CS F212 Database Systems

BITS Pilani
Hyderabad Campus

Prof.R.Gururaj
CS&IS Dept.

Conceptual Database Design (ER Modeling) Ch.7

innovate achieve lead

Content

- *ER Model*
- *Steps in Database Design Process*
- *Entities, Attributes, and Associations*
- *ER Notations*

Major Steps in Database Design Process

innovate achieve lead

Step 1: Requirement analysis
- Understanding the domain
- Identifying the data to be stored
- Identifying the constraints

Step 2: Conceptual Database design

E-R modeling/UML

Step 3: Logical Database Design

Designing tables and relationships

Step 4: Refinement of schema

Step 5: Physical database design
- Indexing
- Clustering
- Storage formats

ER Modeling

ER Model is a popular high-level (conceptual) data model.
It is an approach to designing Semantic Conceptual schema of a Database.
ER model allows us to describe the data involved in a real-world environment in terms of objects and their relationships, which are widely used in design of database.
ER model provides preliminary concepts or idea about the data representation which is later modified to achieve final detailed design.

ER Modeling

Important concepts/notions used in ER modeling are-

Entity is an object in real-world or some idea or concept which can be distinguished from other objects.

Ex.: person, school, class, department, weather, salary, temperature etc.

Entity has independent existence.

Each entity belongs to an *Entity type* that defines the structure.

Entity Set is a Collection of similar objects.

Concepts used in ER

Attribute: reflects a property of an object or entity. We have following types of attributes.

> Simple attribute
> Composite attribute
> Single valued attribute
> Multi-valued attribute
> Derived attribute
> Stored attribute

Key : Is an Attribute of an entity type whose value can uniquely identify an entity in a set.

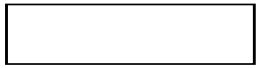
Concepts used in ER

Relationship: The association between entities is known as *relationship*.

Domain of an attribute: The set of possible values is known as domain of an attribute

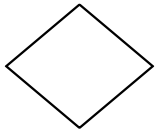
Notations used in ER

Notations used in ER modeling are shown below.

