Query করা যায় ⇒ (i) Top-Down
(ii) Bottom-Top

Transaction Management :

Database Users and Administrators :

Database Architecture

4 প্রকার।

Centralized Database

Client Server Database

ROLL 103466
BOARD DHK
Year 2016

Select *
From Result
where ROLL = "103466"
"DHK" "2016"

Faster Query

Select *
from Result
where Year = "2016"
Board = "DHK"
Roll = "103466"

Parallel Database

Q. Central ও Parallel Database এর পার্থক্য।

⇒ Central Database এ direct database এ hit হয়, where Parallel Database এ direct database এ hit এর পরে আলাদা আলাদা layer এ hit হয় এবং পরে central এর সাথে sync হয়।

Distributed Database

"RAID"

04 May 2017

Chapter-6

ER Modeling

Steps:

(1) Requirement Analysis: Firstly you need to generate an idea about a project, from there you need to develop a High Level Description which is also called ER Modeling. The symbols which are used

in ER Modeling that are representing the Requirements.

So,

Idea → High Level Description → Relationship Schema →
(ER Modeling) (Tables) (DDL/DML)

RDBMS ⇒ Implementation

ER Diagram :

The diagram which represents ER Mode that is ER Diagram.

ERD → Entity Relationship Diagram

There are 3 main parts in ERD.

(i) Entity
(ii) Relationship
(iii) Attribute

There are also many small parts in a ERD.

Scanned by CamScanner

Entity :

→ It's a noun altime

→ After Requirement Analysis we have to findout the Entity and Relationships.

→ Entity will be a real object from Problem Domain and Data will be stored in DB on Entity.

→ Distinguishable from other objects.

→ An entity is described using a set of attributes.

Example:

People	Places
Things	Organizations
Concepts	Events

Scanned by CamScanner

Example:

There are many students in AUST. students can enroll in different courses and receive Grades.

AUST → Specific info
Grades → Values ; no Attributes

Student →

Courses →

→ All entities in a entity set have some set of attributes.

→ Each entity set has a key (Underlined)

→ Each Attributes has a Domain.

Scanned by CamScanner

07 May 2017

Entity set: Collection of Similar Entities.

Entity Type: Collection of Entities that have same attributes.

Example: Student { ID, name, DOB, Phone }

Student

ID	Name	DOB	Phone
1	A	23 June	019
2	B	2 July	017

Relationship:

The interaction between Entity set

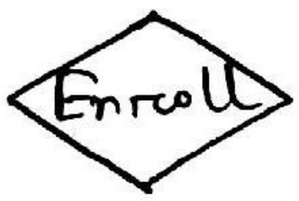
→ Relationship is always a verb

→ It can have own attributes

→ Association among two or more attributes

Example:

There are many [Students] in AUST. [Student] can



in different [Courses] and receive grades.

Student?

A	B
1	α
2	β

Courses?

C	D
3	x
4	y

Enroll

A	B	C	D	E
1	α	3	x	Yes
1	α	4	y	Yes
2	β	3	x	Yes
2	β	4	x	No

Communication
Interaction
Multiplication
Join

Attributes:

The properties of Entity or Relationship

→ Properties are remaining same for entity but value can change

→ If properties / attributes are described in story, use the given ones. Other than guess the attributes and use the meaningfull properties.

→ student

ID, Name, Parent Name, Address, DOB, Blood Group,
Phone, E-mail, Eye color??, Height??, Hair color??,
NID, Birth Certificate etc.

Primary Key:

The attribute which can distinguish one Entity from another can be a primary key.

→ Primary key should be underlined in ERD.

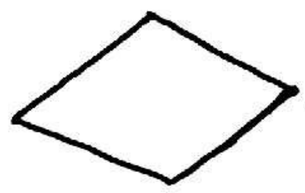
→ More than one attribute can form Primary key.

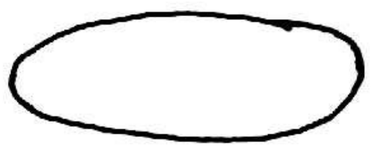
→ Primary key will be minimum number attribute.

→ Primary key can be the attribute which is most frequently use and monthly use to form.

Symbols :

Entity Set : Rectangle

Relationship :  Diamond

Attribute :  Ellipse

Model : Crows Feet
OO ER

* 1টি entity অপর entity এর সাথে connected হতে হলে, তাদের মাঝে অবশ্যই 1টি Relationship থাকা লাগবে।

* Entity => noun

Relationship => Verb

* কোনো 1টি Relationship create করা => Table Create করা।

Scanned by CamScanner

⊛ **Entity :** Student Courses

Relationship : Enroll

Attributes :

Student

Student ID, Student Name, DOB, Phone, Blood Group

Course

Course ID, Course Name, Credit

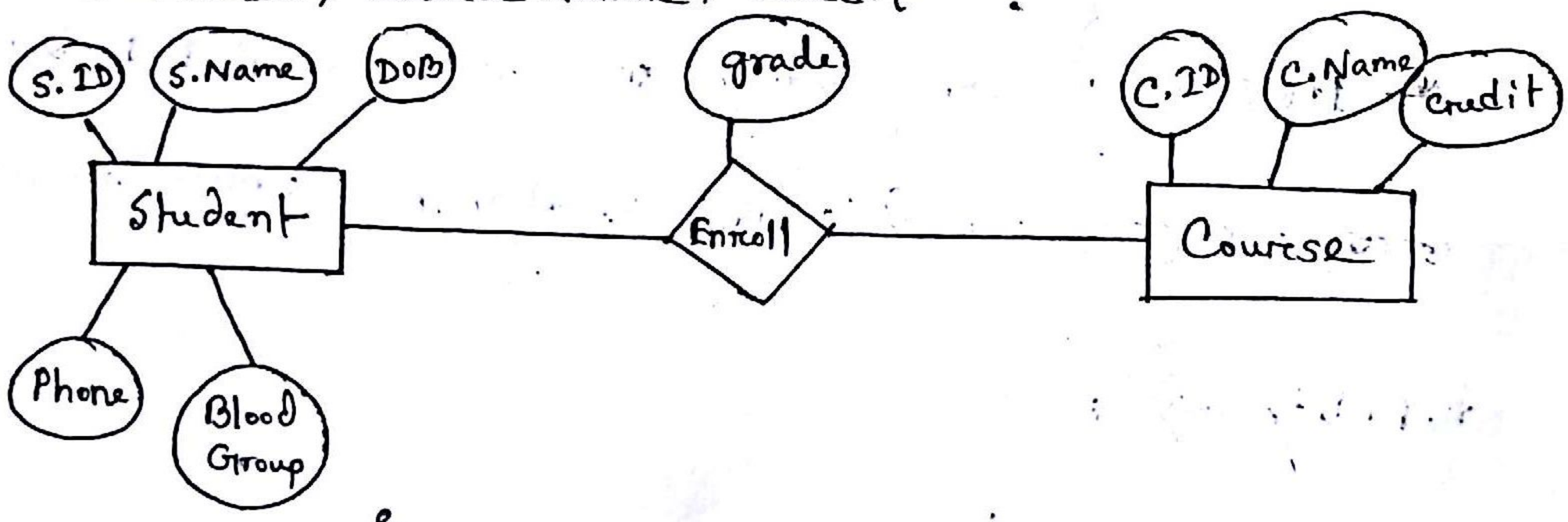


fig : ERD Diagram

S. ID	S. Name	DOB	Phone	Blood G.	C. ID	C. Name	Credit	Grade
1	A	"	"	"	CSE 3109	DB	3.0	A+
2	A	"	017	"	CSE 3109	DB	3.0	A-

08 May 2017

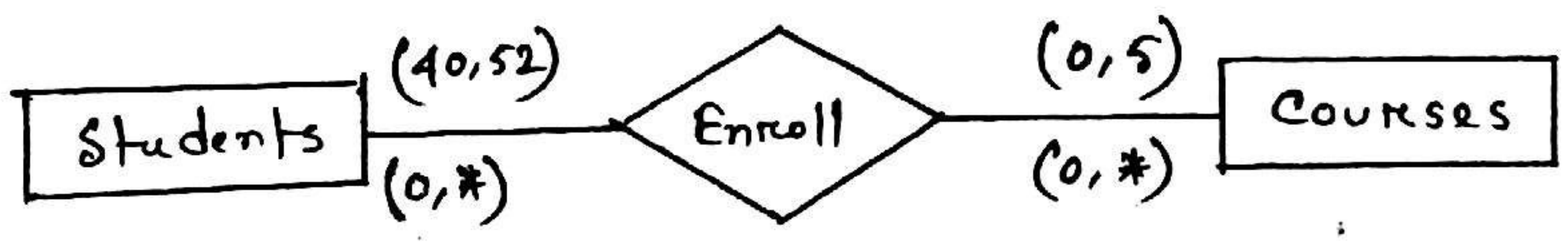
Cardinality Constraint of Relationship

Cardinality Ratio

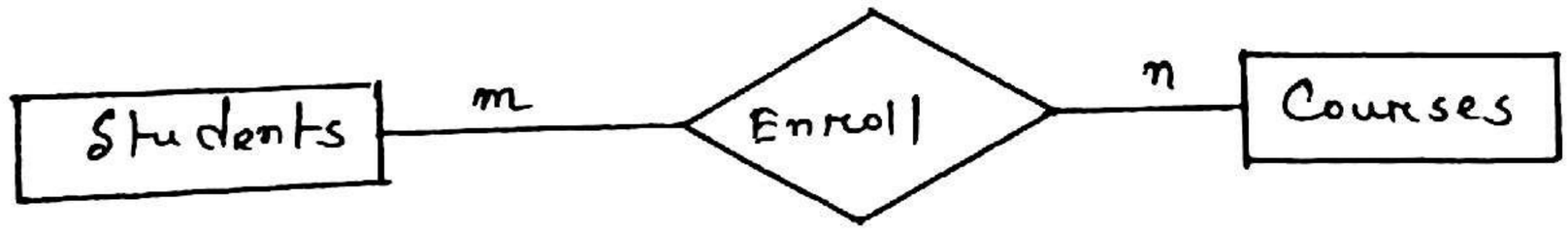
- Express the maximum number of Entity that can be associated with another Entity via a relationship.

4 types of Cardinality Constraints.

