Data Models

Dr. Sambit Bakshi

NIT Rourkela

February 18, 2022

Outline

Data Models

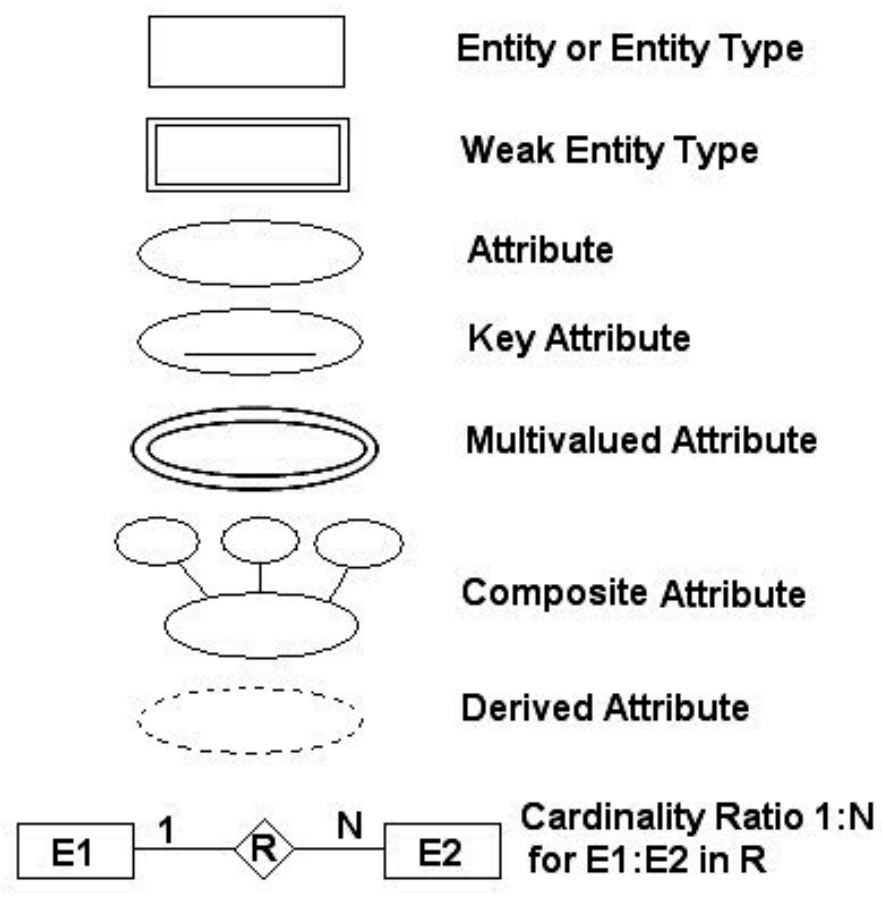
A Collection of concepts that can be used to describe the structure of a database (datatypes, relationships, and constraints that the data should adhere to)

- **High Level Data Models:** Entity-Relationship (E-R) Model, Unified Modeling Language (UML)
- **Representational Data Models:** Network, Hierarchical, Relational [Edgar F. Codd (1970)]
- **Physical Data Models:** Data structures

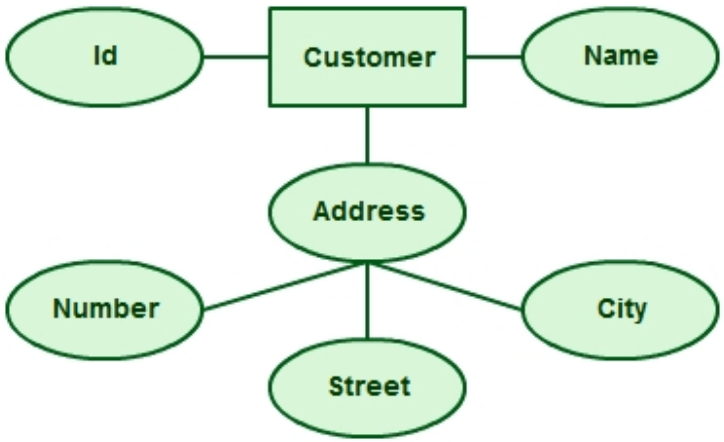
ER Model

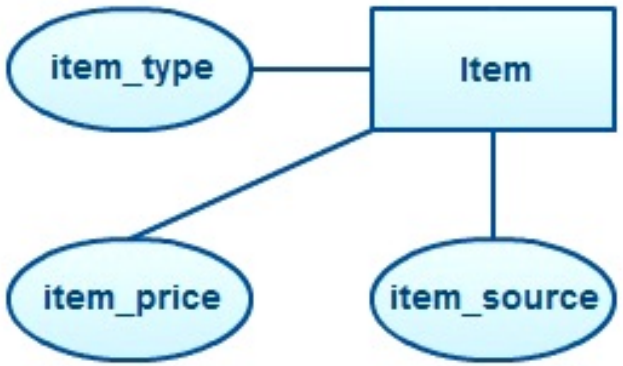
- **E-R Model:** The entity-relationship (E-R) data model is useful in mapping the meanings and interactions of real-world enterprises on to a conceptual schema.
- **Entity:** An entity is a 'thing' or 'object' in the real world that is distinguishable from all other objects.
- An entity is represented by a set of **attributes**.
- **Entity Set:** An entity set is a set of entities of the same type that share the same properties, or attributes.
- **Relationship:** A relationship is an association among several entities.
- **Degree of a relationship:** Number of entities participating in the relationship.
  If the association is among n entities, we call it n-ary relationship.
  If the association is among two entities, we call it binary relationship.

Data Modelling Using ER model

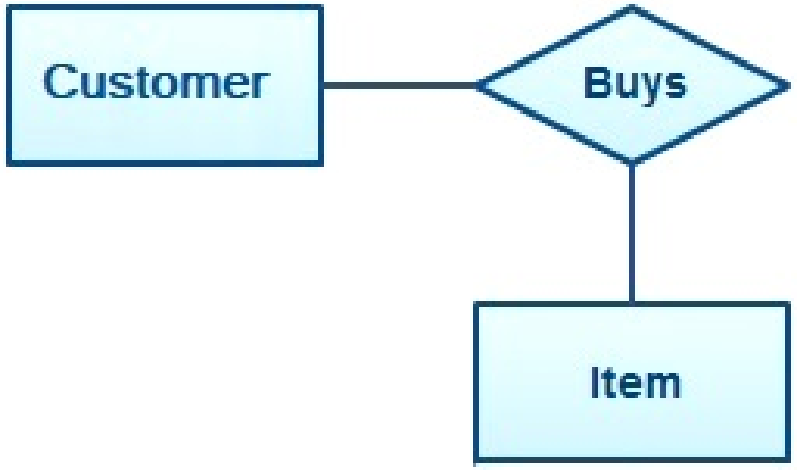


Entity Sets / Entity Types





Relationship Sets



Attributes

- **Attribute:**

  For each attribute, there is a set of permitted values, called the domain, or value set, of that attribute

  Types of Attributes:

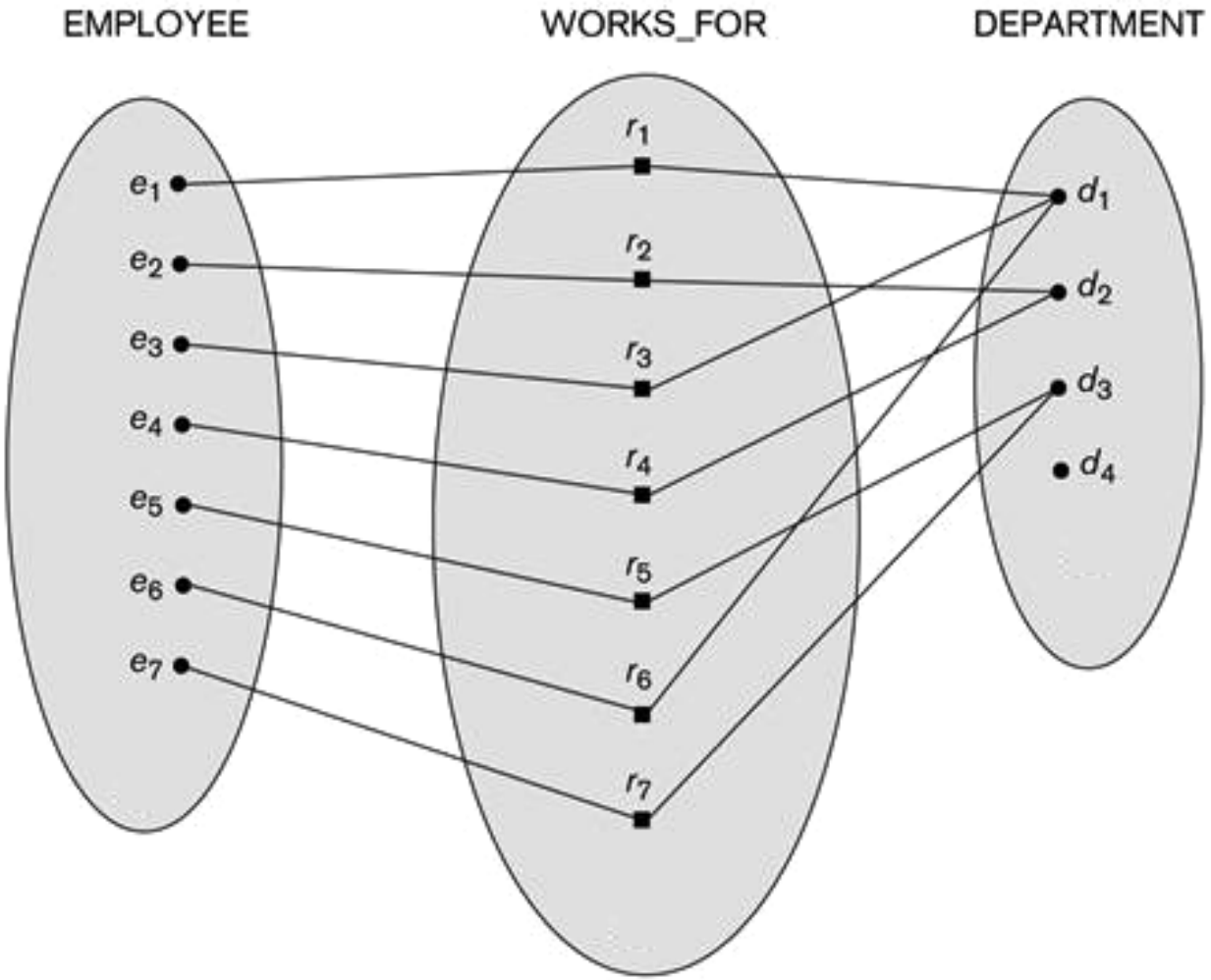
  - Simple and Composite Attribute
  - Single-valued and Multi-valued Attribute
  - Derived Attribute
  - Complex (Multi-valued as well as Composite)