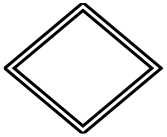


Entity Type

Weak Entity Type



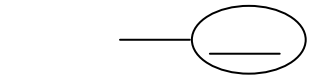
Relationship Type



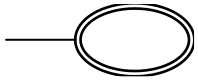
Identifying Relationship type

Attribute

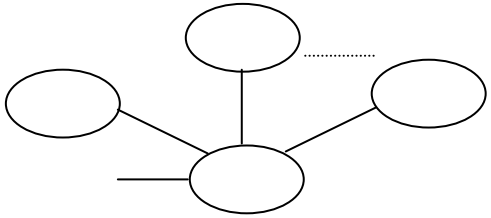
Notations used in ER



Key Attribute



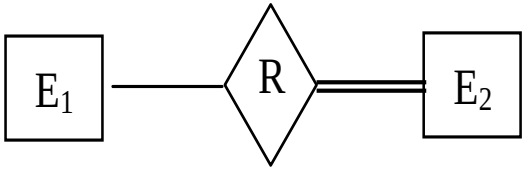
Multivalued Attribute



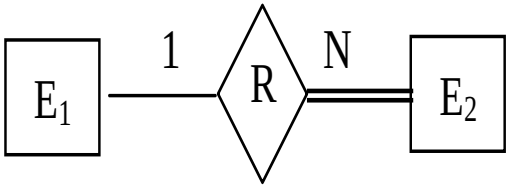
Composite Attribute

Derived Attribute

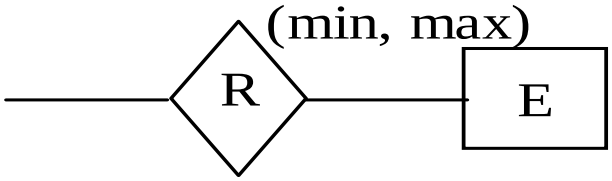
Notations used in ER



Total Participation of E_1 in R



Cardinality ratio 1; N for E_1; E_2 in R

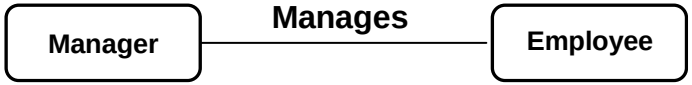


Structural Constraint (min, max) on Participation of E in R

Relationships in ER

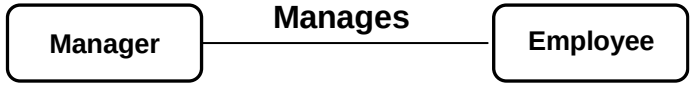
innovate achieve lead

Relationships

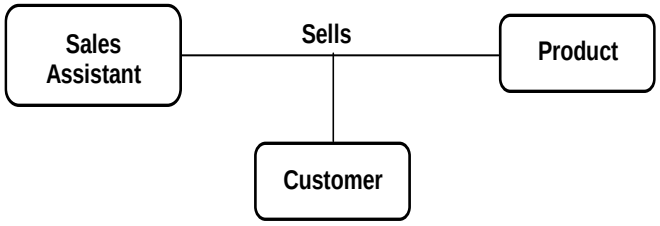


Degree of a Relationship

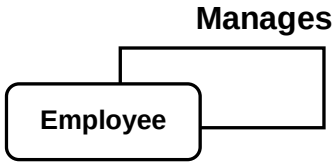
- If there are two entity types involved it is a binary relationship type



- If there are three entity types involved it is a ternary relationship type



- Unary relationships are also known as a recursive relationship



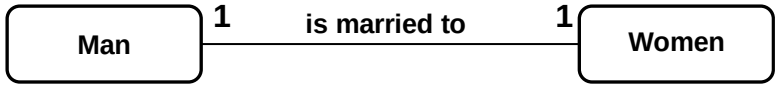
- It is possible to have n-ary relationship (e.g. quaternary or unary)

Relationships in ER

Cardinality of a relationship

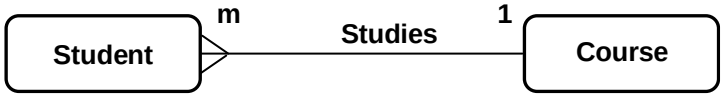
Relationships are rarely one-to-one.
For example, a manager usually manages more than one employee.
This is described by the cardinality of the relationship,
for which there are four possible categories.

One to one (1:1) relationship

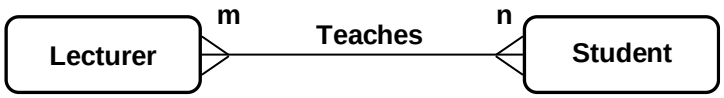


One to many (1:M) relationship

Many to one (M:1) relationship



Many to many (M:N) relationship

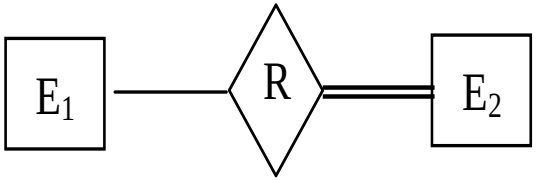


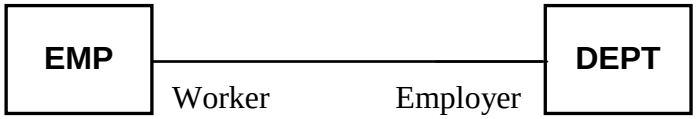
Relationships in ER

Participation Constraint

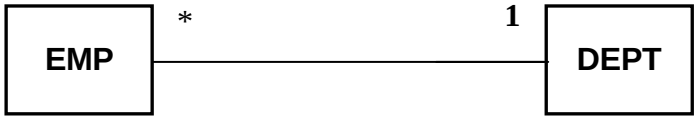
If all the entities of an entity type are involved in the relationship then that entity type's involvement said to be total in that relationship. In the below relationship if each employee is associated with at least one dept. Then the participation of EMP is total. Here, EMP works for DEPT.

If, only few entities of the set are involved the participation is partial.

