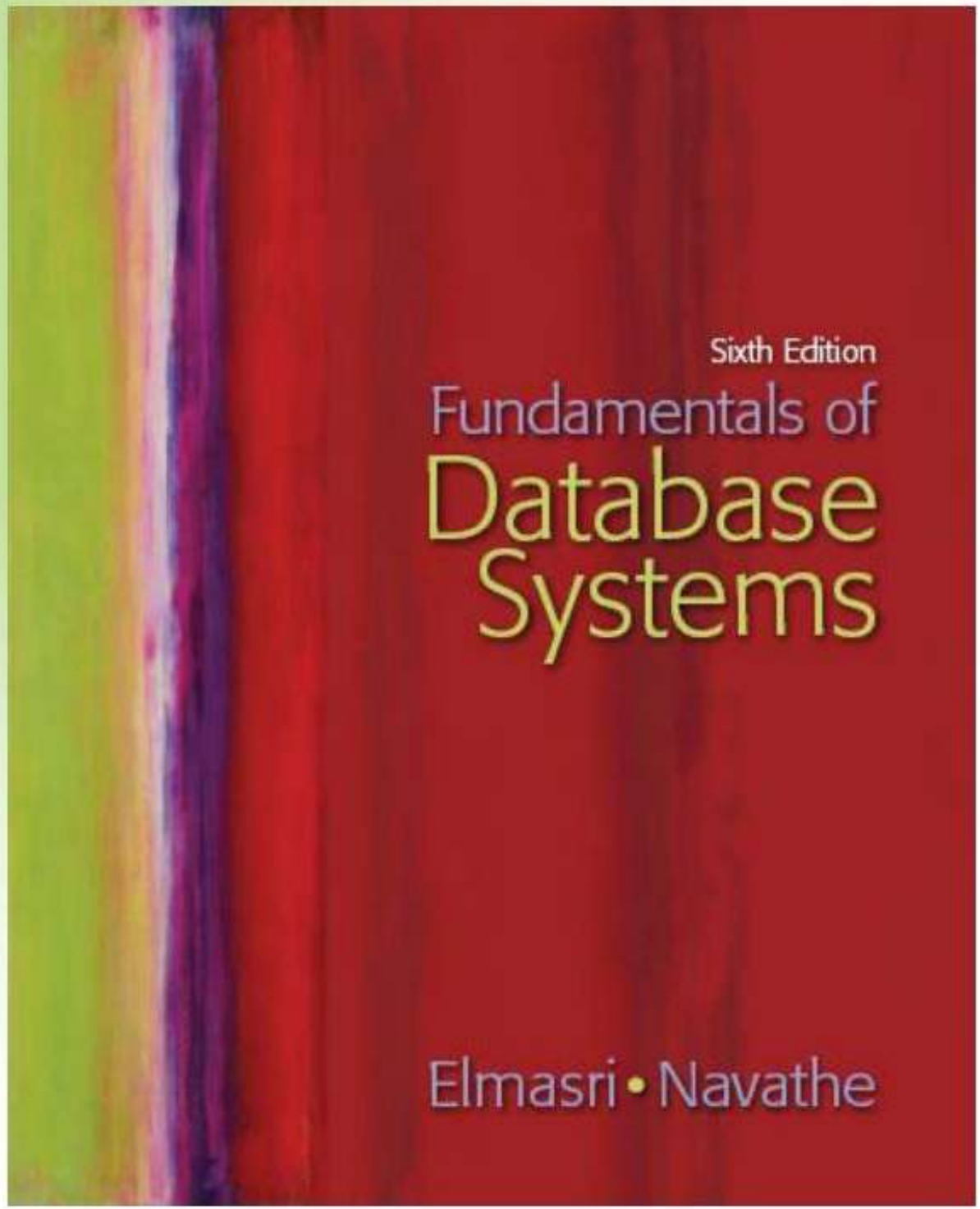
Chapter 8

The Enhanced Entity-Relationship (EER) Model



Addison-Wesley
is an imprint of
PEARSON

The Enhanced Entity-Relationship (EER) Model

- **Enhanced ER (EER) model**
  - Created to design more accurate database schemas
    - Reflect the data properties and constraints more precisely
  - More complex requirements than traditional applications

The Enhanced Entity-Relationship (EER) Model

- Enhanced ER model extends the basic ER model by adding following features:
  - An entity definition is known as a class.
    - A specific occurrence of an entity is an instance of a class.
  - Classes can be formed into superclass/subclass hierarchies using generalization and specialization.
    - The IS-A relationship.
    - Inheritance of attributes: Subclass entity inherits all attributes and relationships of superclass
    - Constraints on subclass membership.
  - Categories are used to represent a union of classes.