Structural Constraints

- **Cardinality Constraints / Maximum Participation / Cardinality ratio**:
  Maximum number an entity in entity set E is allowed to participate in a relationship set R
  One-to-One, One-to-many OR Many-to-one, Many-to-many
- **Modality / Participation Constraints / Minimum Participation / Existence Dependency Constraint**:

  can be Zero for Optional, One or more for Mandatory
  - An entity set E participates in **total** with a relationship set R if every entity in E participates in at least one relationship in R.
  - An entity set E participates in **partial** with a relationship set R if some entity in E participates in at least one relationship in R.

Structural Constraints



Organizational Rule:

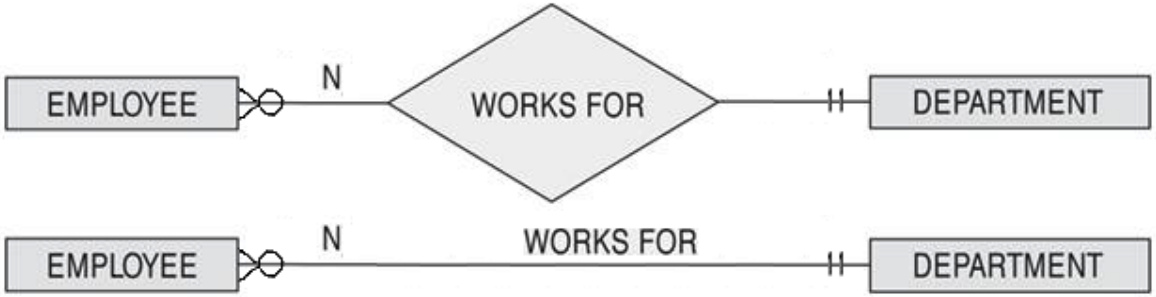
- Each employee works for a particular department
- Many employees can work in a single department
- There can be departments where no employees work

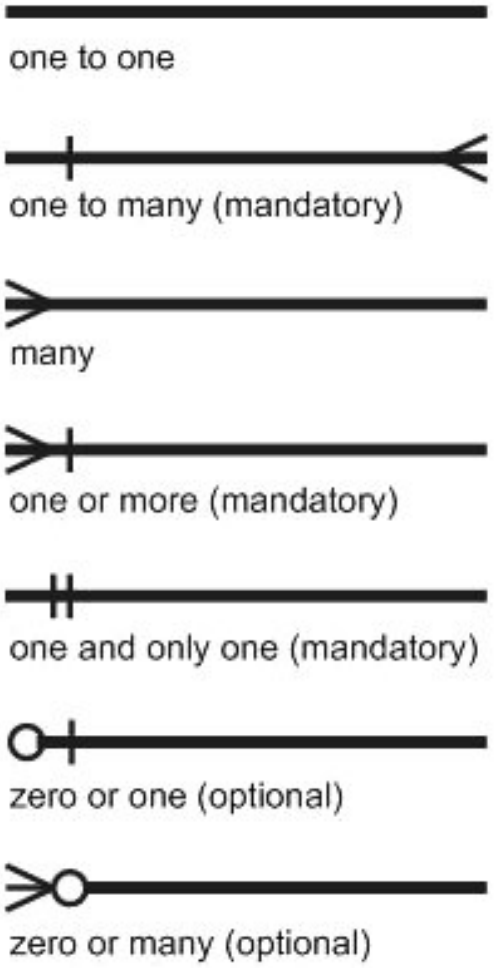
Participation of EMPLOYEE in WORKS_FOR: Max : 1 , Min : 1
Participation of DEPARTMENT in WORKS_FOR: Max : N , Min : 0

Structural Constraints

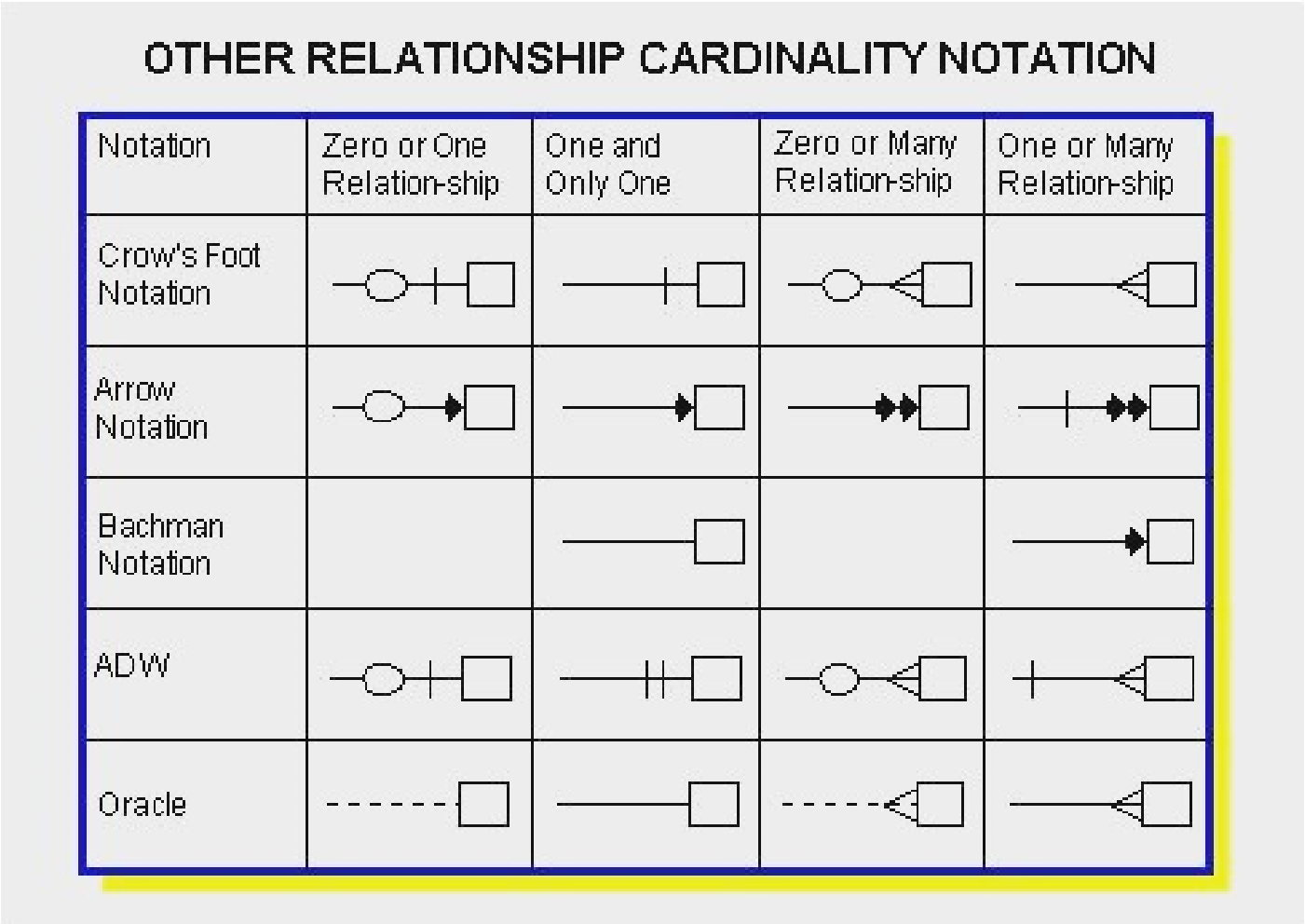
Cardinality Ratio representation: cardinality of two entities written in distance from them (looks like swapped)!