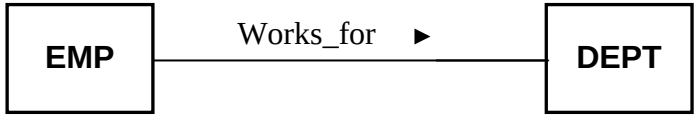




Association Role:



Multiplicity



Association Name & Direction:

Assume a scenario of Clubs (like- Dance Club, Music Club, Photography Club etc.) in an educational Institution, and we need to design a database. The description is follows.

Each Student has name, roll# (unique), address, contact (more than one contact possible), and branch.

There exist clubs (for specialized activities like Dance, Music, Photography etc.), each club has name (unique), start_date, Faculty-Incharge, Funds available.

Clubs conduct events. Each event has start-date, end-date, description, budget. Different clubs may conduct different events during same dates. But no single club will conduct more than one event at a given point of time i.e., a club cannot conduct more than one event with same start and end dates. One event is conducted by one club only. Note that Events will not have any IDs for.

Students can become cub members. One student can be member of more than one club. Some may not have any membership.

Students can become club office bearers (even for clubs where they don't have membership) in specific role like-secretary, treasures etc. One student can be office bearer for only one club. We capture info about present office bearers only

Summary

- ✓ *Various steps in database design process*
- ✓ *What is ER modeling*
- ✓ *Concepts and notations used in ER*

Introduction to EER

innovate achieve lead

Why EER

Some applications like –GIS, CAD/CAM, Telecommunication Have more complex requirements than normal database applications.

To meet the requirements additional modeling concepts were incorporated into Conceptual data modeling such as ER modeling.

The result is EER, stands for

Extended ER modeling or *Enhanced ER modeling*

The additional EER concepts are used to model applications more completely and more accurately.

EER includes some object-oriented concepts, such as inheritance

Additional concepts:

- Subclasses/super classes
- Specialization/generalization
- Attribute and relationship Inheritance

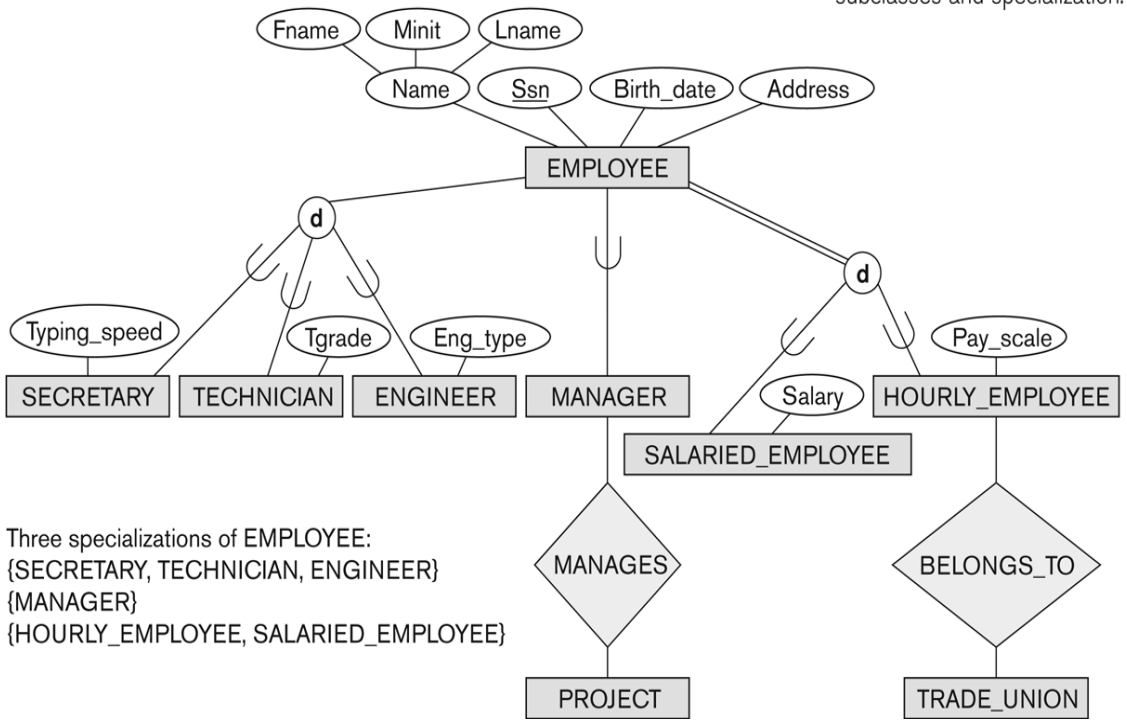
These are fundamental to conceptual modeling.