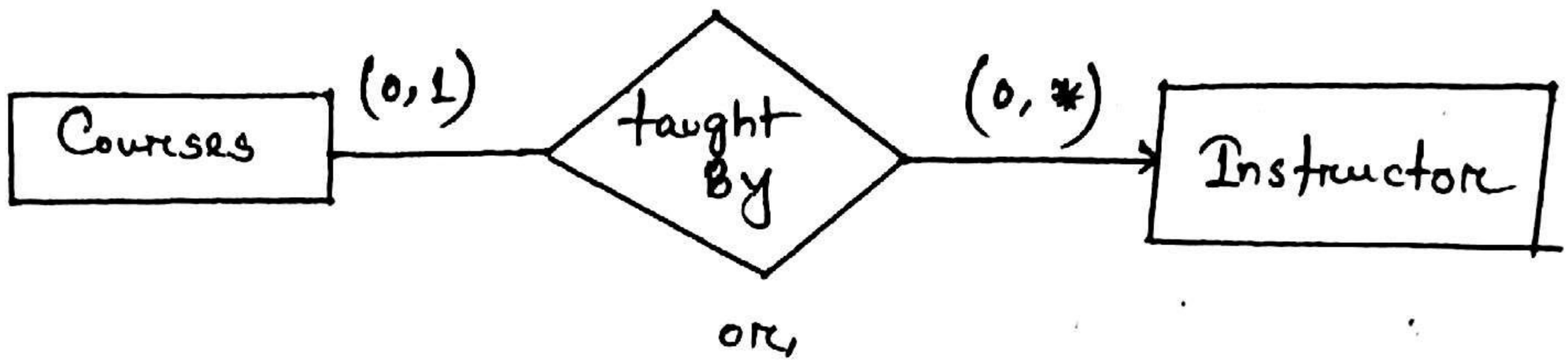
1. Many - To - Many



or,

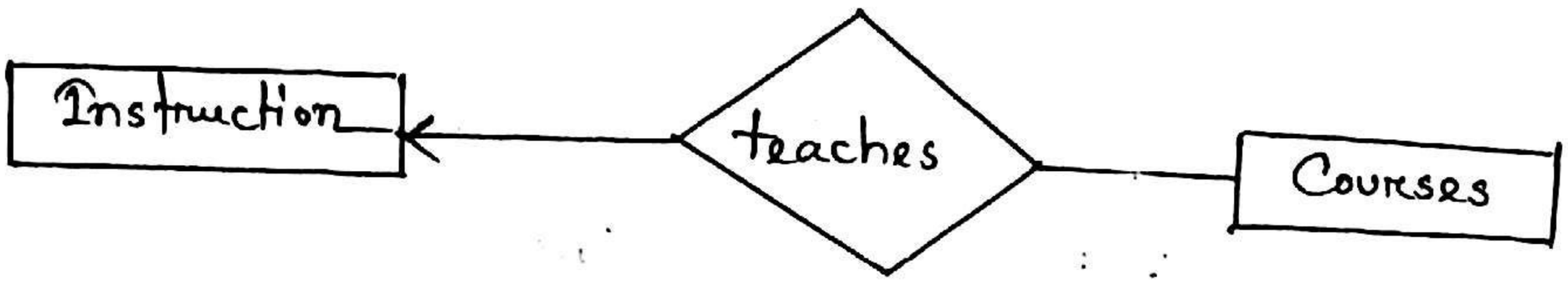


2. Many - to - One

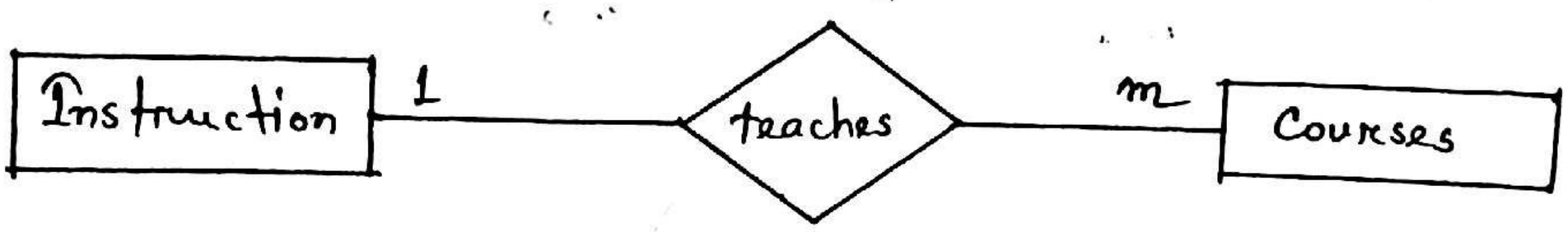


or,

3. One - To - many

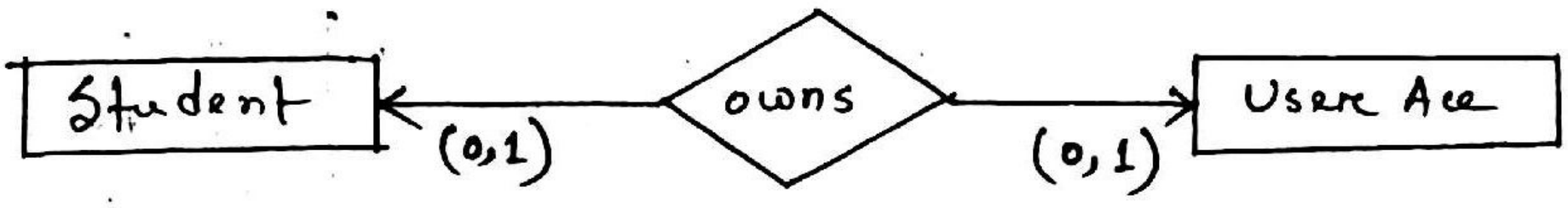


or,

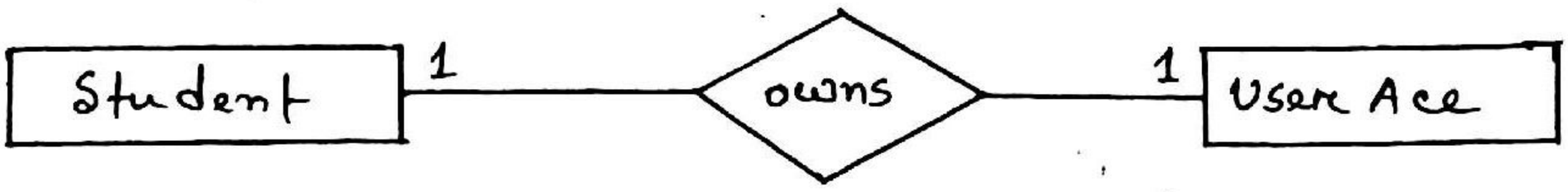


* Arrow এর representation avoid করার চেষ্টা করব।
* 2nd Representation টি follow করব।
* বই এর example গুলো practise করব।

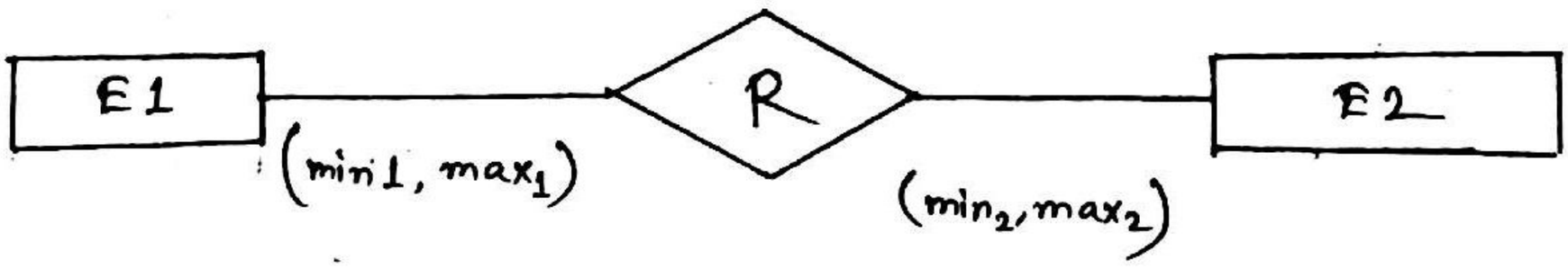
4. One-to-One



or,



Cardinality Limit / Degree of Constraints



Relationship	(min_1, max_1)	(min_2, max_2)
Many-to-Many	(0, *)	(0, *)
Many-to-One	(0, 1)	(0, *)
One-To-Many	(0, *)	(0, 1)
One-To-One	(0, 1)	(0, 1)

Scanned by CamScanner

Quiz-01 [14.05.2017]

Syllabus → slide 1, 2, 3 [uploaded]

[slide is not good to get 20/20]

বই এর Chapter-1, 2

[→ MCQ
→ Short Question
→ 3 or 2 Marks এর Ques আসতে পারে]

(*)

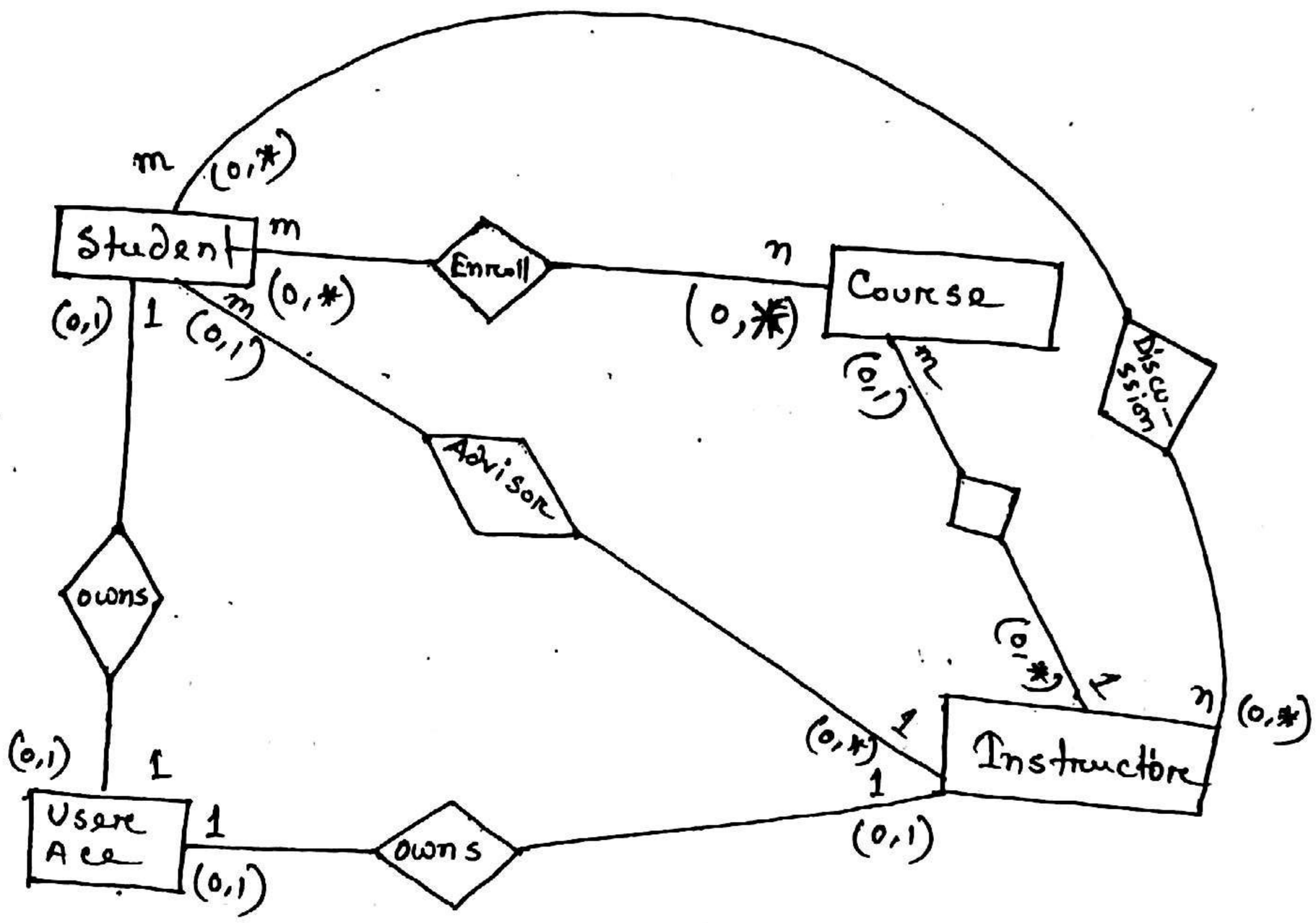


fig: Generic Representation

⊞ '*' থাকলে কোনো Numeric Value use করা যাবে না।

11 May 2017

Example 01: Student take courses. and each student belongs to a particular Department. Students' grade in different courses are stored. Each department has multiple students and a Department offers multiple courses. A course can be offered by a single Department or multiple departments.

Entity:

Student, course, department

Relationship:

take, belongs, offers

Attributes:

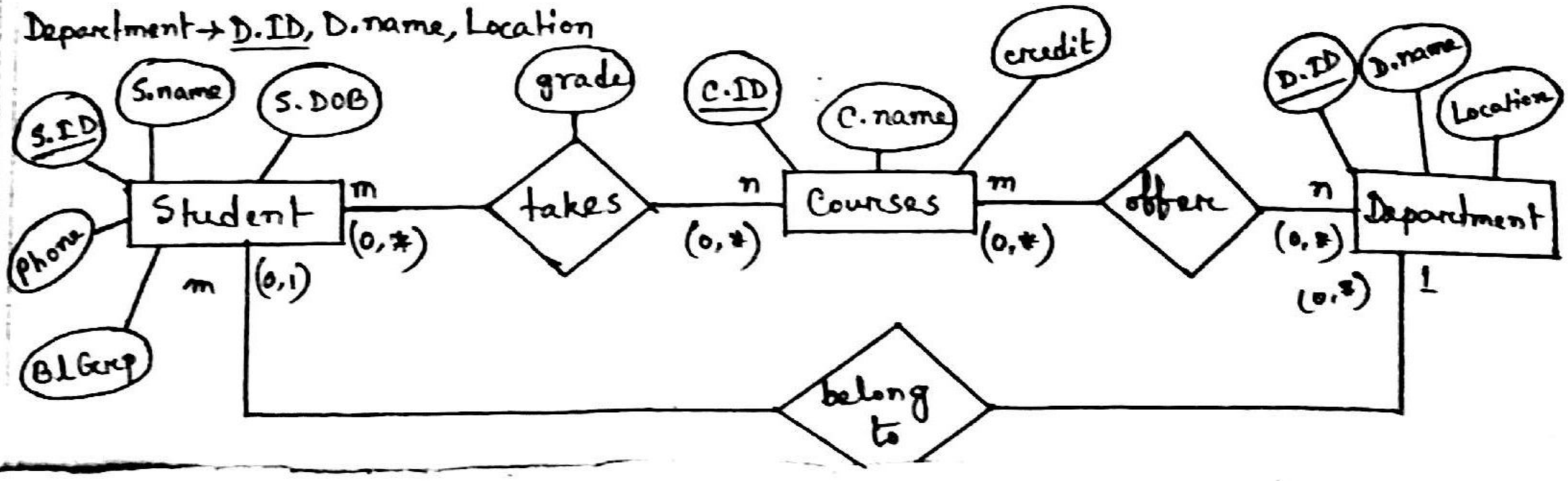
Student → S.ID, S.Name, S.DOB, Phone, Blgrp.