Min-Max Representation: Provides more information! Possible to draw for *n*-ary relationship.



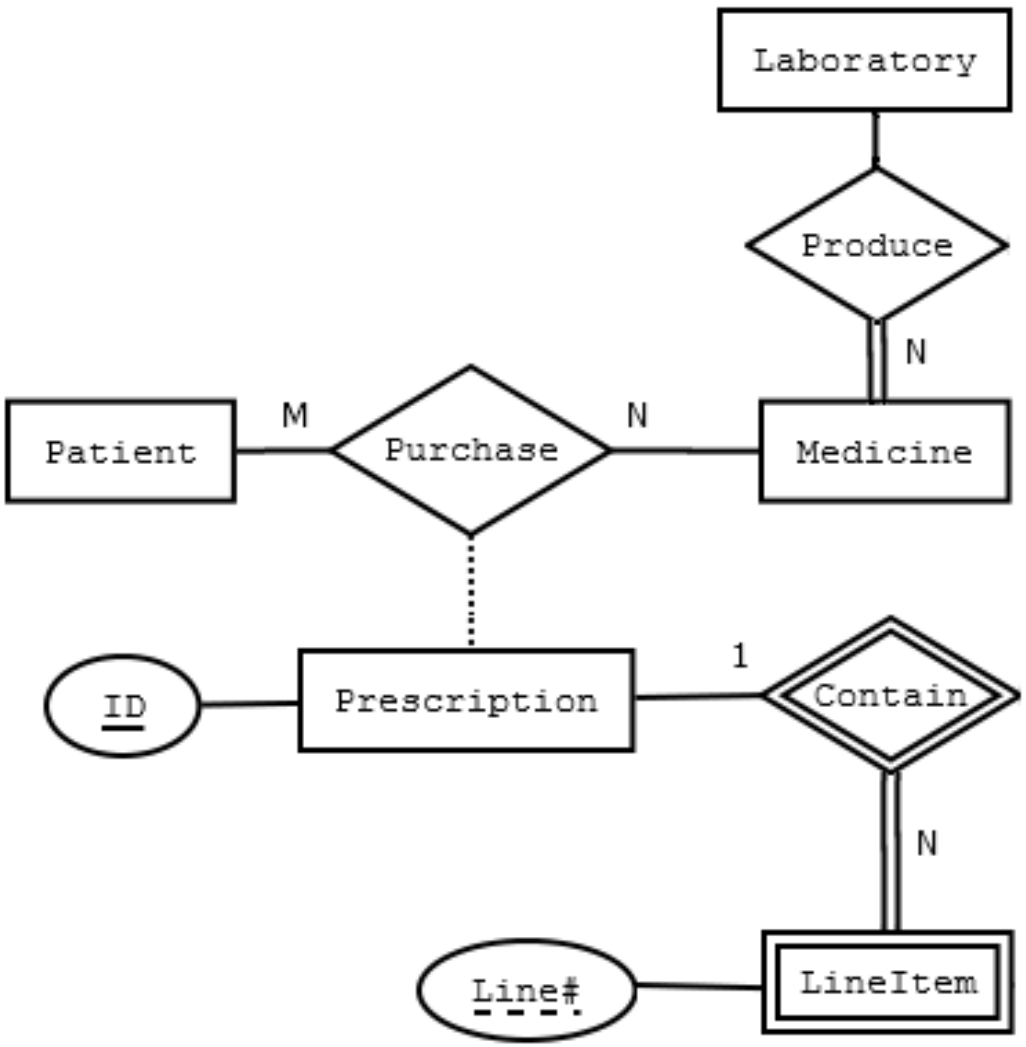


Structural Constraints



We won't learn these!

Structural Constraints



- Medicine participates totally in the Produce relationship, meaning that a medicine can't exist unless Produced by a laboratory. In contrast, a laboratory can exist without producing any medicine - Laboratory participates partially in the Produce relationsip.
- a Patient can Purchase Medicine with or without a Prescription. A Purchase can't exist without a Patient and Medicine, but a Prescription is optional in Purchase relationship.
- a Prescription contains LineItems which are identified by the Prescription's key (ID)and a line number (Line#) together, and not by Line# only. In other words, the LineItems will have a composite key (ID, Line#).

Types of Entity Sets

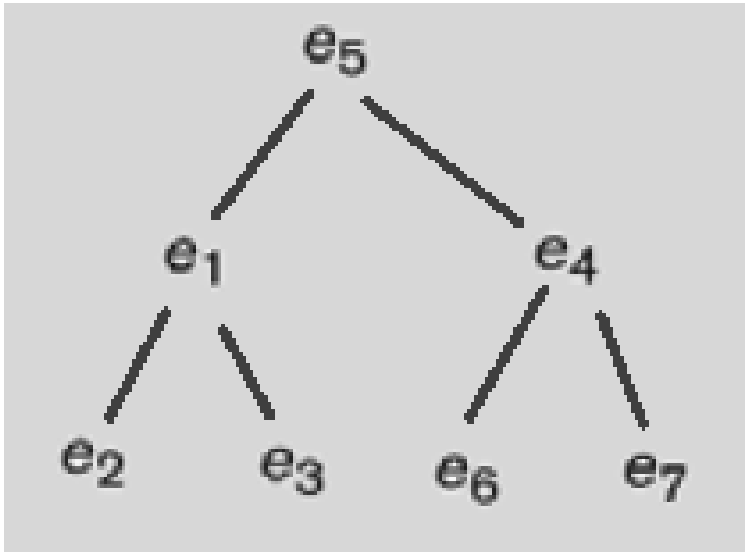
- **Strong Entity Set:** An entity set that has a primary key is termed a strong entity set.
- **Weak Entity Set:** An entity set that does not have sufficient attribute to form a primary key is termed a weak entity set.

- Every weak entity must be associated with an identifying entity, i.e., said to be existence dependent on the identifying entity set.
- The relationship associating the weak entity set with the identifying entity set is called the **identifying relationship**.

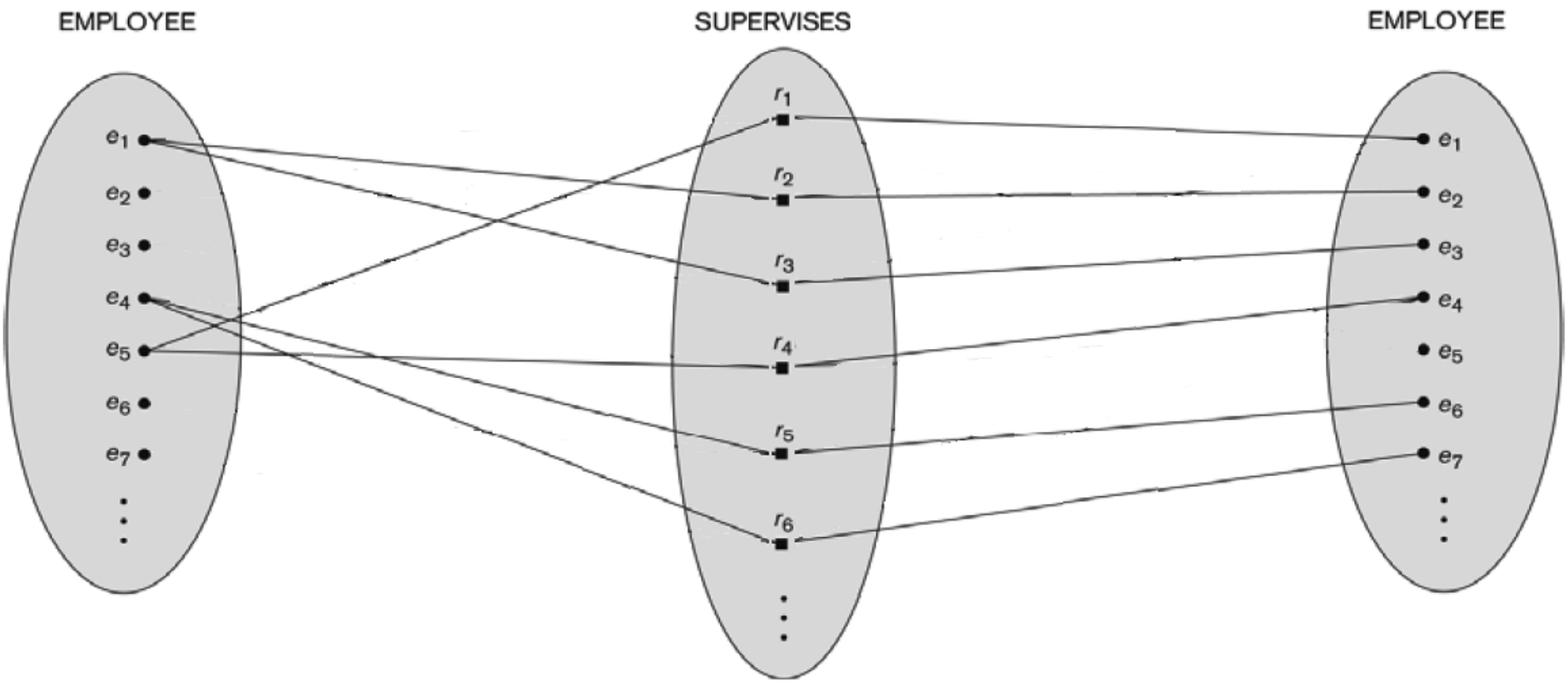
Entity Sets

Q. Which one of the following is used to represent the supporting many-one relationships of a weak entity set in an entity-relationship diagram?
(a) Diamonds with double/bold border
(b) Rectangles with double/bold border
(c) Ovals with double/bold border
(d) Ovals that contain underlined identifiers
[GATE 2020]
ANSWER: (a)