Specialization

- **Specialization**
  - Process of defining a set of subclasses of an entity type
  - Defined on the basis of some distinguishing characteristic of the entities in the superclass
  - Each subclass inherits all relationships and attributes from the super-class.
- Subclass can define:
  - **Specific attributes**
  - **Specific relationship types**
- Certain attributes may apply to some but not all entities of the superclass
- Some relationship types may be participated in only by members of the subclass

Specialization

- The specialization process allows us to do the following:
  - Define a set of subclasses of an entity type
  - Establish additional specific attributes with each subclass
  - Establish additional specific relationship types between each subclass and other entity types or other subclasses

Generalization

- Reverse process of abstraction
- **Generalize** into a single **superclass**
  - Original entity types are special subclasses
- **Generalization**
  - Process of defining a generalized entity type from the given entity types

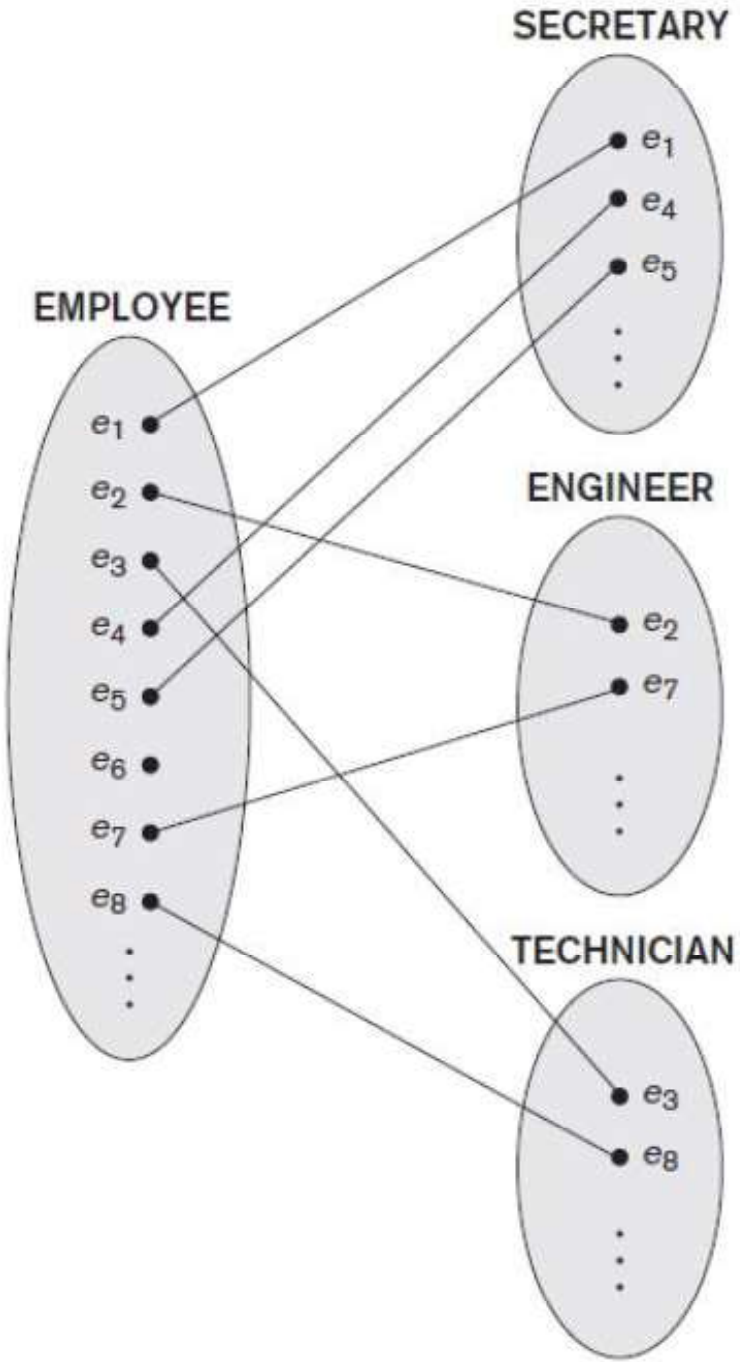
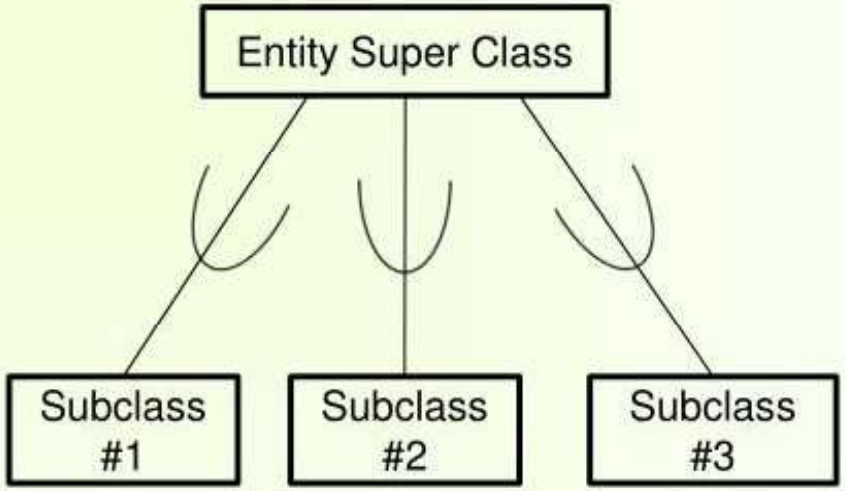