Course → C.ID, C.name, credit

Department → D.ID, D.name, Location

Cardinality Constraints:

ratio



Scanned by CamScanner

Example-02: Patients are treated in a Single ward by the doctors assigned to them. Usually each patient will be assigned by a single doctor but in rare cases they will have two. Assistants also attend to the patients and a number of these assistants are associated with each patient.

Entity: Patient, ward, doctor, Assistant

Relationship: treated, assign to, attend

Attributes:

Patients → P. ID, Name, Age, Bl group, Phone

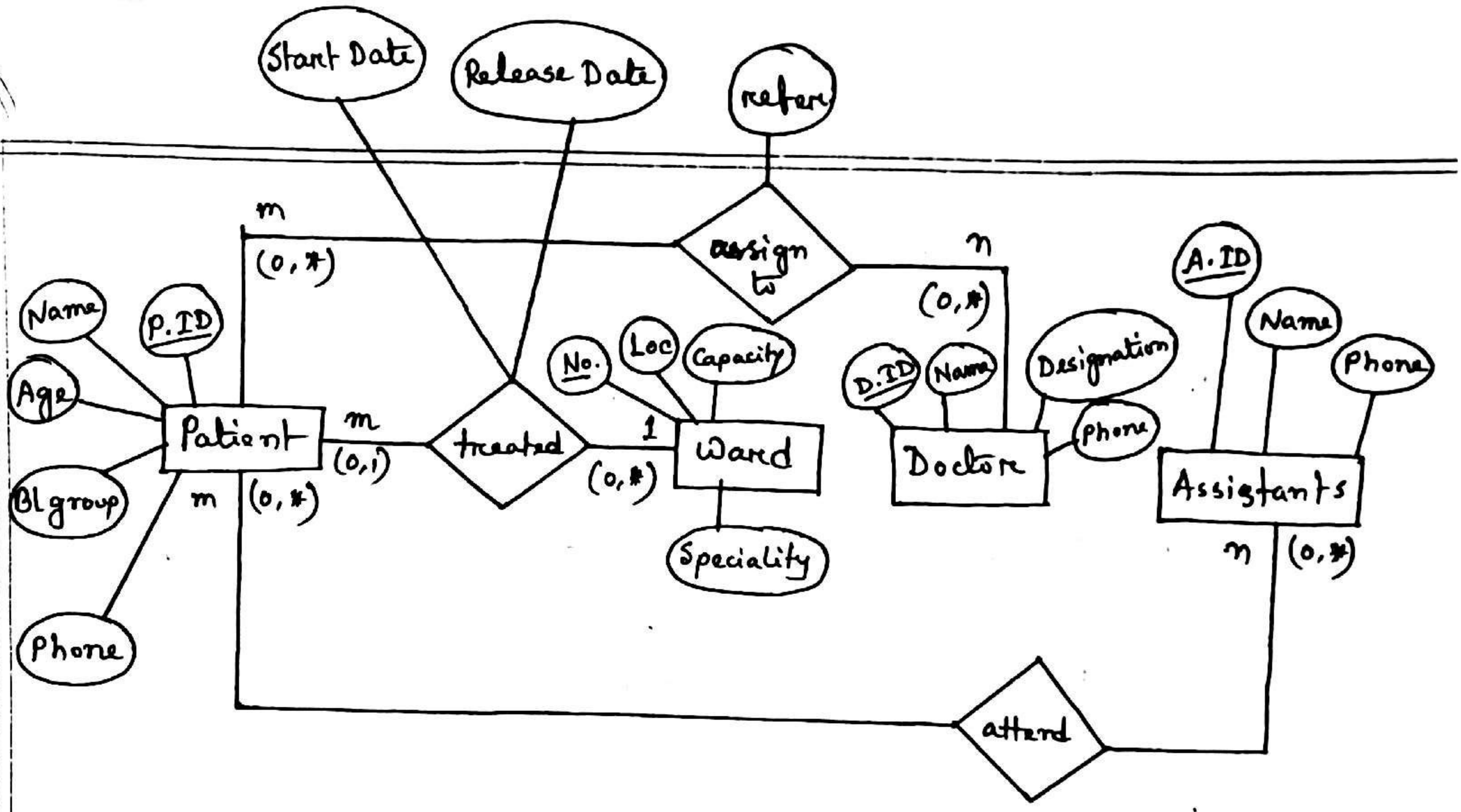
Ward → No., LOC, Capacity, Speciality

Doctor → D. ID, Name, Designation, Phone

Assistant → A. ID, Name, Phone

Cardinality Constraints:

Scanned by CamScanner



* দুটি entity এর মধ্যে multiple relationship থাকতে পারে।

* প্রথমেই noun গুলোকে sentence থেকে বের করে নিবে (Underline করে)।

* exam এ "5/6" টি Attribute তুলে ধরলেই হবে।

Scanned by CamScanner

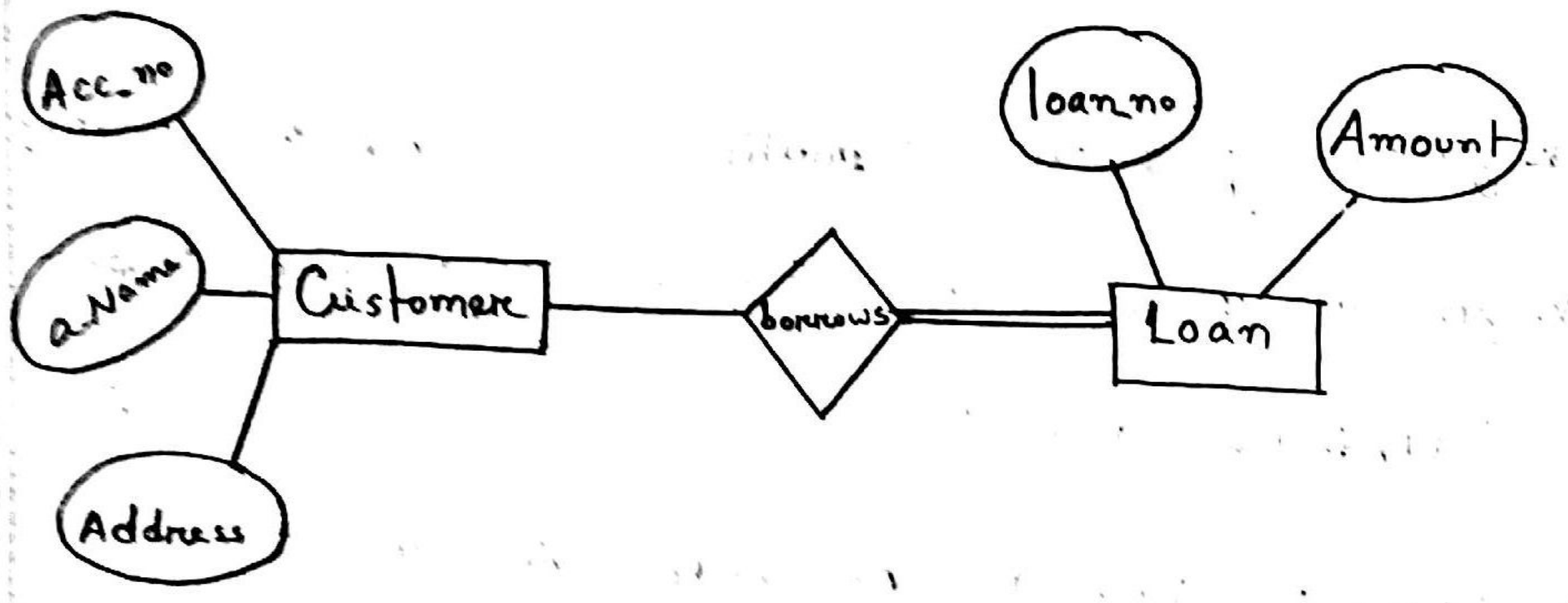
15 May 2017

Participation Constraints

1) Total Participation ═══ (double line)

2) Partial Participation ——— (single line)