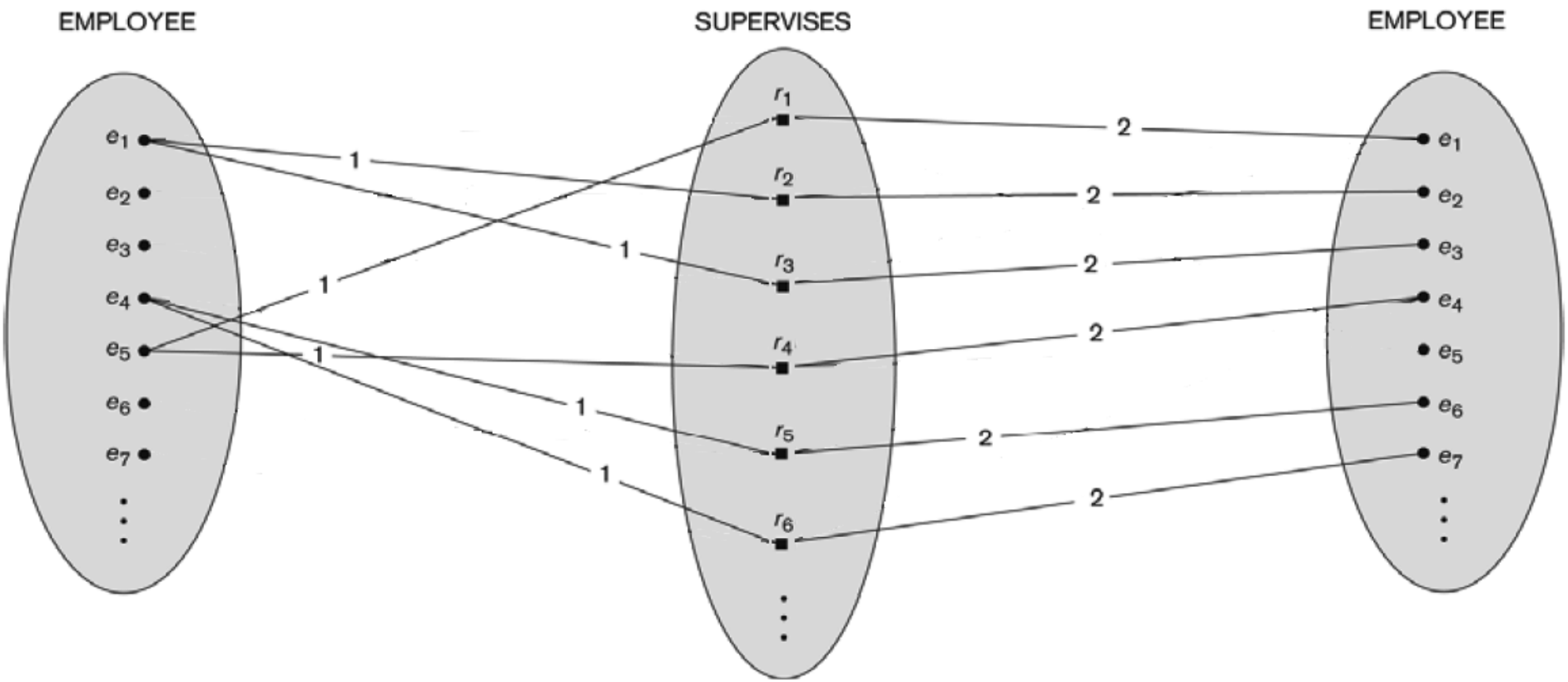
Recursive Relationship



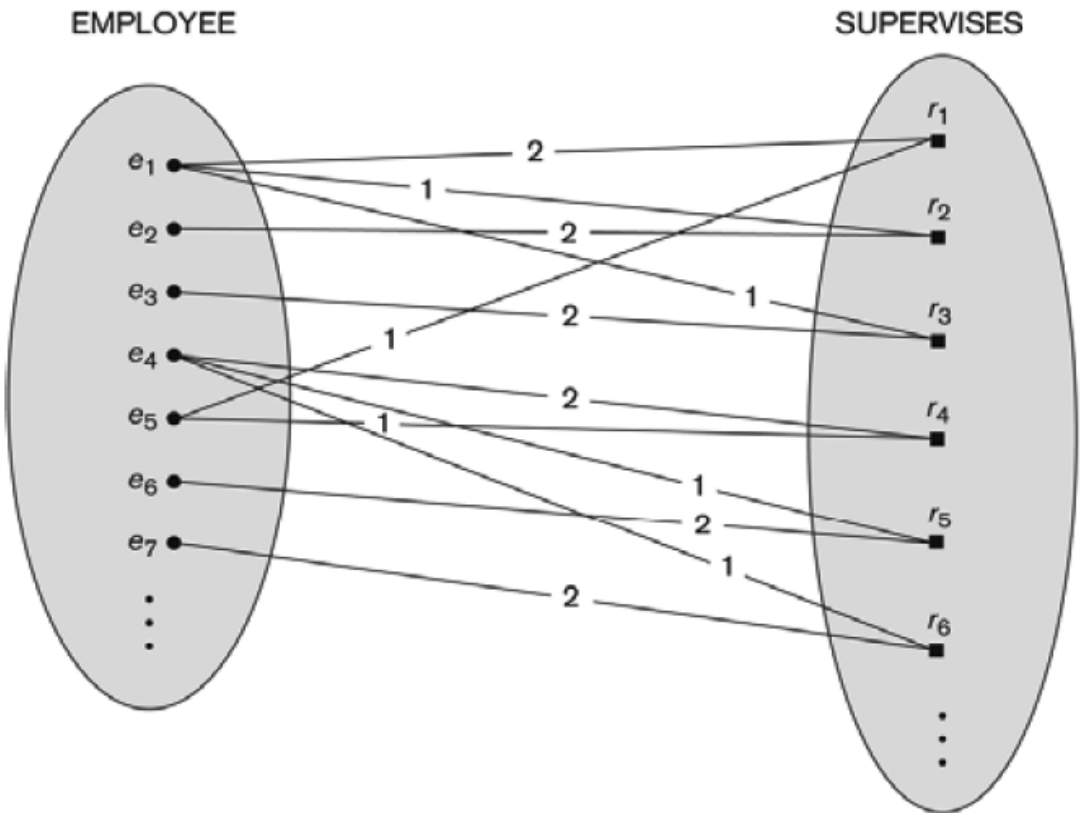
Recursive Relationship

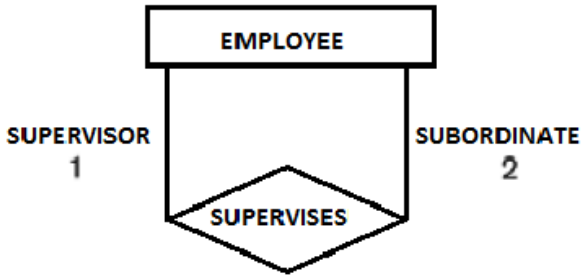


Recursive Relationship



Recursive Relationship

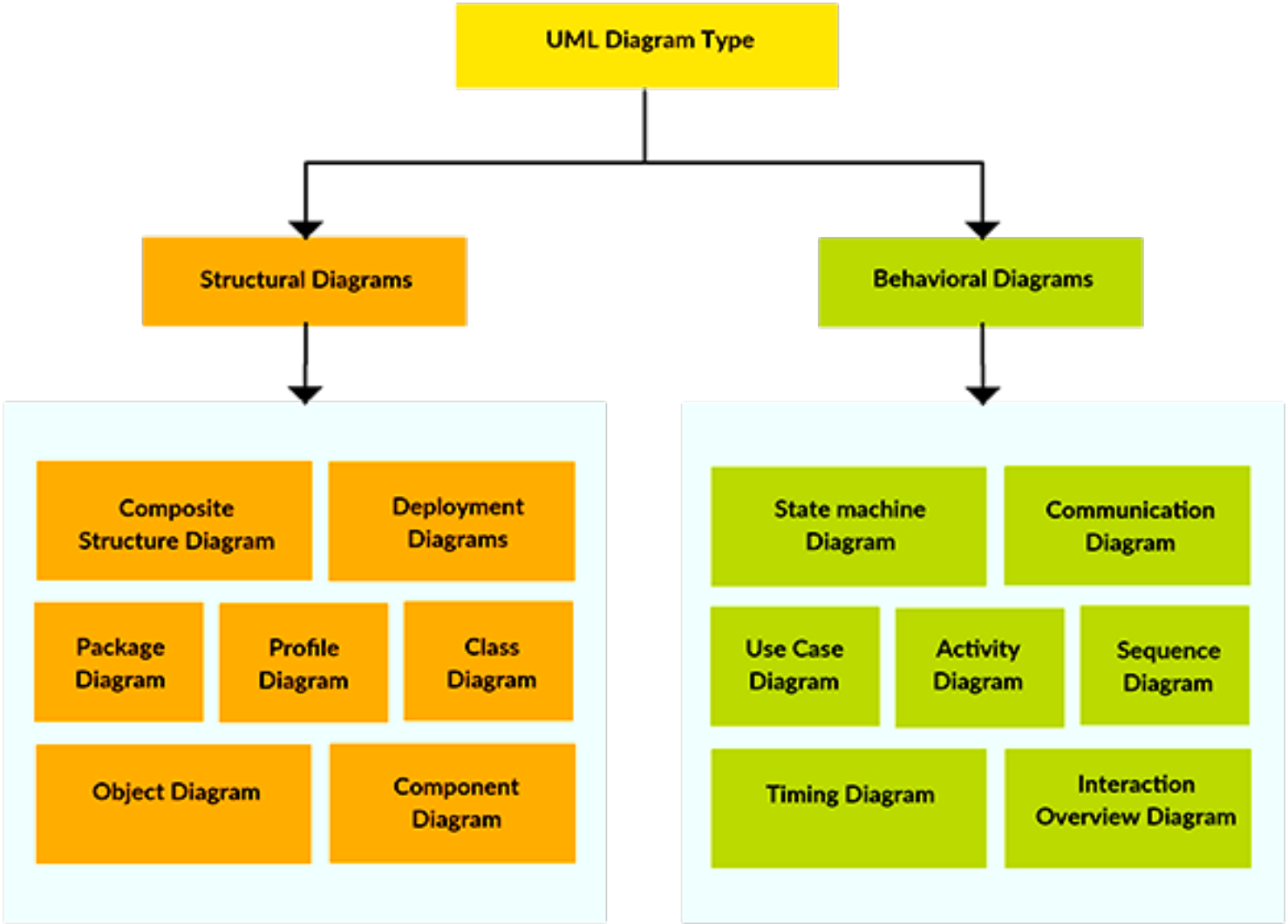




Extending E-R Features

- **Specialization** is the process of classifying a class of objects into more specialized subclasses.
- **Generalization** is the process of generalizing several classes into a higher-level abstract class that includes the objects in all these classes.
- **Specialization** is conceptual refinement, whereas **generalization** is conceptual synthesis.
- **Aggregation** is an abstraction concept for building composite objects from their component objects.
- **Association** is an abstraction concept used to associate objects from several independent classes.

Types of UML Diagram



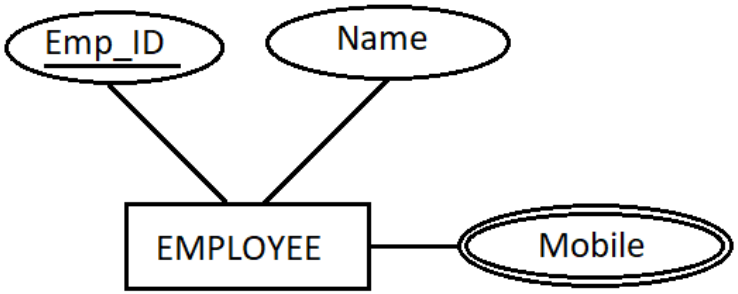
ER Diagram to Relational Model

ER Model	**Relational Model**
Entity Type	Entity Relation
1:1 or 1:N relationship type	foreign key (or relation)
M:N relationship type	a relation with 2 foreign keys
n-ary relationship type	a relation with *n* foreign keys
simple attribute	attribute in relation
composite attribute	set of attributes in relation
multivalued attribute	relation and foreign key
value set	domain
key attributes	keys / candidate keys / primary keys

Properties of Relational Model

- There are no duplicate tuples.
- Tuples are unordered.
- Attributes are unordered.
- All attribute values are atomic.