Subclasses in EER

Subclass of an Entity type

An entity type may have additional meaningful subgroupings of its entities.

Figure 4.1
EER diagram notation to represent subclasses and specialization.



- These are called superclass/subclass relationships
- Note: An entity that is member of a subclass represents the same real-world entity as some member of the superclass:
  - ❖ The subclass member is the same entity in a *distinct specific role*
  - ❖ An entity cannot exist in the database merely by being a member of a subclass; it must also be a member of the superclass
  - ❖ A member of the superclass can be optionally included as a member of any number of its subclasses

Attribute Inheritance in Superclass / Subclass Relationships

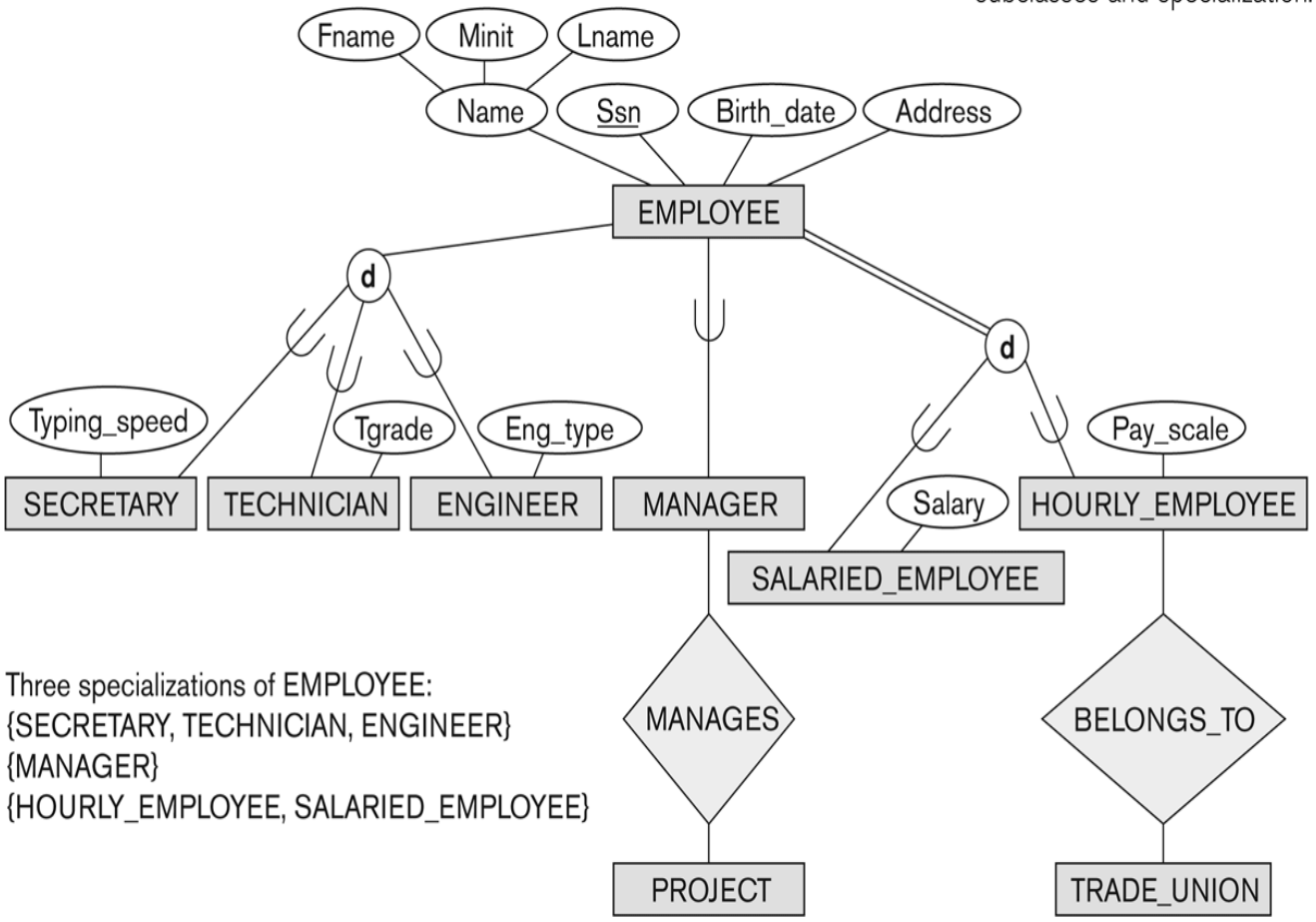
Specialization in EER

Specialization is the process of defining a set of subclasses of a superclass.

Generalization is the reverse of the specialization process.

Several classes with common features are generalized into a superclass. Original classes become its subclasses

Figure 4.1
EER diagram notation to represent subclasses and specialization.