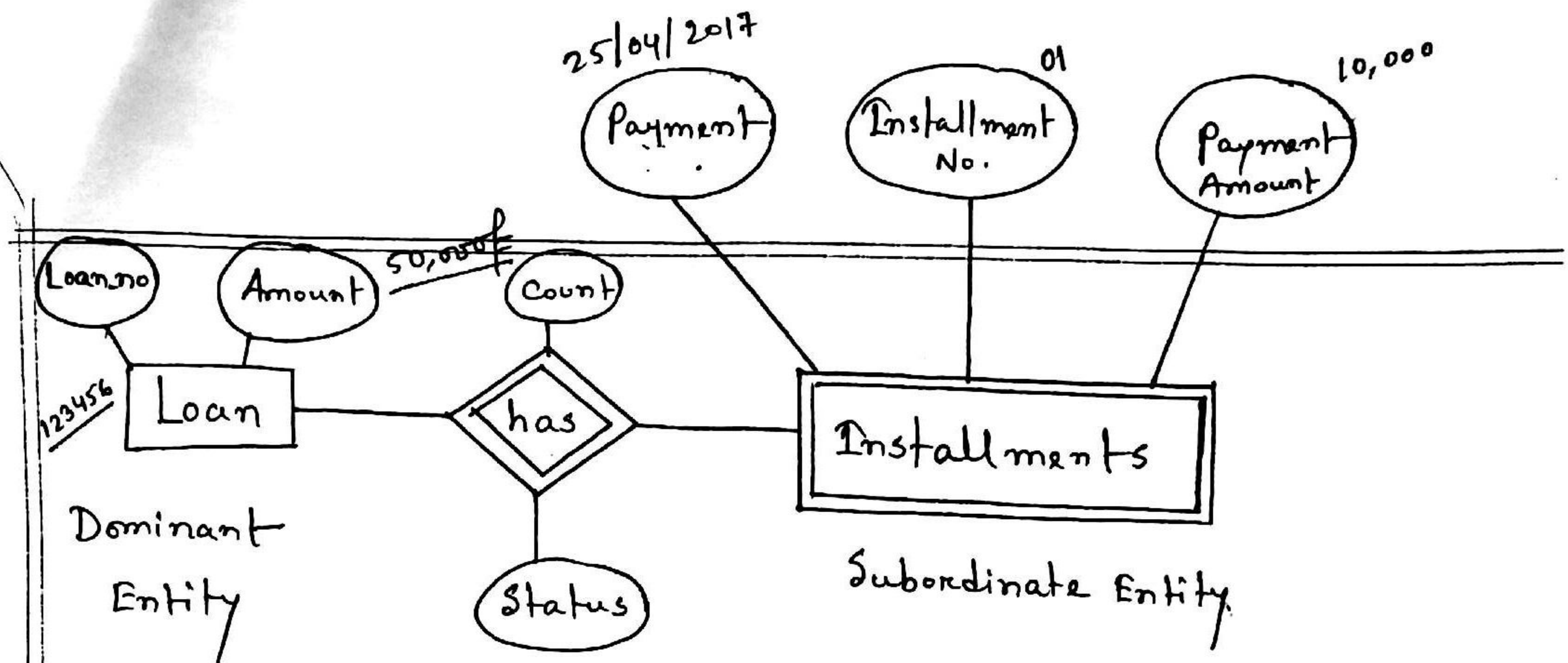
Example : Bank



Weak Entity Set: One-to-many relationship

Dominant Subordinate

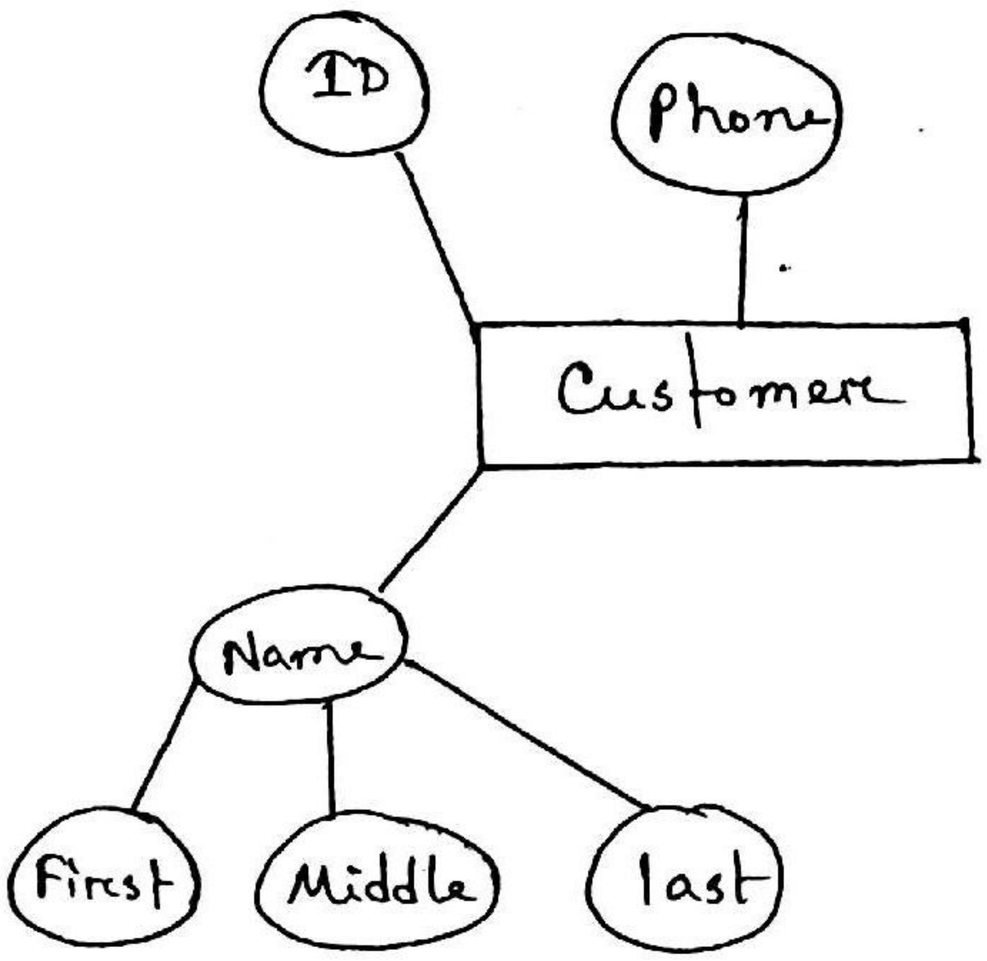
— Weak, entity, indentifying entity, discriminator, indentify relationship.