ER Diagram to Relational Model with example

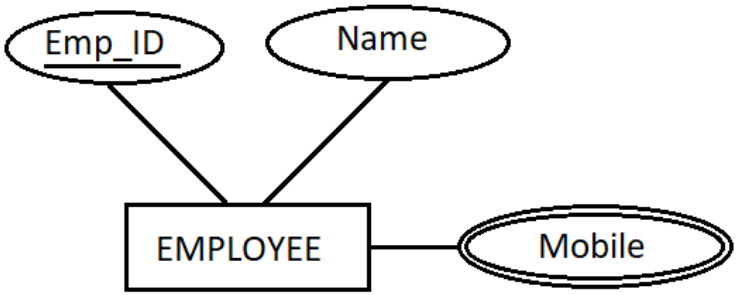


EMPLOYEE

Emp_ID	Name	Mobile
123	De Yu	2032
123	De Yu	2033
134	Min Jiang	2066
134	Min Jiang	2067
134	Min Jiang	2068
135	Bin Zhu	NULL

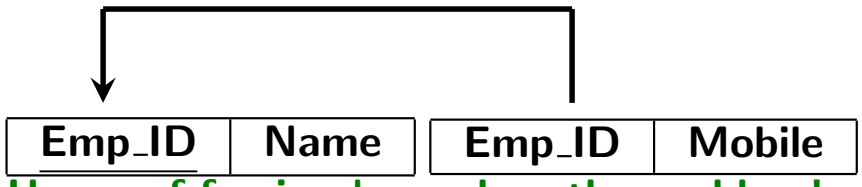
Problem with multi-valued attribute!

ER Diagram to Relational Model with example



EMPLOYEE

Emp_ID	Name	Mobile
123	De Yu	2032
123	De Yu	2033
134	Min Jiang	2066
134	Min Jiang	2067
134	Min Jiang	2068
135	Bin Zhu	NULL



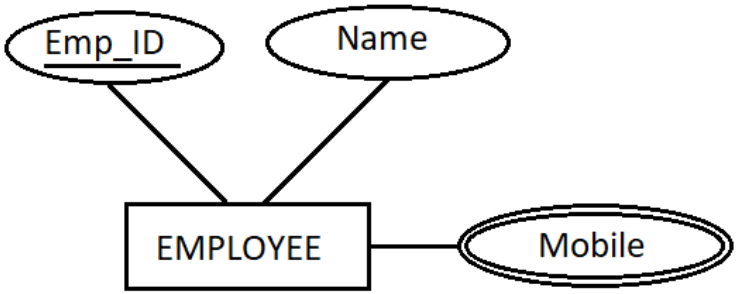
Usage of foreign key solves the problem!

Emp_ID	Name
123	De Yu
134	Min Jiang
135	Bin Zhu

Emp_ID	Mobile
123	2032
123	2033
134	2066
134	2067
134	2068
~~136~~	~~2099~~

But number of table increases!

ER Diagram to Relational Model with example



EMPLOYEE

Emp_ID	Name	Mobile
123	De Yu	2032
123	De Yu	2033
134	Min Jiang	2066
134	Min Jiang	2067
134	Min Jiang	2068
135	Bin Zhu	NULL

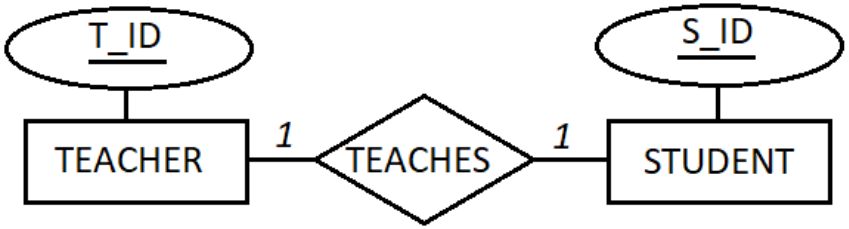
Emp_ID	Name	Mobile_1	Mobile_2
123	De Yu	2032	2033
134	Min Jiang	2066	2067
135	Bin Zhu	NULL	NULL

An organization may not care for a third mobile number.

We loose one phone number (2068) of Min Jiang, but number of table remains same!

This might be a suboptimal solution, but not an equivalent one.

ER Diagram to Relational Model with example



TEACHER

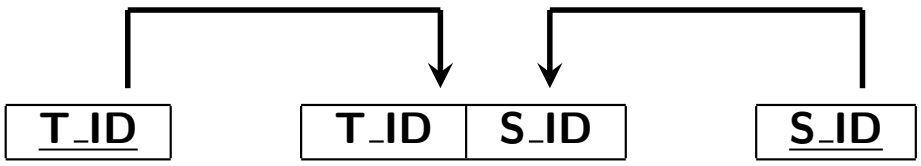
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
~~T_7~~	~~S_1~~
T_4	NULL
NULL	S_2

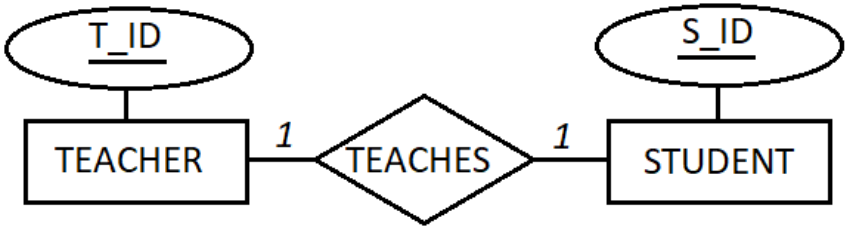
STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



Can these three tables be reduced?