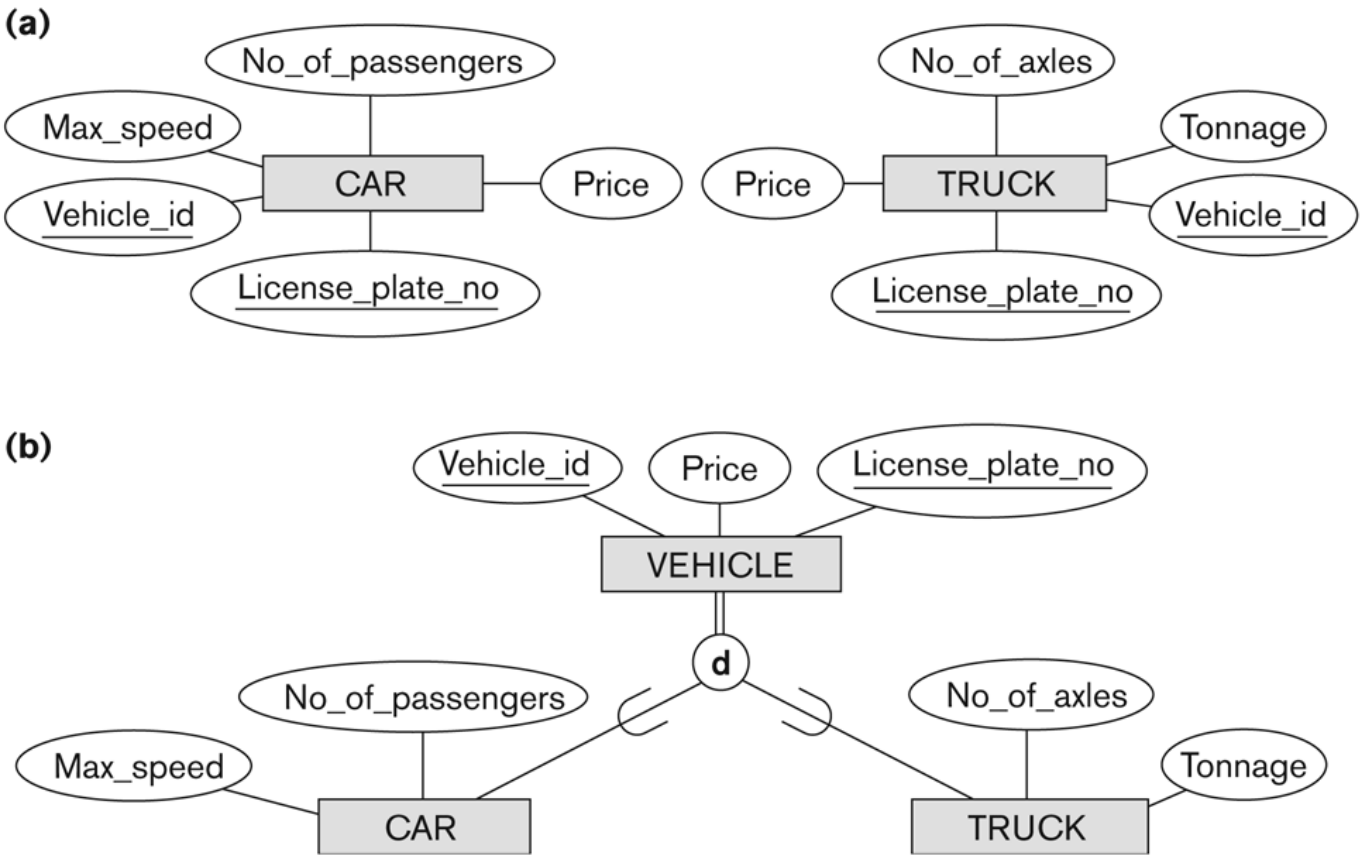


Figure 4.3
Generalization. (a) Two entity types, CAR and TRUCK.
(b) Generalizing CAR and TRUCK into the superclass VEHICLE.

Constraints

Two basic constraints can apply to a specialization/generalization:

- Disjointness Constraint:
- Completeness Constraint:

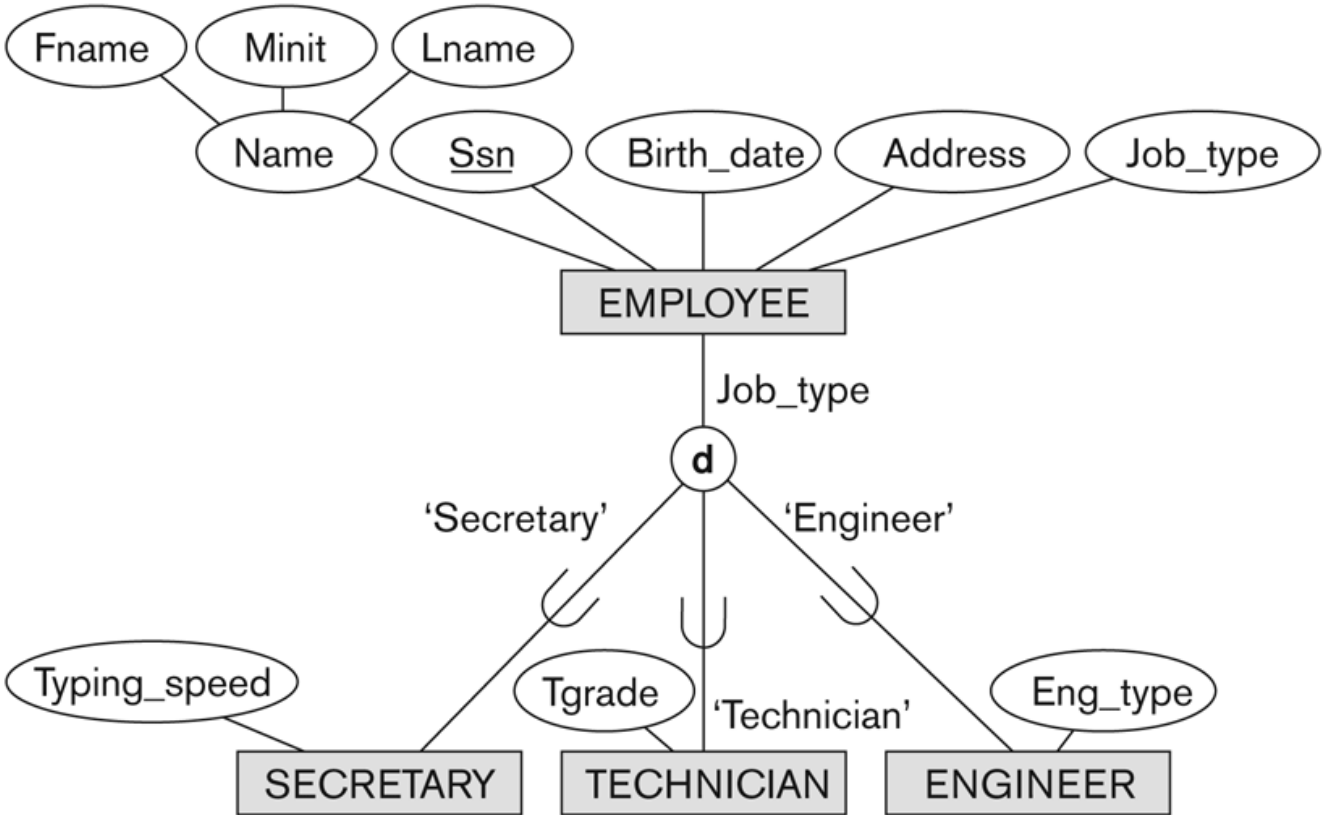
Hence, we have four types of specialization/generalization:

- Disjoint, total
- Disjoint, partial
- Overlapping, total
- Overlapping, partial

Note: Generalization usually is total because the superclass is derived from the subclasses.