loan	Amount	Account
123456	50,000	ABC1200
123789	60,000	XYZ987

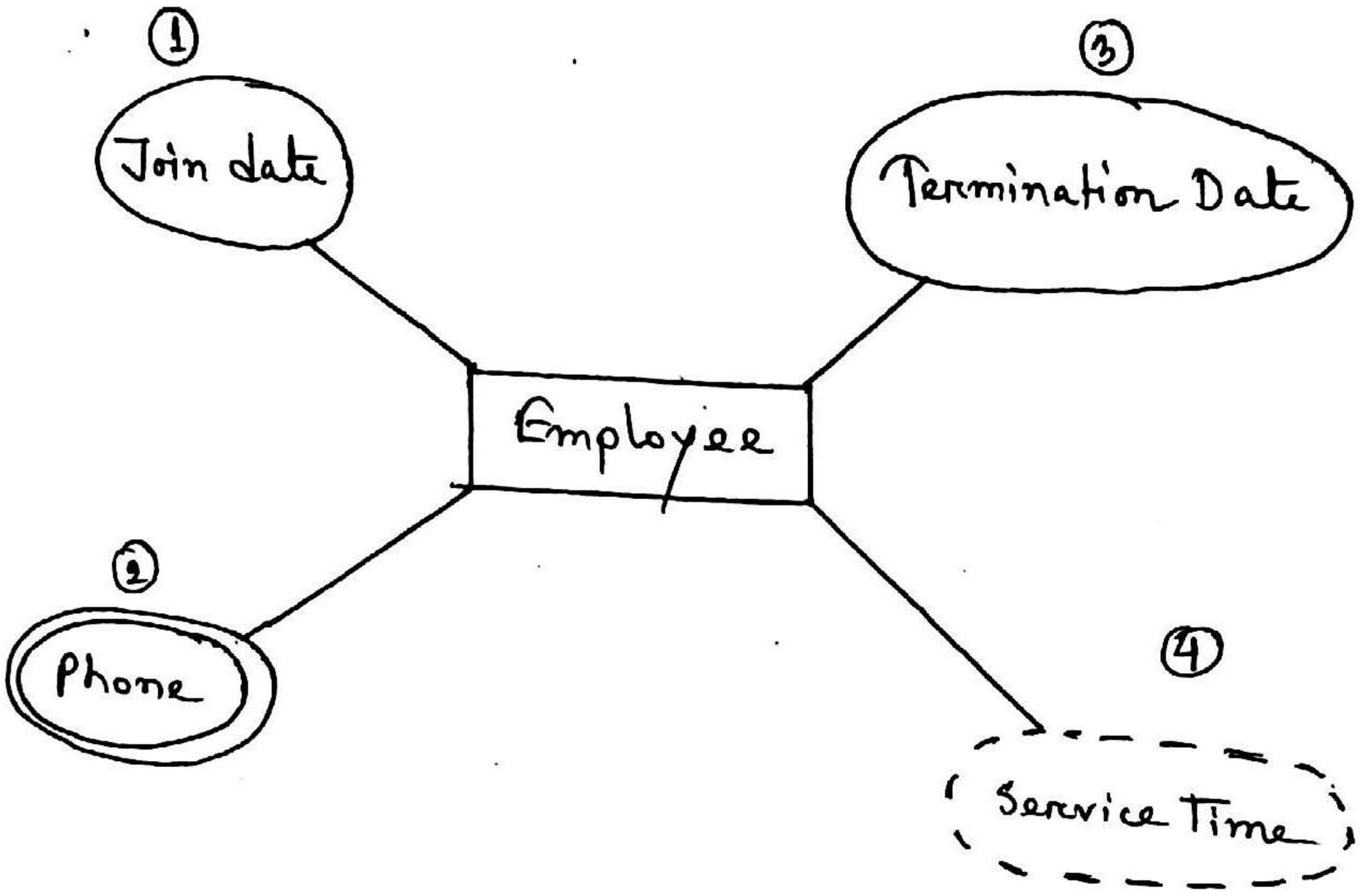
Attribute Type

Simple & Composite



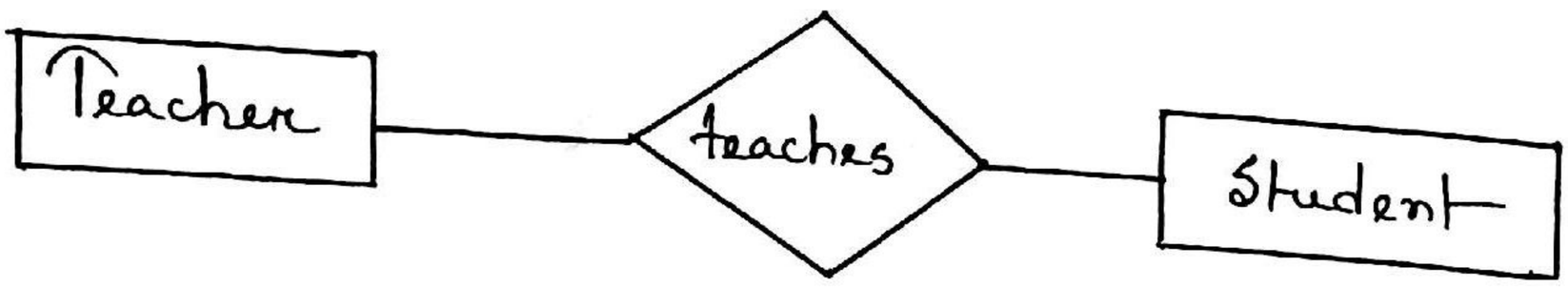
Scanned by CamScanner

① Single Value, ② Multi Value, ③ Null and ④ derived



Degree of Relationship

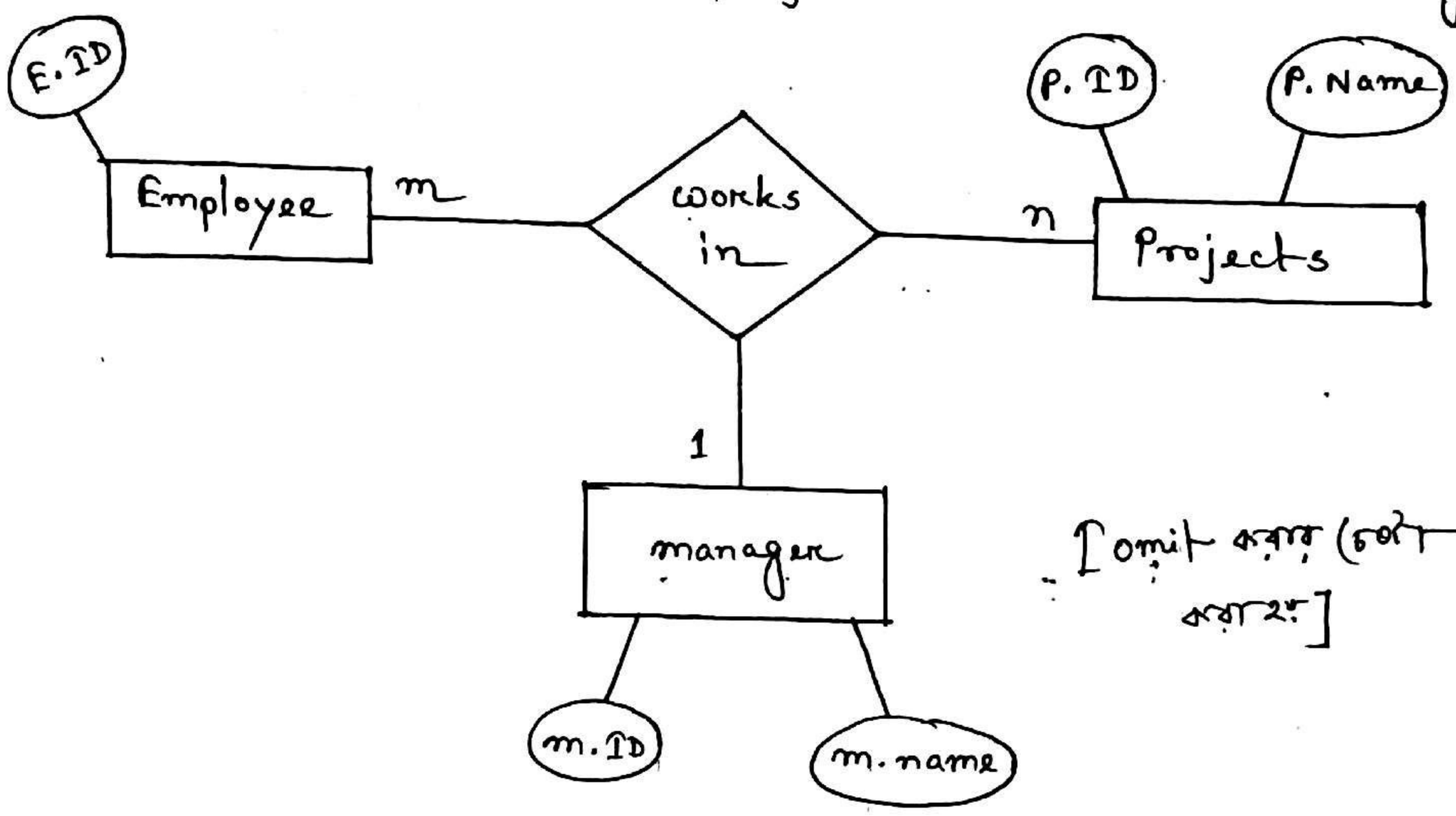
Binary



Scanned by CamScanner

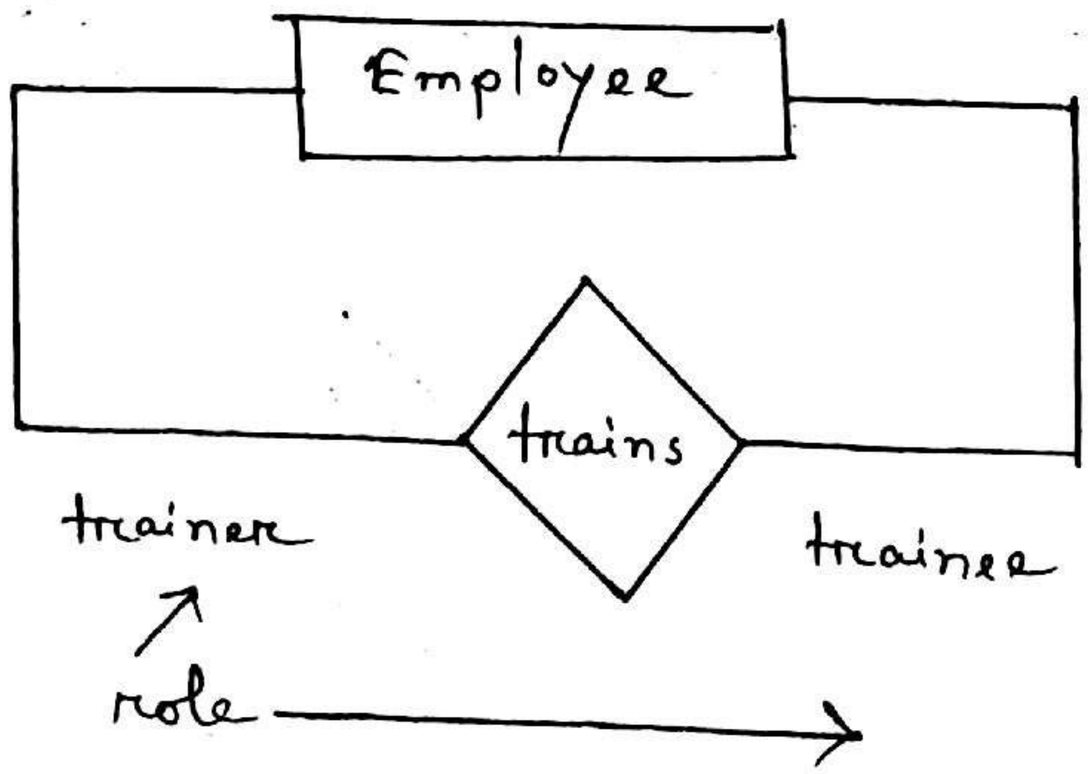
Ternary:

A company has several projects, employees can work on several projects under one manager.



[omit করার (চেষ্টা করা হয়]

Unary / Recursive

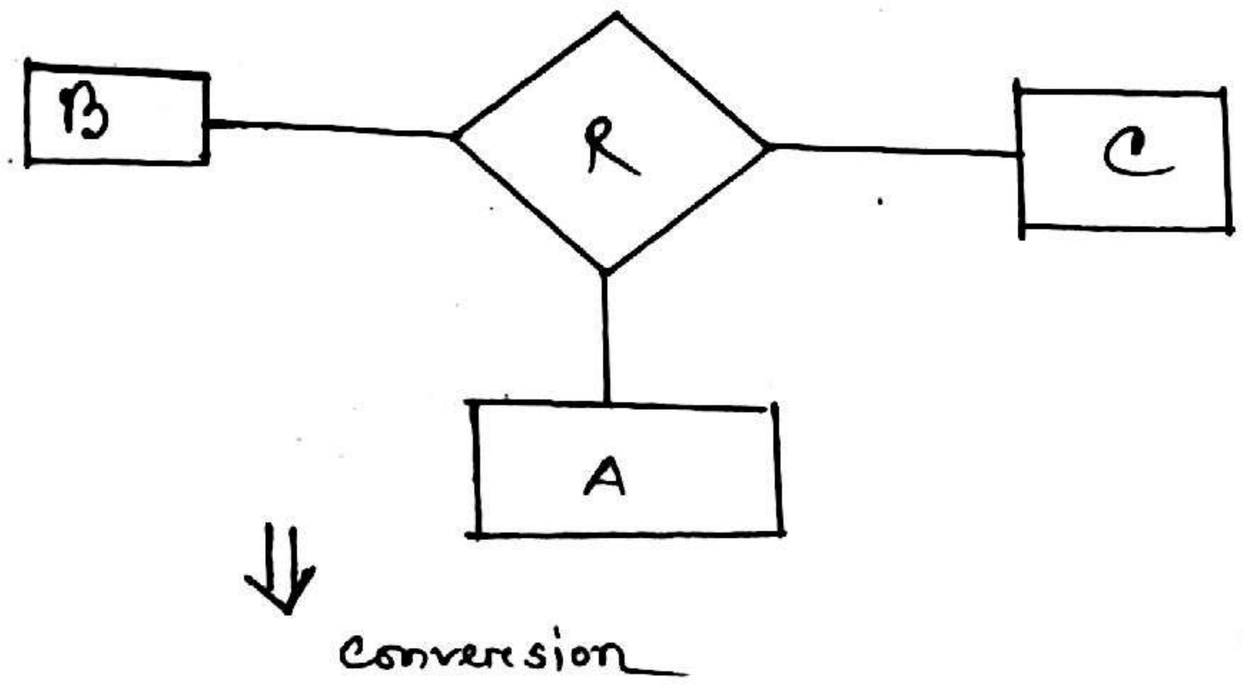


Scanned by CamScanner

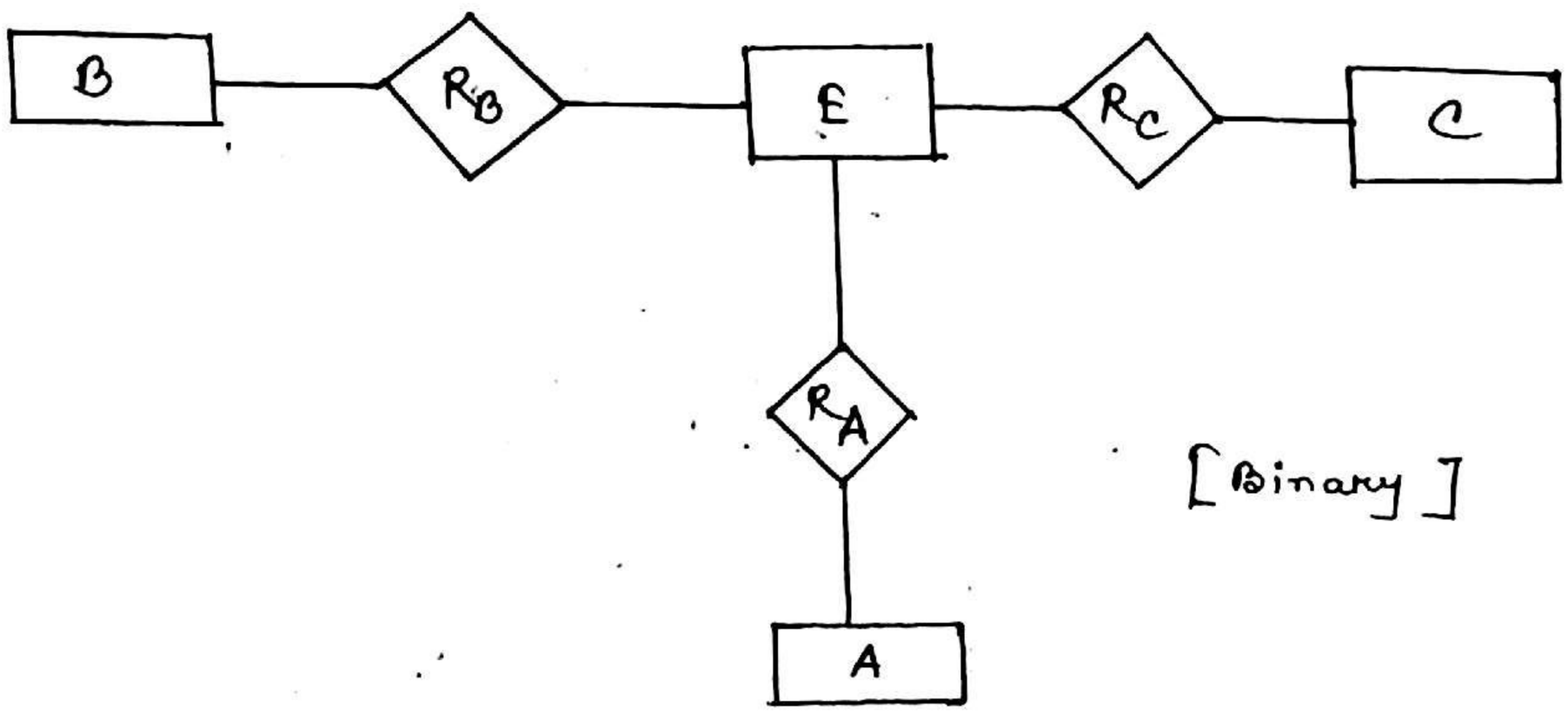
N-ary : More than 3.

⇒ Try to avoid Unary and Ternary relationships as their implementation is not easy all time.

Converting non-Binary to Binary Form



[non-Binary]

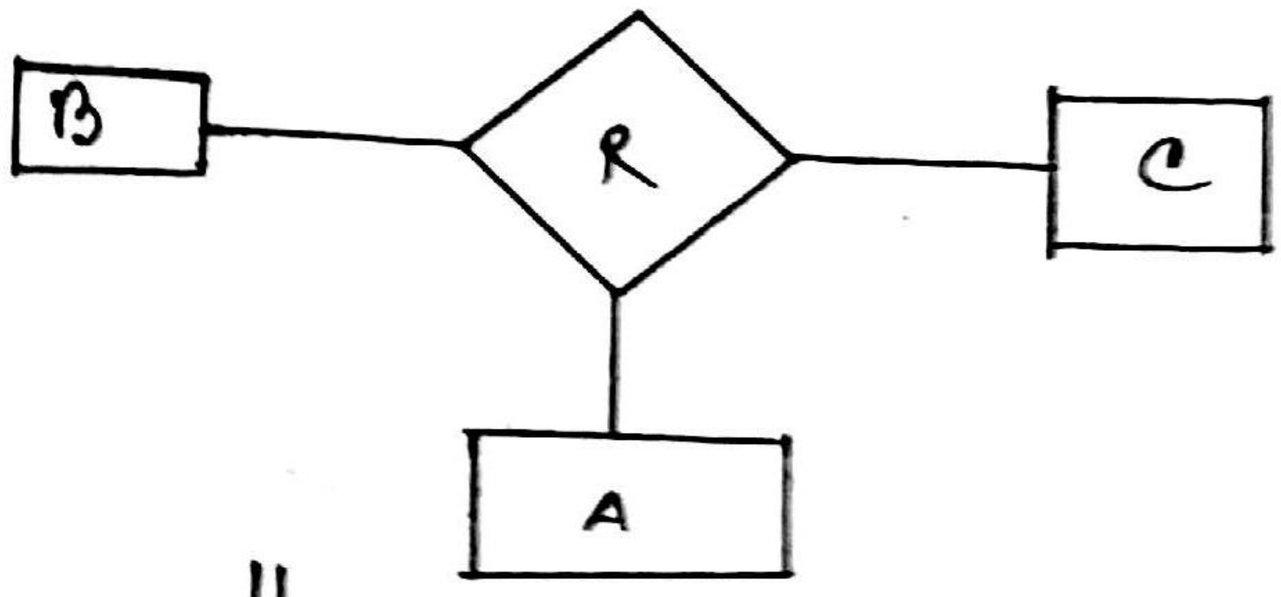


[Binary]

N-ary : More than 3.

⇒ Try to avoid Unary and Ternary relationships as their Implementation is not easy all time.