Figure 8.2
Instances of a specialization.

Enhanced ER Diagram

- Specialization/Generalization:



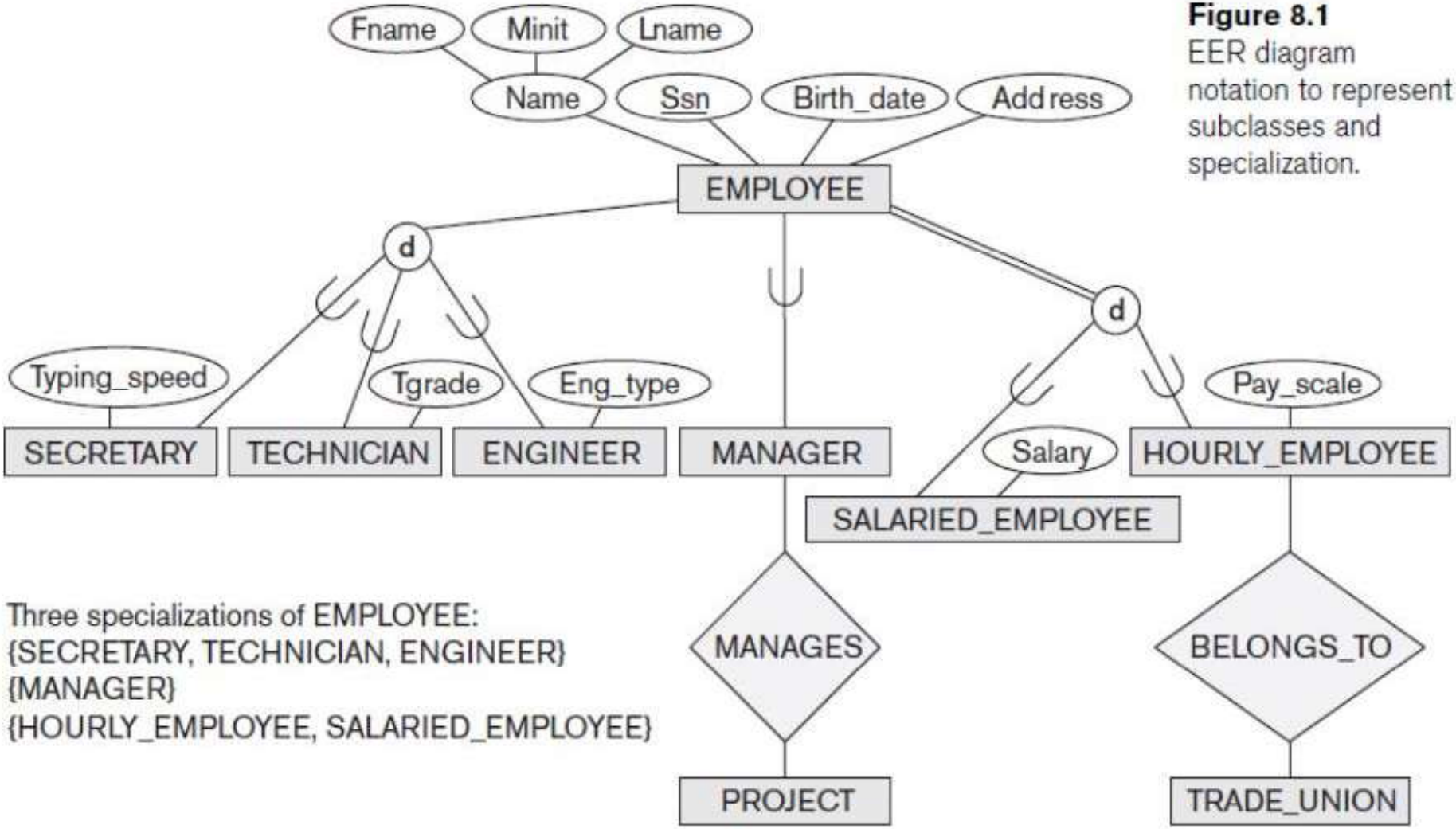


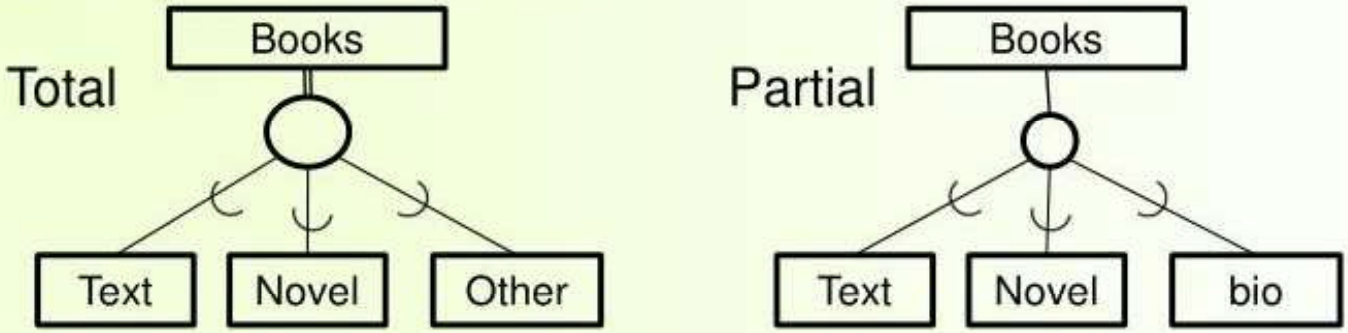
Figure 8.1
EER diagram notation to represent subclasses and specialization.

Constraints on Specialization and Generalization

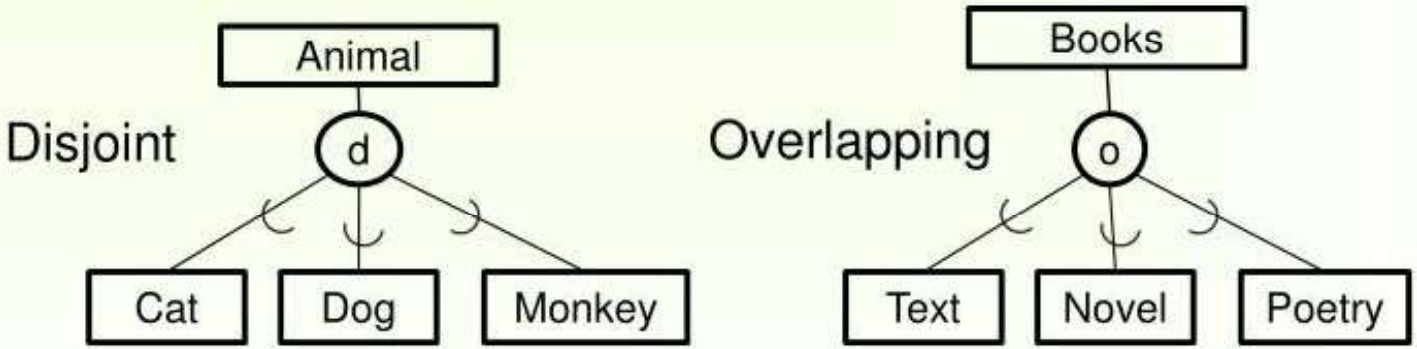
- **Completeness** (or **totalness**) **constraint**
  - **Total Specialization** – Every member of the super-class must belong to at least one subclass.
  - **Partial Specialization** – a member of the super-class may not belong to one of the subclasses.
- **Disjoint** – every member of the super-class can belong to at most one of the subclasses.
- **Overlapping** – a member of the super-class can belong to more than one of the subclasses.
- Disjointness and completeness constraints are independent

ER Diagram Notation (cont.)