Example of disjoint partial Specialization

Figure 4.4
EER diagram notation for an attribute-defined specialization on Job_type.



Example of overlapping total Specialization

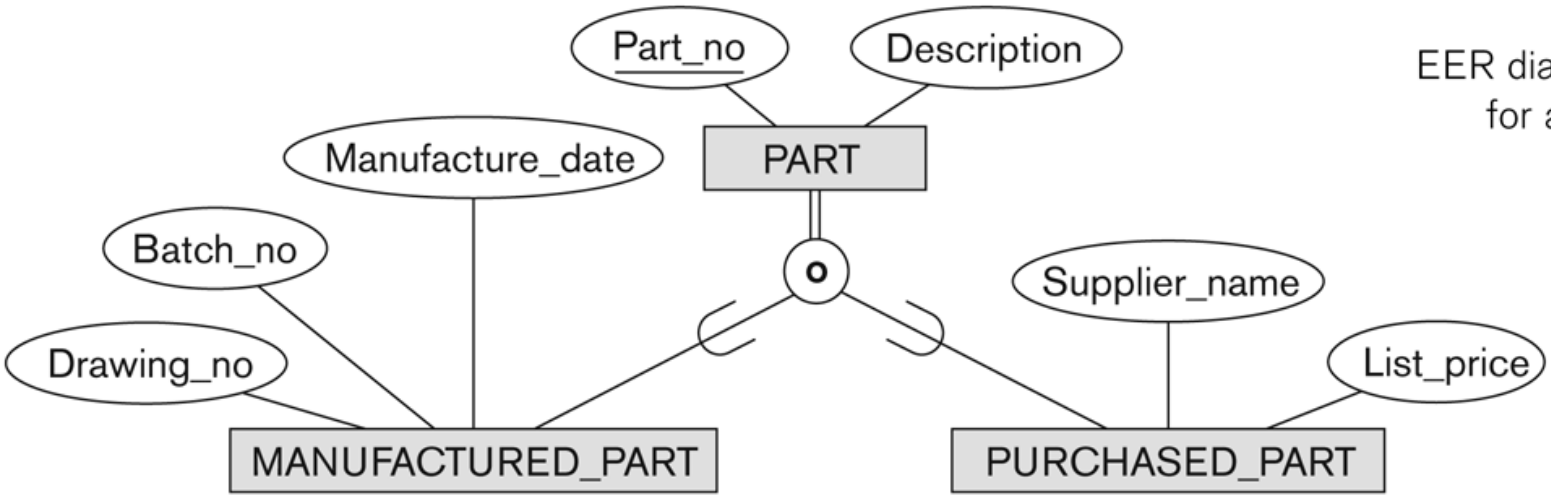


Figure 4.5
EER diagram notation for an overlapping (nondisjoint) specialization.

Exercise Problem

Q. Assume that we need to capture the data about a *Non-profit organization*(NPO) with following details/requirements.

i) The NPO depends on a number of different types of persons for its successful operations.

ii) The NPO is interested in following attributes of a person- SSN (identifier), name, address (consists of house#/city/state/zip components), and telephone.

iii) Three types of persons are of great interest to the NPO- employees, volunteers, and donors.

iv) Employees have only a date_hired attribute, volunteers have skill attribute, and donors have donor category as specific attribute.

v) Donors have a specific relationship (called Donates) with an item entity type(assume suitable attributes for Item entity type including a key).

vi) A donor must have donated one or more items, and an item may have no donors, or one or more donors.

vii) There are persons other than employees, volunteers, and donors who are of interest to the NPO, so that a person need not belong to any of these three types/groups. On the other hand, at a given time a person may belong to two or more of the above groups.

We discussed:

- What is EER
- Subclasses/super classes
- Specialization/generalization
- Attribute and relationship Inheritance
- Participation