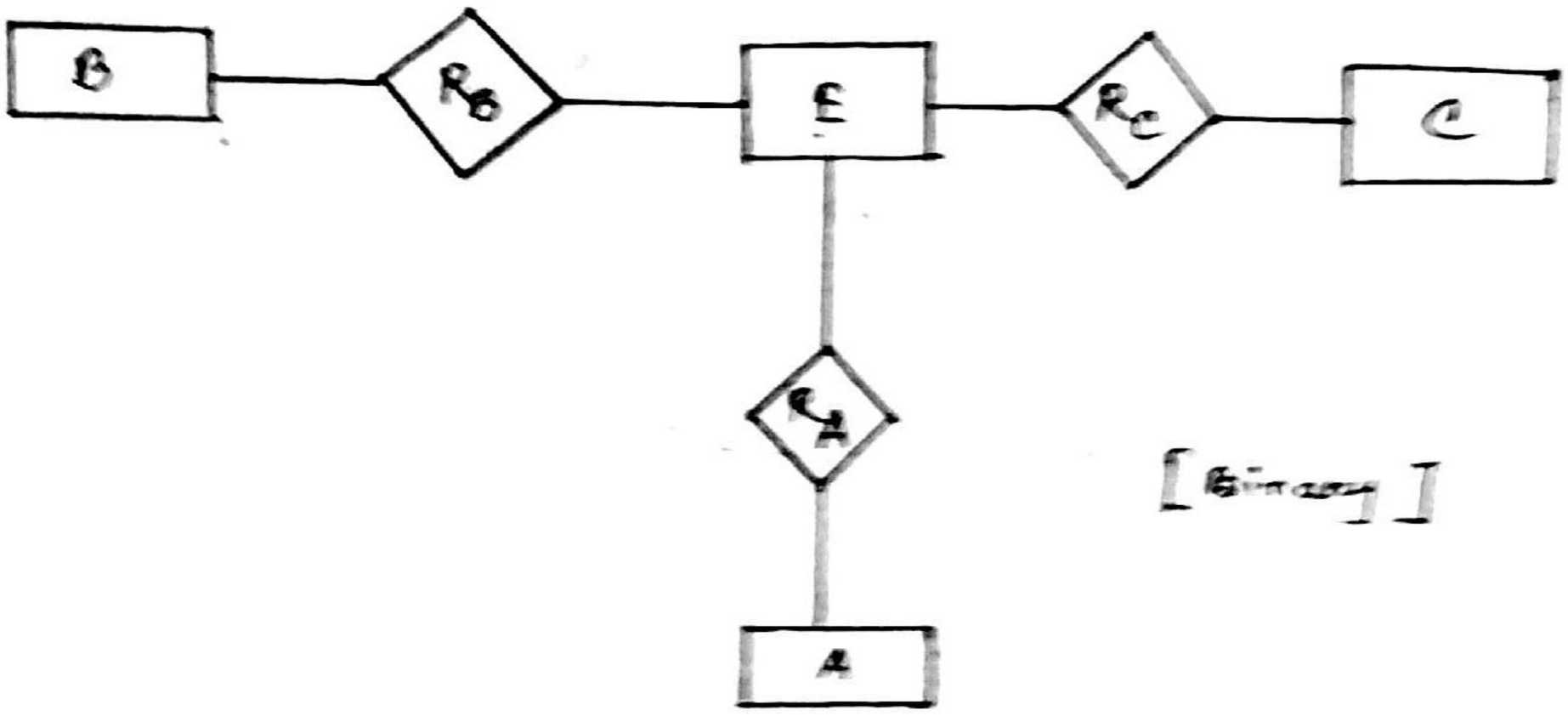
Converting non-Binary to Binary form



[non-Binary]

⇓

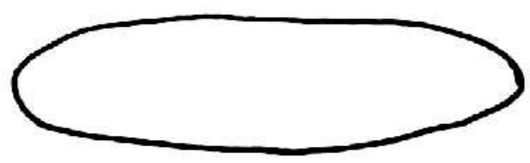
Conversion



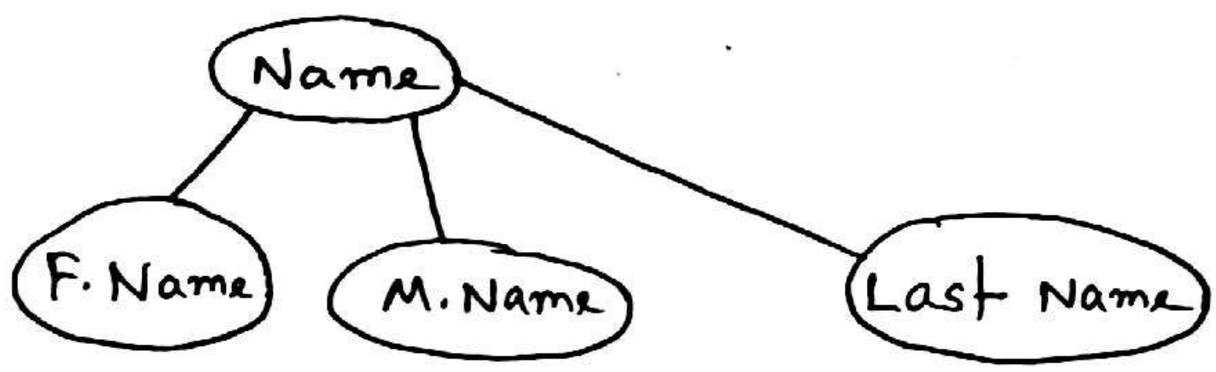
[Binary]

Explanation

Simple



Composite



→ single value

→ Null value

→ Multi value

→ Derives [Calculation করতে হয়]

Quiz-02 [04 JUNE 2017] (Sunday)

Syllabus → ERD

18 May 2017

Keys :

Suppose, A = {1, 2, 3, 4}

A → Entity Set

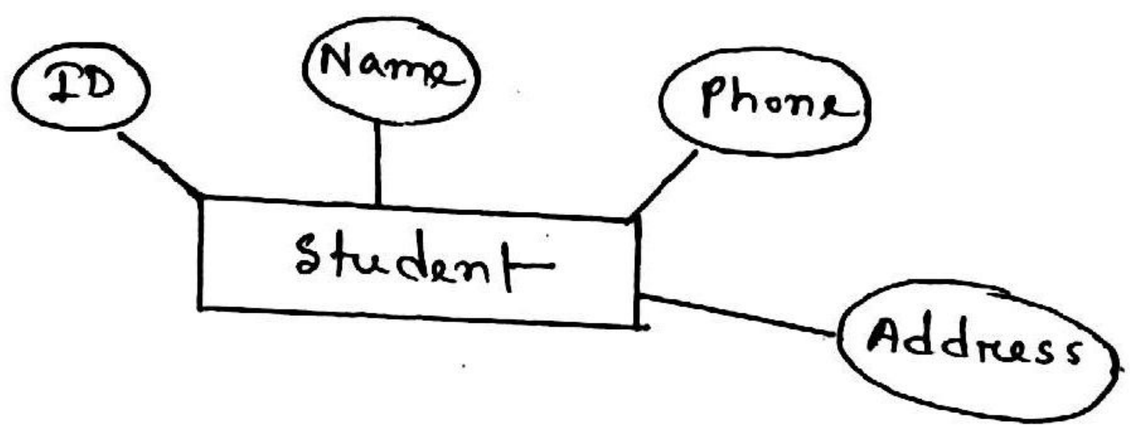
1, 2, 3, 4 → Attributes

$2^4 = 16$

2^n

P(A) = ? {1} {2} {3} {4} {1, 2}, ...

i) Super Keys : (কোন Entity Set এর সকল Attribute কে নিয়ে গঠিত



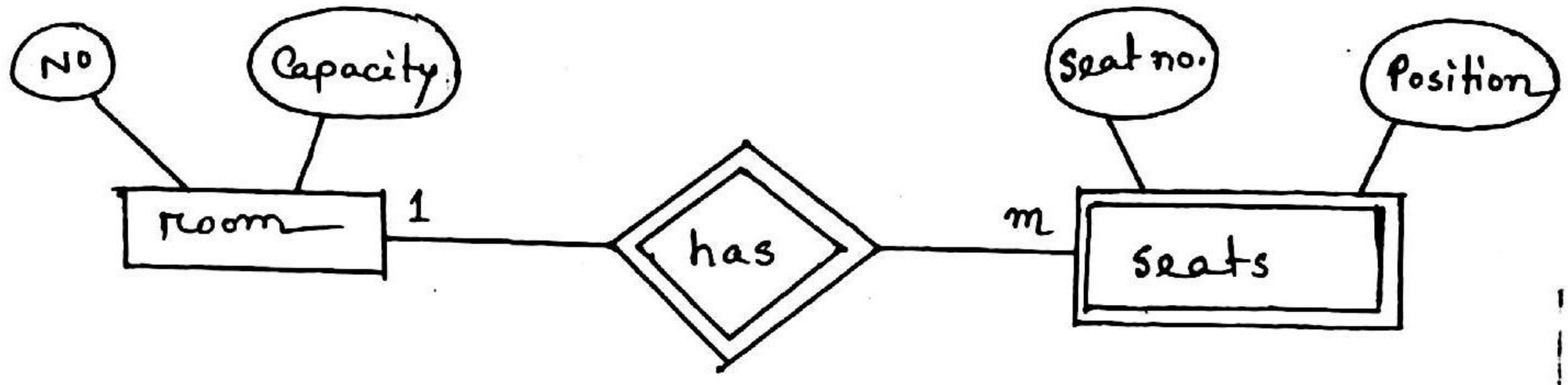
ii) Candidate Keys (সমস্ত super key, যার value minimal)
{ID}, {Phone}, {ID-Name}, {ID-Phone}

iii) Primary Key (সমস্ত Candidate Key যার value Unique)

iv) Alternative Key (Primary Key select করার পর rest of the candidate keys are Alternative Keys)

v) Composite Key (একাধিক Attribute নিয়ে 1টি Primary key declare করলে, তা Composite Key)

vi) Foreign Key (এমন একটি key যা অন্য table থেকে আসে এবং সে table এতে আসে তার Primary key)



* Foreign key is mostly used in Weak Entity set.

Enhanced ER Modeling

→ In some cases, an Entity type has a number of subgroups.

→ Relationships and attributes of superclass are inherited to sub-class.

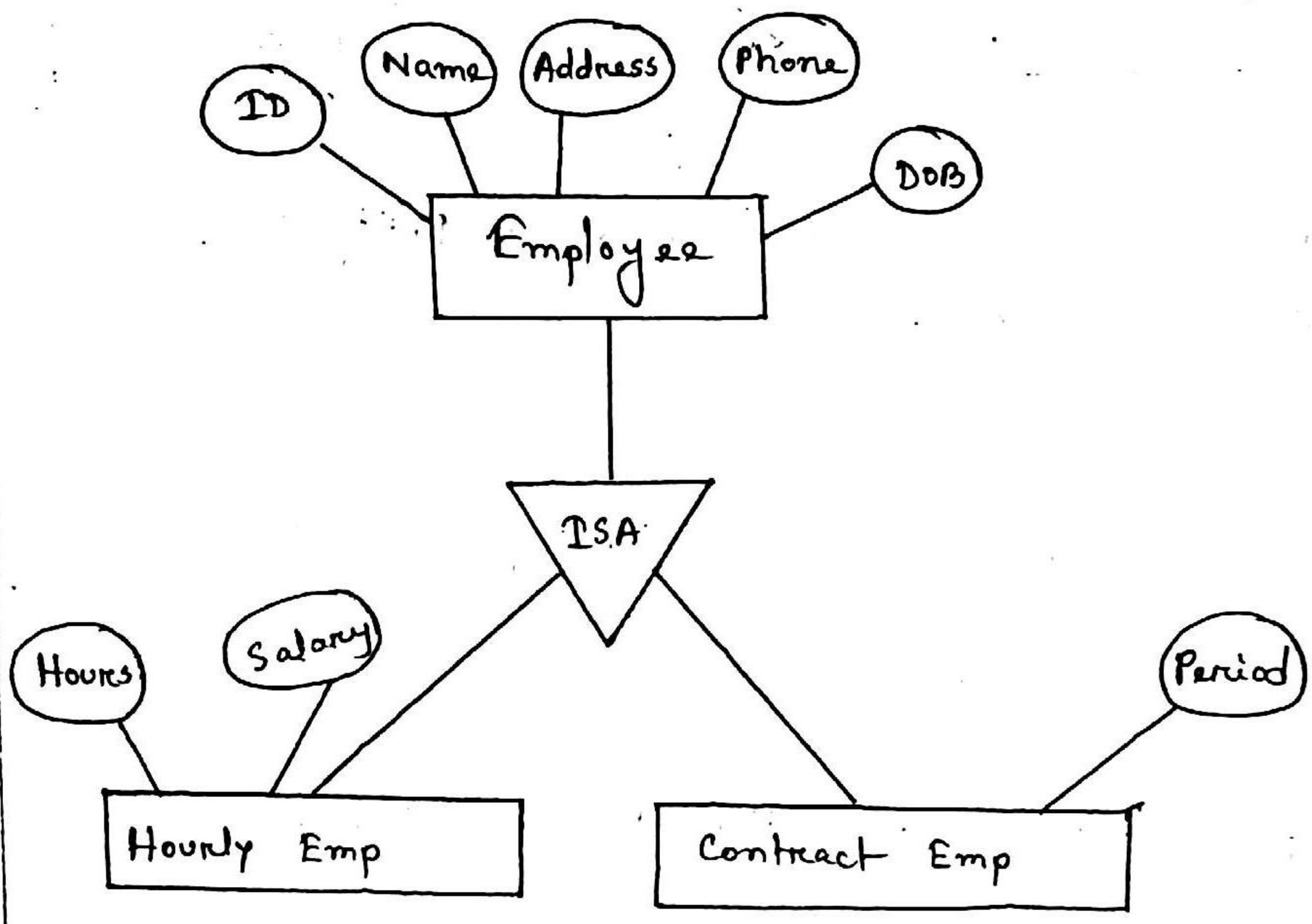
→ Sub class can have additional attributes and relationship.