- Total or partial specialization:



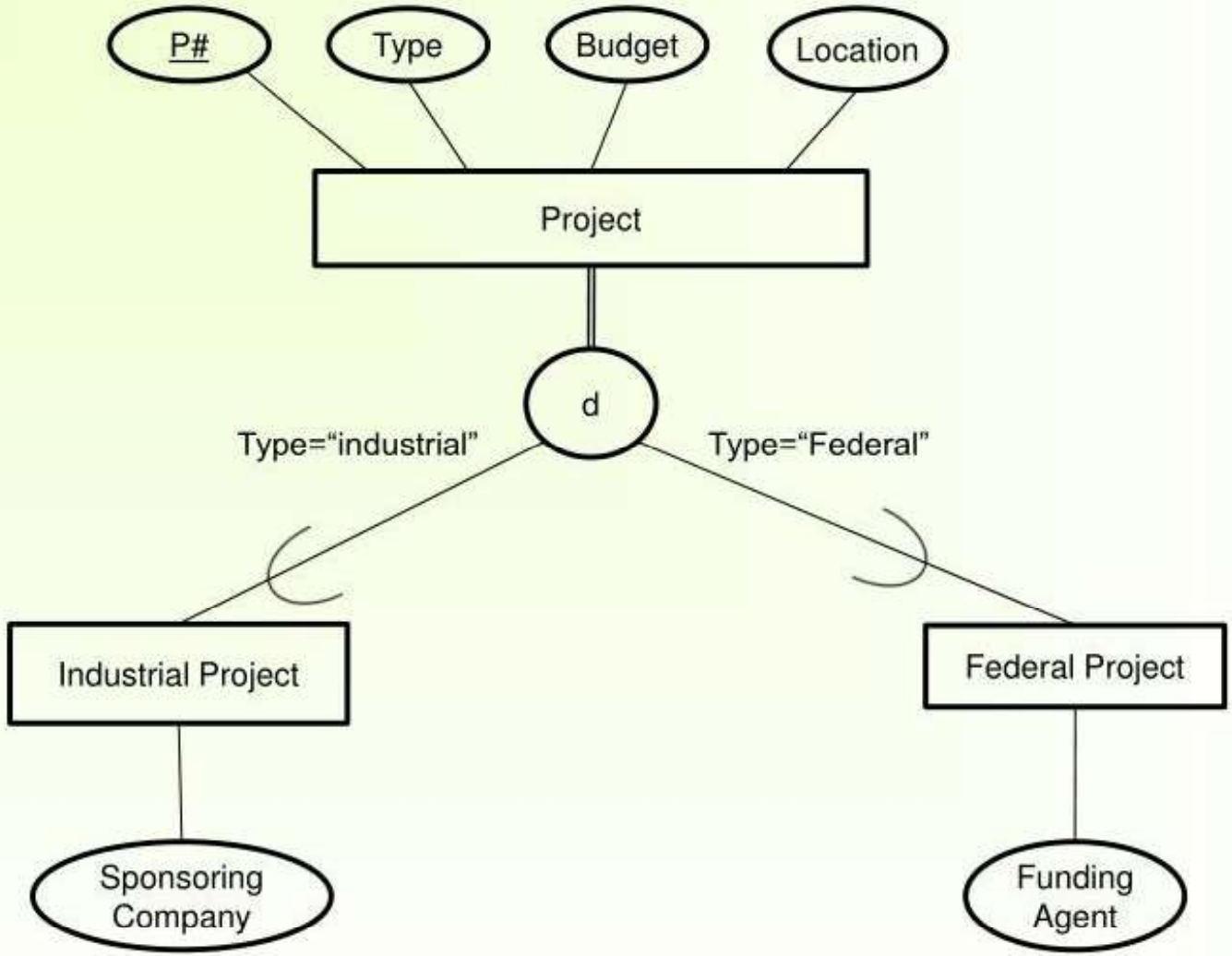
- Disjoint or overlapping specialization:



Constraints on Specialization and Generalization

- May be several or one subclass
- Determine entity subtype:
  - **Predicate-defined** (or c**ondition-defined**) **subclasses:** we can determine exactly the entities that will become members of each subclass by placing a condition on the value of some attribute of the superclass.
  - **Attribute-defined specialization:** all subclasses in a specialization have their membership condition on the same attribute of the superclass.
  - **User-defined:** we do not have a condition for determining membership in a subclass.