ER Diagram to Relational Model with example



TEACHER

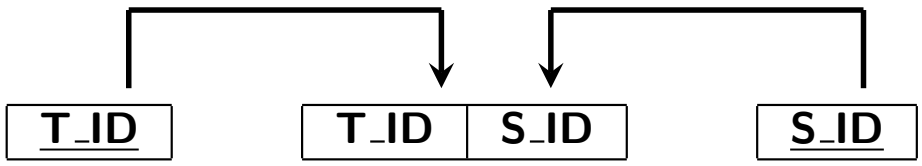
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
~~T_7~~	~~S_1~~
T_4	NULL
NULL	S_2

STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



TEACHER-TEACHES

T_ID	S_ID
T_1	S_7
T_2	NULL
T_3	S_5
T_4	NULL
T_5	NULL
T_6	S_1
T_7	NULL

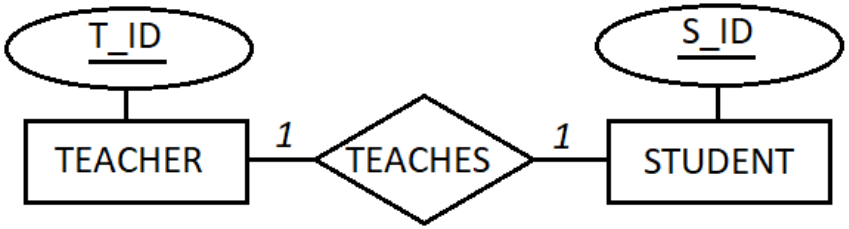
STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8

Option 1: Pushing the information of the relationship to the TEACHER side

Usage of foreign key solves the problem!

ER Diagram to Relational Model with example



TEACHER

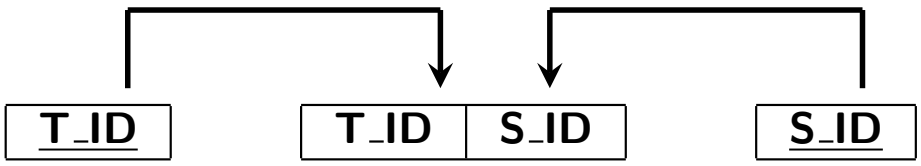
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
~~T_7~~	~~S_1~~
T_4	NULL
NULL	S_2

STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



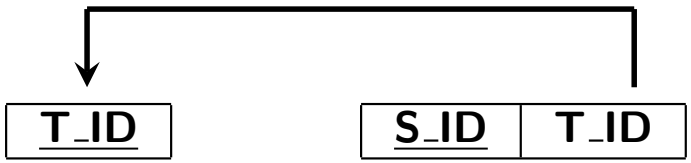
TEACHER

T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

STUDENT-TAUGHTBY

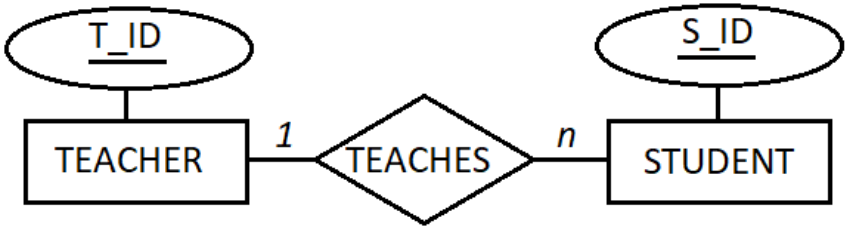
S_ID	T_ID
S_1	T_6
S_2	NULL
S_3	NULL
S_4	NULL
S_5	T_3
S_6	NULL
S_7	T_1
S_8	NULL

Option 2: Pushing the information of the relationship to the STUDENT side



Usage of foreign key solves the problem!

ER Diagram to Relational Model with example



TEACHER

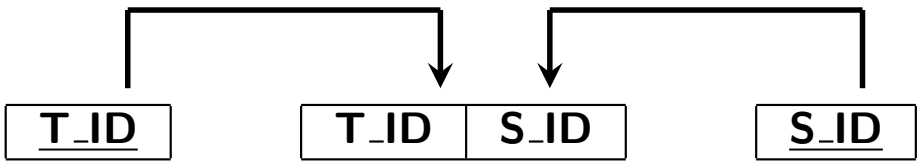
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
T_1	S_3
~~T_7~~	~~S_1~~
T_4	NULL
NULL	S_2

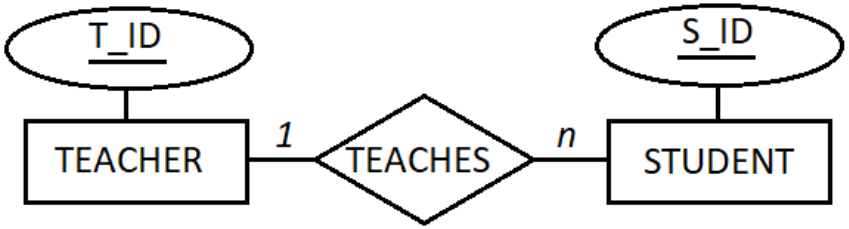
STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



Can these three tables be reduced?

ER Diagram to Relational Model with example



TEACHER

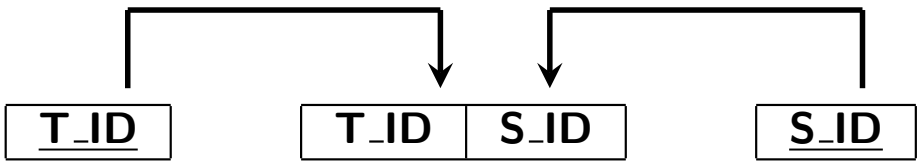
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
T_1	S_3
~~T_7~~	~~S_1~~
T_4	NULL
NULL	S_2

STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



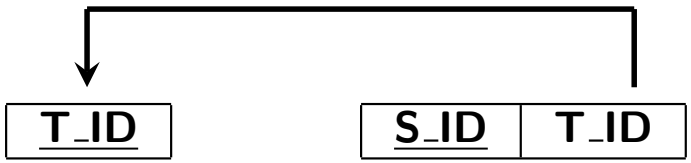
TEACHER

T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

STUDENT-TAUGHTBY

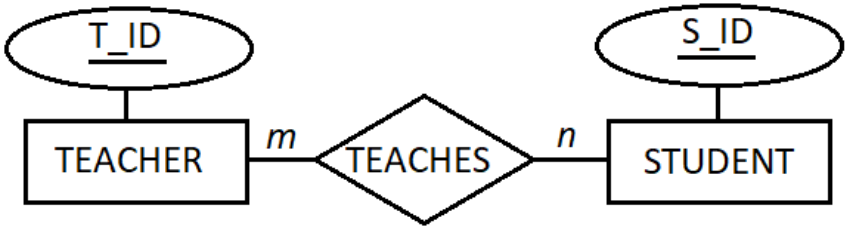
S_ID	T_ID
S_1	T_6
S_2	NULL
S_3	T_1
S_4	NULL
S_5	T_3
S_6	NULL
S_7	T_1
S_8	NULL

Only option: Pushing the information of the relationship to the STUDENT side



Usage of foreign key solves the problem!

ER Diagram to Relational Model with example



TEACHER

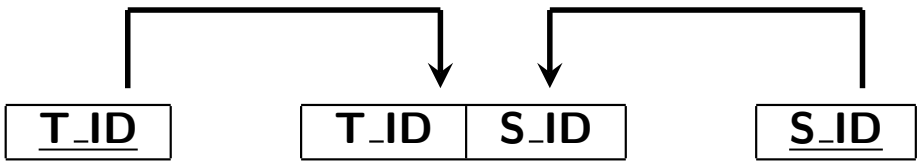
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
T_1	S_3
T_7	S_1
T_6	S_5

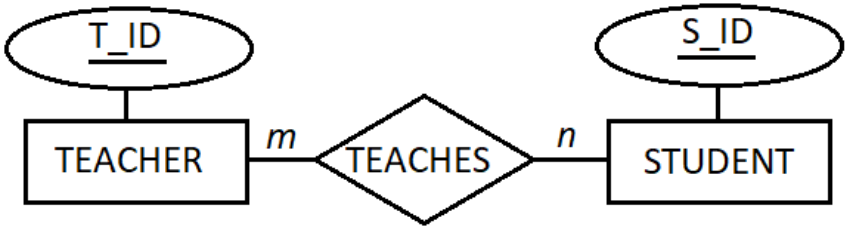
STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



Can these three tables be reduced?