Two types: i) Specialization

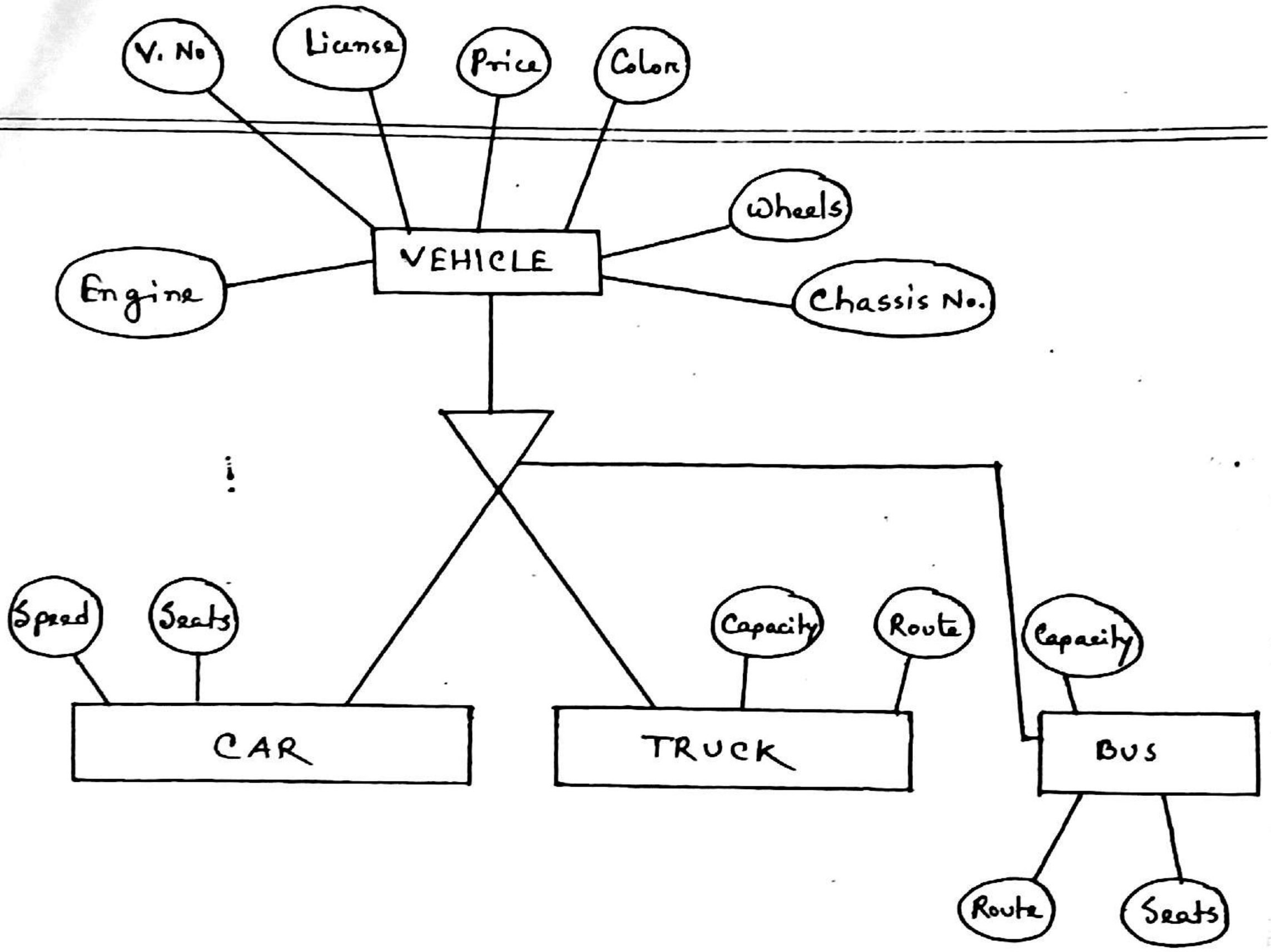
ii) Generalization

Scanned by CamScanner

Specialization: Process of defining a set of sub classes of an Entity type (top down ↓)



Generalization: Reverse process of Specialization (↑ bottom up). Identify common features of Entity types and generalize them into a single super class (Including Primary Key)



Scanned by CamScanner

21 May 2017

Relational Algebra:

The relational algebra is a theoretical language with operations that work on one or more relations (table) to define another relation without changing the original relation. It is used to manipulate relations to obtain a desired result.

Relational Algebra Operations:

Union, intersection, difference, Division, renaming Selection, projection, cartesian product, join etc.

* Unary Operation:

Operate in one relation (1 table).

Example: Selection, projection etc.

* Binary Operation:

Operate on two relations

Example: Union, Intersection, join etc.

Scanned by CamScanner

* N-ary Operation:

(দুই এর অধিক table নিয়ে Operation)

Selection Operation:

(Condition নিয়ে কাজ করে)
(Specific কোন কোন Row কে select করাই)

Syntax:

$\sigma_{predicate}(R)$

here,
σ = Selection (sigma)
predicate = Condition
R = relation

Symbols:

< > or ≠
>, <, ≥, ≤, =
∨, ∧, ¬
↓ ↓ ↓
OR AND NOT

Example:

student

Roll	Name	GPA
01	Rupa	4.5
02	Rimi	3.5
03	Shuvo	2.5
04	Sami	3.2
05	Poly	2.7
06	Keya	2.9

Scanned by CamScanner

Find student whose GPA is > 300

$\sigma_{GPA > 3.0}$ (Student) ✓ (Just এই line টি Ans করব xm এ)

ROLL	Name	GPA
01	Rupa	4.5
02	Rimi	3.5
04	Sami	3.2

[xm এ Table আঁকা লাগবে না]

Employee

Name	Age	Salary
Marly	25	9000
Luky	40	3000
Mark	36	4500
John	42	3900

1. Find the Employee whose age is less than 30 years.

   ⇒ $\sigma_{Age < 30}$ (Employee)

2. Find the Employee whose salary is more than 4000

   ⇒ $\sigma_{salary > 4000}$ (Employee)

3. Find the Employee whose age is less than 30 and get 4000 salary?

   ⇒ $\sigma_{Age < 30 \wedge salary = 4000}$ (Employee)

4. Find the Employee whose get 4000 or age below 30?

   ⇒ $\sigma_{salary = 4000 \vee age < 30}$ (Employee)

Scanned by CamScanner

5. Find the Employee whose name is 'Luky'.

$\Rightarrow \sigma_{name = 'Luky'} (Employee)$

Projection Operation:

Syntax:

$$\Pi_{a_1, a_2, \ldots, a_n} (R)$$

Here,

Π = projection (pie)

$a_1, \ldots, a_n$ = attributes

staff

Name	Gender	Date of Birth	Salary
Rony	M	01/05/86	20,000
Tony	M	05/07/80	30,000
Jony	M	07/02/95	40,000
Rupa	F	03/09/96	35,000

Scanned by CamScanner

* Show the names of staff Relation.

$\Rightarrow \Pi_{name}(staff)$

Name
Rony
Tony
Jony
Rupa

* Show the names and the salary of staff Relation.

$\Rightarrow \Pi_{name,\ salary}(staff)$

* Show the names of the staff whose salary is greater than 30,000 in the relation staff.

$\Rightarrow \Pi_{name}\left(\sigma_{salary > 30,000}(staff)\right)$

Scanned by CamScanner

* Projection operation is used to show the columns of a relation.

Select Name

From staff

where . . . . .

22 May 2017

Task

Customer

Name	Street	City	Phone	Age
Ruhi	abc	DHAKA	123	12
Adnan	def	KHULNA	456	16
Upoma	ghi	RAJSHAHI	789	17
Sazzad	jkl	DHAKA	135	19

Scanned by CamScanner

i) Find the customers name who lives in Dhaka.

Ans. $\Pi_{name}(\sigma_{city='dhaka'}(customer))$

ii) Find the customer name who age is more than 15 and lives Dhaka where his/her phone numbere starts with 1.

Ans. $\Pi_{name}(\sigma_{age>15 \wedge city='dhaka' \wedge phone='1\%'}(customer))$

Binary Operation Compatibility :

To perform union, intersection, difference operations, relations should be UNION compatible. 2 relations are UNION compatible if they have some number of attributes and belong to the same domain.

→ Column number need to be same.

→ Domain Type need to be same.

Scanned by CamScanner

R

A	B

S

A	B

Compatible

R

A	B

S

C	D

Not Compatible

Suppose,

R

A	B
α	1
α	2
β	1

S

A	B
α	2
β	3

Scanned by CamScanner

Union Operation :

$R \cup S$ is

A	B
α	1
α	2
β	1
β	3

Intersection Operation :

$R \cap S$ is

A	B
α	2

Difference Operation :

$R - S$ is

A	B
α	1
β	1

$S - R$ is

A	B
β	3

Scanned by CamScanner

Rename Operation : [table এর এক instance create করে
Attribute " " Attribute Name change করে]

→ When tables are not compatible

→ When need to compare same relation

To rename a table. $\rho_{A}(B)$ → old table-name
↓
new table-name

To rename attribute

$$\rho_{\text{old name} \rightarrow \text{new name}}(R)$$

Example:

Employee (name, branch, salary)

after,

$$\rho_{\text{branch, salary} \rightarrow \text{location, pay}}(\text{Employee})$$

So, we will get,

Employee (name, location, pay)

Scanned by CamScanner

Division Operation:

R

A	B
a_1	b_1
a_1	b_2
a_2	b_1
a_3	b_2

ভাগ্য

S

B
b_1
b_2

ভাজক

$R/S =$

A
a_1

ভাগফল

$$\frac{R}{S} = \frac{AB\cancel{C}}{\cancel{C}}$$

= AB Column

$$\frac{R}{S} = \frac{A\cancel{B}}{\cancel{B}} = A \text{ Col}$$

* এখন A হল Ans Col.
* Answer হল ভাজক এর কোন কোন element গুলো ভাগ্য এর সাথে shared আছে তা।
* Common না থাকলে Ans 0 হবে (এই Ans Col এর Attribute ফাঁকা থাকবে!

(*)

A

S_{no}	P_{no}
S_1	P_1
S_1	P_2
S_1	P_3
S_1	P_4
S_2	P_1
S_2	P_2
S_3	P_2
S_4	P_2
S_4	P_4

B_1

P_{no}
P_2

B_2

P_{no}
P_4

A/B_1 =

Sno
S_1
S_2
S_3
S_4

A/B_2 =

Sno.
S_1
S_4

Cartesian Product Operation:

R

A
a
b

S

B
1
2
3

R×S =

A	B
a	1
a	2
a	3
b	1
b	2
b	3

Scanned by CamScanner

* $[2 \times 2 = 4]$

R

A	B
α	1
α	2

S

B	C
1	x
2	4

R×S =

A	B	D	C
α	1	1	x
α	1	2	4
α	2	1	x
α	2	2	4

$\rho_{B \to D(S)}$

* কোনটা কোন table থেকে আসছে তা বোঝার জন্য rename করতে হবে (in case of common name)

25 May 2017

Example:

Borrower

Name	Loan no
KAMAL	L-17
JAMAL	L-23

Loan

Loan no	Branch	Amount
L-14	KAKRAIL	75,000
L-23	MOTIJHEEL	50,000

Relational Algebra

Find the name of the Customers of a Bank (in-example) who have a loan at MOTIJHEEL Branch.

$\rho_{Loan.no \rightarrow borrow.Loan}$ (borrower)

$\rho_{Loan.no \rightarrow Loan.Loan\ no}$ (Loan)

Scanned by CamScanner

(Borrower × Loan) = new_table → optional
(না করাই ভালো)

Name	Borrow. Loan no	Loan. Loan no	Branch	Amount
KAMAL	L-17	L-14	KAKRAIL	75,000
KAMAL	L-17	L-23	MOTIJHEEL	50,000
JAMAL	L-23	L-14	KAKRAIL	75,000
JAMAL	L-23	L-23	MOTIJHEEL	50,000

Borrow. Loan no = Loan. Loan no

Branch = MOTIJHEEL

$$\Pi_{name}\left(\sigma_{borrow.Loan\ no\ =\ Loan.loan\ no\ \wedge\ Branch = MOTIJHEEL}\left(Borrow \times Loan\right)\right)$$

* table গুলো হলে => "Cartesian"
table বড় হলে => "Join"

* Cartesian এ row এর পর row এর B match করবে।

28 May 2017

Natural Join (⋈)

Example:

Common Column

R

Name	Department
Smith ✓	CSE ✓
John ✓	EEE ✓
Paul ✓	EEE ✓

S

Department	Head
✓ EEE	Brown
✓ CSE	Alen
MCE	white

Joining এর সময় কোনো একটি table এর specific কোনো row বাদ গেলে, তাকে "Dangling Tuple" বলে।

R ⋈ S

Bridge

Name	Department	Head
Smith	CSE	Alen
John	EEE	Brown
Paul	EEE	Brown

Scanned by CamScanner

Use the tables of the last class:

Borrower (Name, Loan no)
Loan (Loan No, Branch, Amount)

Find the name of the customer who have a loan at MOTIJHEEL Branch using natural Join.

⇒

Name	Loan No	Branch	Amount
JAMAL	L-23	MOTIJHEEL	75,000

$$\Pi_{name}\left(\sigma_{Branch = MOTIJHEEL}\left(Borrow \bowtie Loan\right)\right)$$

Scanned by CamScanner

Theta Join:

Notation $(R \bowtie S)$

or,

$$\sigma_{condition}(R \bowtie S)$$

Example:

Depositor (CustomerNo, name, $\boxed{\text{accountno}}$)

Account ($\boxed{\text{accountno,}}$ branch, balance)

— Find the names of the customer who have an account in the bank having balance greater than 5000?

$$\pi_{name}\left(\sigma_{Balance > 5000}(Depositor \bowtie Account)\right)$$

or,

$$\pi_{name}(Depositor \bowtie Account)_{Balance > 5000}$$

Scanned by CamScanner

Semi Join (Left side Priority ম্যাচ, resulting table → 3 Half left side এর Column সংখ্যক Column থাকে)

Notation $R \ltimes_F S$, F is the condition

Say,

R

E.No	E. name	Designation
E1	Rupa	Programmer
E2	Nipa	Analyst
E3	Jhuma	DB Admin
E4	Sadat	Consultant

S

Designation	Salary
Programmer	25000
Consultant	50,000
Manager	75000

$R \ltimes S$

R designation = S designation

E.No.	E. Name	Designation
E1	Rupa	Programmer
E3	Sadat	Consultant

Outer Join

(Left এলাকার table full থাকবে, Right এর value match করলে হবে, otherwise NULL হবে)

1) Left outer Join (⟕)
2) Right Outer Join (⟖)
3) Full Outer Join (⟗)

i) Left Outer Join

$R_1 ⟕ R_2$

E. Name	Department	Head
Smith	Sales	NULL
Black	Production	Mary
white	Production	Mary

ii) Right Outer Join

$R1 ⟖ R2$

E. Name	Department	Head
Black	Production	Mary
NULL	Purchasing	Brown
White	Preoduction	Mary

Example

R1

E. Name	Department
Smith	Sales
Black	Production
white	Production

R2

Department	Head
Production	Mary
Purchasing	Brown

Scanned by CamScanner

29 May 2017

3) Full Outer Join (Left Outer + Right Outer)

Name	Department	Head
Smith	Sales	NULL
Black	Production	Mary
White	Production	Mary
NULL	Purchase	Brown

4 Aggregation (मतलब एक Operation करना) :

Aggregate function take a collection of values and return a single value as a result.

Column/Attribute Value (Same Domain).

Notation

$$ {}_{G}\mathcal{G}_{F(A)}(R) $$

$\mathcal{G} \rightarrow$ Aggregate Symbol

$F \rightarrow$ functions

$A \rightarrow$ Attribute

$R \rightarrow$ Relation

$G \rightarrow$ group By

Functions can be

- max
- min
- sum
- count
- count_distinct
- average

etc.

Example

Instruction

ID	Name	Department	Salary
01	Alen	CSE	20,000
21	Brown	EEE	25,000
51	Cook	EEE	35,000
22	Daurson	ME	10,000
43	Enly	CE	15,000
56	Frank	CSE	40,000
65	Givson	ME	30,000

Scanned by CamScanner

- Find the total salary of the Instructor?

$\Rightarrow \mathcal{G}_{sum(salary)}$ (Instruction)

- How many instruction are there in Instruction Relation?

$\Rightarrow \mathcal{G}_{count\text{-}distinct(ID)}$ (Instructor)

- How many departments are there in Instruction Relation?

$\Rightarrow \mathcal{G}_{count\text{-}distinct(department)}$ (Instructor)

- Find the average salary of the instructors according to the department?

$\Rightarrow \mathcal{G}_{average(salary)}$ (Instructor)

Department → Group By (as footer-back)

Scanned by CamScanner

ID	Name	Department	Salary
43	Enly	CE	15,000
01	Alen	CSE	20,000
56	Frank	CSE	40,000
21	Brown	EEE	25,000
51	Cook	EEE	35,000
22	Dawson	ME	10,000
65	Gierson	ME	30,000

→

Department	Average Salary
CE	15,000
CSE	30,000
EEE	30,000
ME	20,000

(Group)

01 June 2017

CHAPTER: NORMALIZATION

Functional Dependency: Attribute of a relation depends on other attributes of the same relation.

- determinant
- objects of determinant

} not vice versa.

Scanned by CamScanner

Example :

Student ID → {First name, last name}

Course number → Course name, Teacher name

Student ID, Course number → marks

marks → grade

Course number → Teacher name

$\underbrace{\text{Course Number, Course Name}}_{\text{Composite determinant (Composite Determinant)}}$ → Teacher name

Partial Functional Dependency :

Attribute, is removed from the composite determinant but the dependency is not impacted.

Course name, Course number → Teacher name

* Course Attribute are "Not Database" "Impacted" By :

Full Functional Dependency: If an attribute is removed from the composite determinant and the dependency is impacted.

Student ID, Course Number $\longrightarrow$ Marks

Transitive Dependency: When attributes are creating a chain of dependency.

$R(A, B, C)$

$A \rightarrow B$, $B \rightarrow C$

$\therefore A \rightarrow C$

Staff

$\overset{A}{\text{Staff no}} \rightarrow \overset{B}{\text{designation}} \rightarrow \overset{C}{\text{Salary}}$

$\overset{A}{\text{Staff no}} \rightarrow \overset{C}{\text{salary}}$

Normal Form (যখন কোনো table এর data inconsistent হয়)

1. 1NF (First Normal Form)
2. 2NF
3. 3NF
4. BCNF (Boyce Codd)
5. 4NF
6. 5NF
7. 6NF

Scanned by CamScanner

Normalization:

Normalization is needed to stop data inconsistency. For 3 reasons problem is created.

- Insert anomaly
- Delete anomaly
- Update anomaly

Person ID	Project Budget	Project	Time Spent by person for project
S75	60 (32)	P1	7
S75	40	P2	3
S79	32	P1	4
S79	27	P3	1
S96	40	P2	5
→ S96	50	P1	7

Scanned by CamScanner

Steps in NORMALIZATION

1) Specify the primary key of the relation.

2) Specify the Functional Dependency of a relation.

3) Apply the definition of each Normal Form (Starting with INF)

4) If a relation fails to meet the definition of Normal Form, change relation until it meets the definition.

5) Re-Test the modified/new relation to ensure they meet the defination of each Normal Form.

Scanned by CamScanner

05 June 2017

First Normal Form (1NF)

A relation (Table) is in 1NF if it is satisfied the following conditions.

→ Contains only atomic (single) values

→ There are no repeating groups

Product

ProductID	Colour	Price
1	red, green	15.99
2	yellow	23.95
3	green	17.50
4	blue, green	9.99
5	red	29.99

Book

Book.id	Author 1	Author 2	Author 3
1	Abraham	John	Paul
2	Cardlon	Smith	Jack
3	Bella	Naomi	John

* Atomic value নেই, So 1NF এ আনতে হবে।
then "Duplicacy" check করতে হবে।
* table তৈরির পর table-name change হয়ে যাবে।

determinant → object of determinant

Product ID → Colour
Product ID → price } break করতে হবে

Product-colour

Product ID	Colour
1	red
1	green
2	yellow
3	green
4	blue
4	green
5	red

Product-Price

Product ID	Price
1	15.99
2	23.95
3	17.50
4	9.99
5	29.99

Book-Author

Book ID	Author
1	Abraham
1	John
1	Paul
2	Candlon
2	Smith
2	Jack
3	Jack
3	Jack
3	Jack

Second Normal Form (2NF)

A relation is in 2NF if it satisfies the following conditions -

→ It is in 1NF

→ All non-key attributes are fully functionally dependent on primary key.

* "Normalization" mainly "Data-Inconsistency" নিয়ে কাজ করে।

Scanned by CamScanner

Purchase Detail

Customer. ID	Store ID	Location
1	1	DHAKA
1	3	KHULNA
2	1	DHAKA
3	2	SYLHET
4	3	KHULNA

* Multiple Value Database ও single Row তে insert না করাই ভালো, Because ওতে NORMALIZATION করা যায় না।

Determinant		Object of Determinant
CustomerID, StoreID	→	Location

CustomerID → Location ✗

StoreID → Location ✓

CustomerID → StoreID ✓

Scanned by CamScanner

Purchase

Customer ID	Store ID
1	1
1	3
2	1
3	2
4	3

Store

Store ID	Location
1	DHAKA
2	Sylhet
3	KHULNA

* একই primary key পরবর্তীতে আবার use করা যায়—
(Duplicacy Possible)

Scanned by CamScanner

06 July 2017

Third Normal Form (3NF) : A relation

is in 3NF if it follows:

- It is in 2NF.
- No transitive dependency exists within the relation.

Transitive

$A \rightarrow B$

$B \rightarrow C$

$\therefore A \rightarrow C$

Book_details

Book_id	Genre_id	Genre Type	Price
1	1	Story	120
2	2	Sports	140
3	1		130
4	3	Travel	350
5	2	Sports	250

Scanned by CamScanner

Functional Dependency

$\underset{A}{\underline{Book_id}} \rightarrow \underset{B}{\underline{Genre_id}}$, price

$\underset{B}{\underline{Genre_id}} \rightarrow \underset{C}{\underline{GenreType}}$

$\therefore A \rightarrow C$

Book → GenreType

Transitive হয়ে যাচ্ছে। So, 3NF হবে না। অতএব, Convert করতে হবে।

Book_price

Book_id	Price	Genre_id
1	120	1
2	140	2
3	130	1
4	350	3
5	250	2

Genre_details

Genre_id	Genre_type
1	Story
2	Sports
3	Travel

Not transitive. So, 3NF.

Scanned by CamScanner

Employee

Emp_id	Emp_name	Emp_zip	Emp_state	Emp_city	Emp_district
1001	John	201212	NR	RAJ	RANGPUR
2001	Lora	201208	SU	DHK	MYMENSING
2019	Steve	192001	SU	DHK	MANIKGANJ
1012	Mark	201219	ES	KHL	SATKHIRA
1051	Smith	202002	WS	CTG	COX'S BAZAR

Q. Normalisation কেন দরকার?

Q. Normalisation এর steps গুলো লিখ।

Q. Partial Functional Dependency & Fully Functional Dependency এর পার্থক্য লিখ।

Q. Composite Determinant কাকে বলে?

Q. Normalization এর কয়টা form আছে?

Q. কোন Normal এ আছে? কোনটায় নাই? কেন নাই? Normal কর। (7 Marks)

কী Rule মানলে হবে Normal.