ER Diagram to Relational Model with example



TEACHER

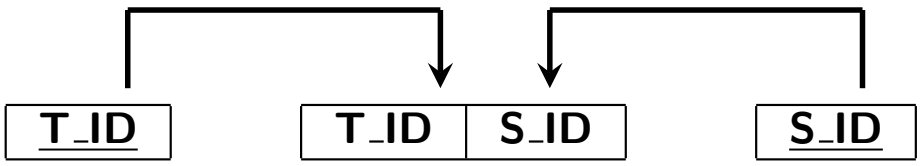
T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

TEACHES

T_ID	S_ID
T_1	S_7
T_3	S_5
T_6	S_1
T_1	S_3
T_7	S_1
T_6	S_5

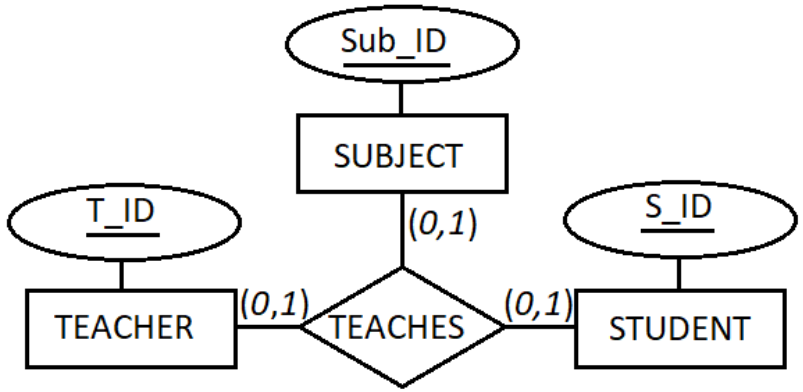
STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8



No option to reduce these three tables to a lesser number.
Two foreign keys are used.

ER Diagram to Relational Model with example



TEACHER

T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8

SUBJECT

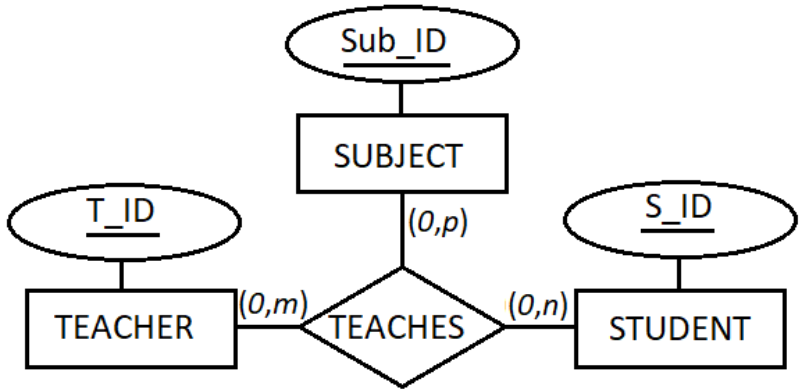
Sub_ID
CS_1
CS_2
CS_3
CS_4
CS_5

TEACHES

T_ID	S_ID	Sub_ID
T_1	S_7	CS1
T_3	S_5	CS2
T_6	S_1	CS3

To reduce these four tables to a lesser number, push the relationship to any entity. Two foreign keys are used.

ER Diagram to Relational Model with example



TEACHER

T_ID
T_1
T_2
T_3
T_4
T_5
T_6
T_7

STUDENT

S_ID
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8

SUBJECT

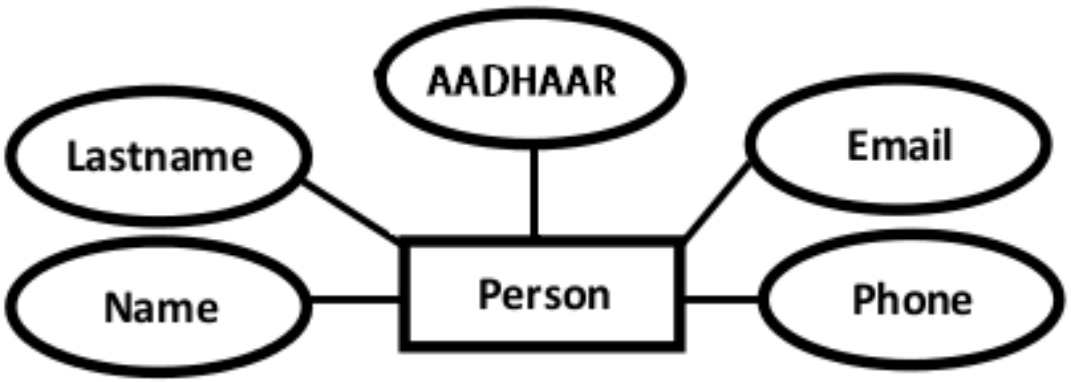
Sub_ID
CS_1
CS_2
CS_3
CS_4
CS_5

TEACHES

T_ID	S_ID	Sub_ID
T_1	S_7	CS1
T_3	S_5	CS2
T_6	S_1	CS3
T_1	S_6	CS1
T_4	S_7	CS1
T_3	S_1	CS4

Cannot reduce these four tables to a lesser number.
Three foreign keys are used.

ER Diagram to Relational Model



1. Verify the type of attributes and select only simple ones
2. Identify multivalued and complex attributes

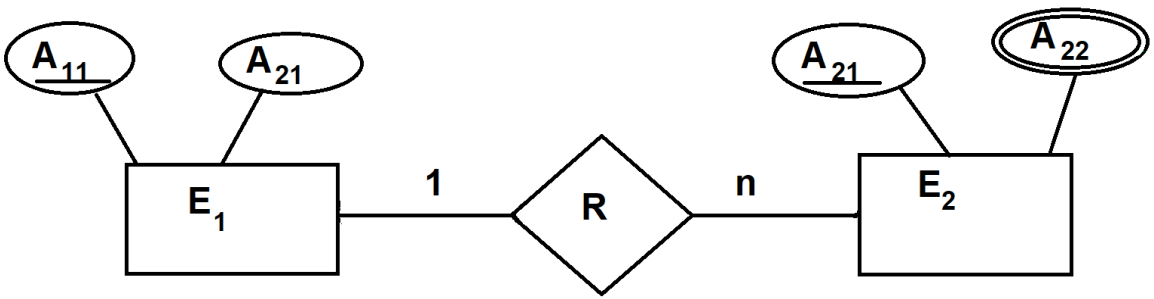
ER Diagram to Relational Model

Q. Given the basic ER and relational models, which of the following is INCORRECT?
(a) An attribute of an entity can have more than one value
(b) An attribute of an entity can be composite
(c) In a row of a relational table, an attribute can have more than one value
(d) In a row of a relational table, an attribute can have exactly one value or a NULL value
[GATE 2012]

ER Diagram to Relational Model

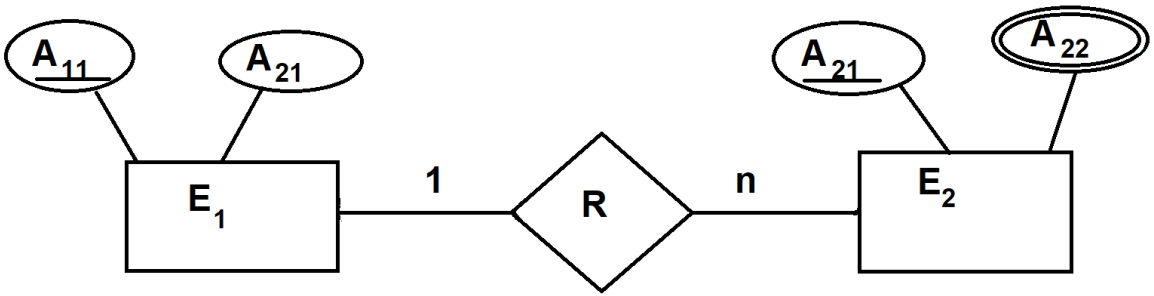
Q. Given the basic ER and relational models, which of the following is INCORRECT?
(a) An attribute of an entity can have more than one value
(b) An attribute of an entity can be composite
(c) In a row of a relational table, an attribute can have more than one value
(d) In a row of a relational table, an attribute can have exactly one value or a NULL value
[GATE 2012]
ANSWER (c)

ER Diagram to Relational Model



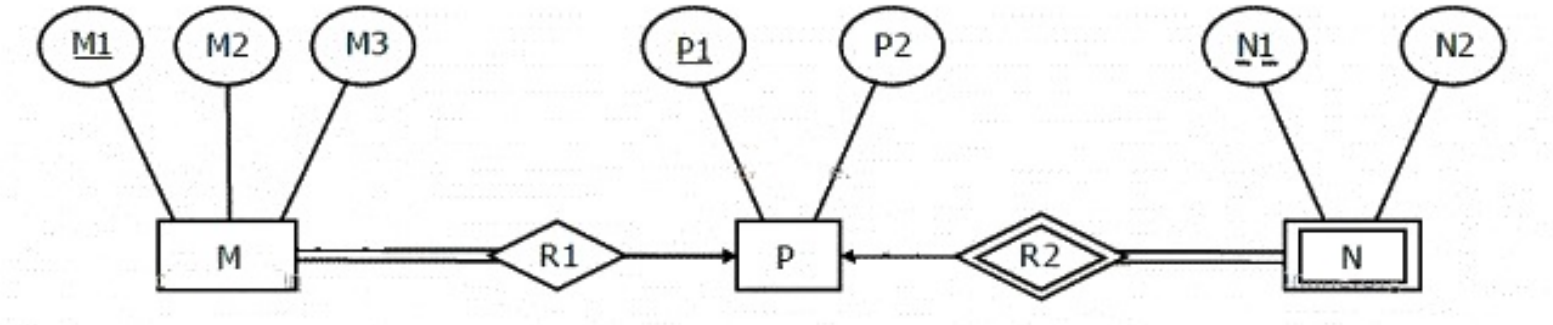
Q1. Maximum how many relations (tables) would be required when this ER diagram will be converted to relational model?
Q2. Minimum how many relations (tables) would be required when this ER diagram will be converted to relational model?
[GATE 2004]

ER Diagram to Relational Model



Q1. Maximum how many relations (tables) would be required when this ER diagram will be converted to relational model?
Q2. Minimum how many relations (tables) would be required when this ER diagram will be converted to relational model?
[GATE 2004]
ANSWER to Q1: 3
ANSWER to Q2: 2

ER Diagram to Relational Model

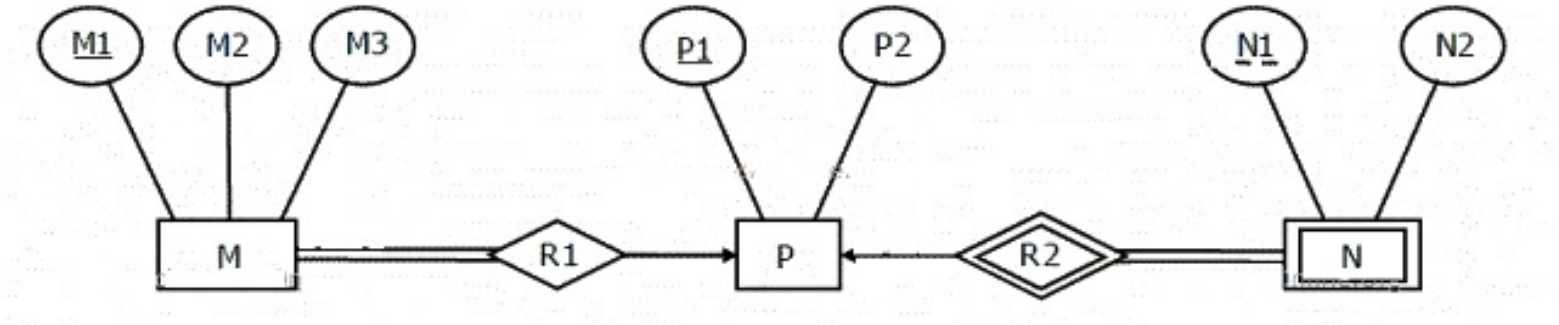


Q1. Minimum how many relations (tables) would be required when this ER diagram will be converted to relational model?
Q2. Which of the following is a correct attribute set for one of the tables for the correct answer to the above question?

(a) M1,M2,M3,P1 (b) M1, P1, N1, N2
(c) M1, P1, N1 (d) M1, P1

[GATE 2008]

ER Diagram to Relational Model



Q1. Minimum how many relations (tables) would be required when this ER diagram will be converted to relational model?
Q2. Which of the following is a correct attribute set for one of the tables for the correct answer to the above question?

(a) M1,M2,M3,P1 (b) M1, P1, N1, N2
(c) M1, P1, N1 (d) M1, P1

[GATE 2008]
ANSWER Q1. 2
ANSWER Q2. (a) M1,M2,M3,P1

ER Diagram to Relational Model

Q. E_1 and E_2 are two entities (with only simple attributes) having R_1 and R_2 as two relationships between them.
R_1 is a one-to-many and R_2 is a many-to-many relationship. R_1 and R_2 do not have any attribute of their own.
What is the minimum number of tables required to convert this ER model to relational model? [GATE 2005]

ER Diagram to Relational Model

Q. E_1 and E_2 are two entities (with only simple attributes) having R_1 and R_2 as two relationships between them.
R_1 is a one-to-many and R_2 is a many-to-many relationship. R_1 and R_2 do not have any attribute of their own.
What is the minimum number of tables required to convert this ER model to relational model?
[GATE 2005]
ANSWER: 3

ER Diagram to Relational Model

A	C
2	4
3	4
4	3
5	2
7	2
9	5
6	4

Q. The table has two attributes A and C where A is the primary key and C is a foreign key referencing A with on-delete cascade.
The set of all tuples that must be additionally deleted to preserve referential integrity when the tuple $(2, 4)$ is deleted is:
(a) (3,4),(6,4)
(b) (5,2),(7,2)
(c) (5,2),(7,2),(9,5)
(d) (3,4),(4,3),(6,4)
[GATE 2005]

ER Diagram to Relational Model

A	C
2	4
3	4
4	3
5	2
7	2
9	5
6	4

Q. The table has two attributes A and C where A is the primary key and C is a foreign key referencing A with on-delete cascade.
The set of all tuples that must be additionally deleted to preserve referential integrity when the tuple $(2, 4)$ is deleted is:
(a) (3,4),(6,4)
(b) (5,2),(7,2)
(c) (5,2),(7,2),(9,5)
(d) (3,4),(4,3),(6,4)
[GATE 2005]
ANSWER: (c)

ER Diagram to Relational Model

Table T1

P	Q
2	2
3	8
7	3
5	8
6	9
8	5
9	8

Table T2

R	S
2	2
8	3
3	2
9	7
5	7
7	2

Q. In table T1, P is the primary key, Q is the foreign key referencing R in table T2 with on-delete cascade. In table T2, R is the primary key and S is the foreign key referencing P in the table T1 with on-delete cascade. In order to delete record $(3, 8)$ from table T1, numbers of additional record that need to be deleted from table T1 is

(a) 0
(b) 1
(c) 2
(d) 3

ER Diagram to Relational Model

Table T1

P	Q
2	2
3	8
7	3
5	8
6	9
8	5
9	8

Table T2

R	S
2	2
8	3
3	2
9	7
5	7
7	2

Q. In table T1, P is the primary key, Q is the foreign key referencing R in table T2 with on-delete cascade. In table T2, R is the primary key and S is the foreign key referencing P in the table T1 with on-delete cascade. In order to delete record $(3, 8)$ from table T1, numbers of additional record that need to be deleted from table T1 is
(a) 0
(b) 1
(c) 2
(d) 3
ANSWER: (c)

ER Diagram to Relational Model

Table T1

P	Q
2	2
3	8
7	3
5	8
6	9
8	5
9	8

Table T2

R	S
2	2
8	3
3	2
9	7
5	7
7	2

Q. In table T1, P is the primary key, Q is the foreign key referencing R in table T2 with on-delete cascade and on-update cascade. In table T2, R is the primary key and S is the foreign key referencing P in the table T1 with on-delete set NULL and on-update cascade. In order to delete record $(3, 8)$ from table T1, numbers of additional record that need to be deleted from table T1 is

(a) 0
(b) 1
(c) 2
(d) 3

[GATE2017]

ER Diagram to Relational Model

Table T1

P	Q
2	2
3	8
7	3
5	8
6	9
8	5
9	8

Table T2

R	S
2	2
8	3
3	2
9	7
5	7
7	2

Q. In table T1, P is the primary key, Q is the foreign key referencing R in table T2 with on-delete cascade and on-update cascade. In table T2, R is the primary key and S is the foreign key referencing P in the table T1 with on-delete set NULL and on-update cascade. In order to delete record $(3, 8)$ from table T1, numbers of additional record that need to be deleted from table T1 is
(a) 0
(b) 1
(c) 2
(d) 3
[GATE2017]
ANSWER: (a)

ER Diagram to Relational Model

Q. Let R(a,b,c) and S(d,e,f) be two relations. 'd' is the foreign key of S that refers to primary key of R. Consider the following four operations on S and R:
(1) Insert into R
(2) Insert into S
(3) Delete from R
(4) Delete from S
Which of these operations may cause violation of referential integrity constraint?
(A) None of (1), (2), (3) or (4) can cause its violation
(B) All of (1), (2), (3) and (4) can cause its violation
(C) Both (1) and (4) can cause its violation
(D) Both (2) and (3) can cause its violation
[GATE 1997]

ER Diagram to Relational Model

Q. Let R(a,b,c) and S(d,e,f) be two relations. 'd' is the foreign key of S that refers to primary key of R. Consider the following four operations on S and R:
(1) Insert into R
(2) Insert into S
(3) Delete from R
(4) Delete from S
Which of these operations may cause violation of referential integrity constraint?
(A) None of (1), (2), (3) or (4) can cause its violation
(B) All of (1), (2), (3) and (4) can cause its violation
(C) Both (1) and (4) can cause its violation
(D) Both (2) and (3) can cause its violation
[GATE 1997]
ANSWER: (D)

ER Diagram to Relational Model

Q. An ER model of a database consists of entity types `A` and `B`. These are connected by a relationship `R` which does not have its own attribute. Under which of the following conditions, can the relational table for `R` be merged with that of `A`?
(a) Relation R is one-to-many and the participation of `A` in `R` is total
(b) Relation R is one-to-many and the participation of `A` in `R` is partial
(c) Relation R is many-to-one and the participation of `A` in `R` is total
(d) Relation R is many-to-one and the participation of `A` in `R` is partial
[GATE 2017]

ER Diagram to Relational Model

Q. An ER model of a database consists of entity types A and B. These are connected by a relationship R which does not have its own attribute. Under which of the following conditions, can the relational table for R be merged with that of A?
(a) Relation R is one-to-many and the participation of A in R is total
(b) Relation R is one-to-many and the participation of A in R is partial
(c) Relation R is many-to-one and the participation of A in R is total
(d) Relation R is many-to-one and the participation of A in R is partial
[GATE 2017]
ANSWER: (c) and (d)