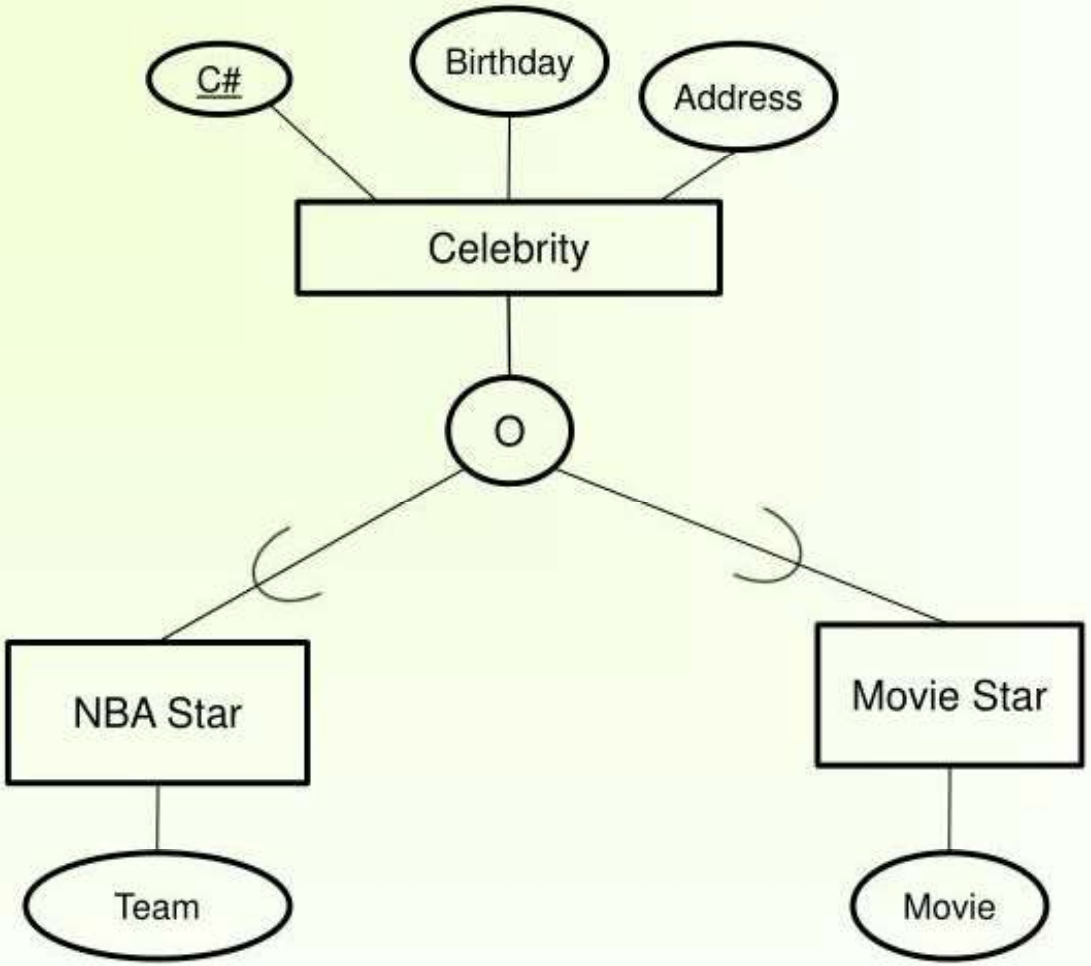
Attribute-defined Subclass



Rules for Attribute-defined Subclass

- If the specialization attribute at the superclass level is single-valued, membership at the subclass level is always disjoint.
- If the specialization attribute at the superclass level is multi-valued, membership at the subclass level is always overlapping.
- If the specialization is total, the attribute value in the superclass is required.
- If the specialization is partial, the specialization attribute value in the superclass is optional. The presence of a value, however, implies automatic insertion at the subclass level.

User-defined Subclass



Specialization and Generalization Hierarchies and Lattices

- **Specialization hierarchy**
  - Every subclass participates as a subclass in only one class/subclass relationship
  - Results in a **tree structure** or **strict hierarchy**
- **Specialization lattice**
  - Subclass can be a subclass in more than one class/subclass relationship

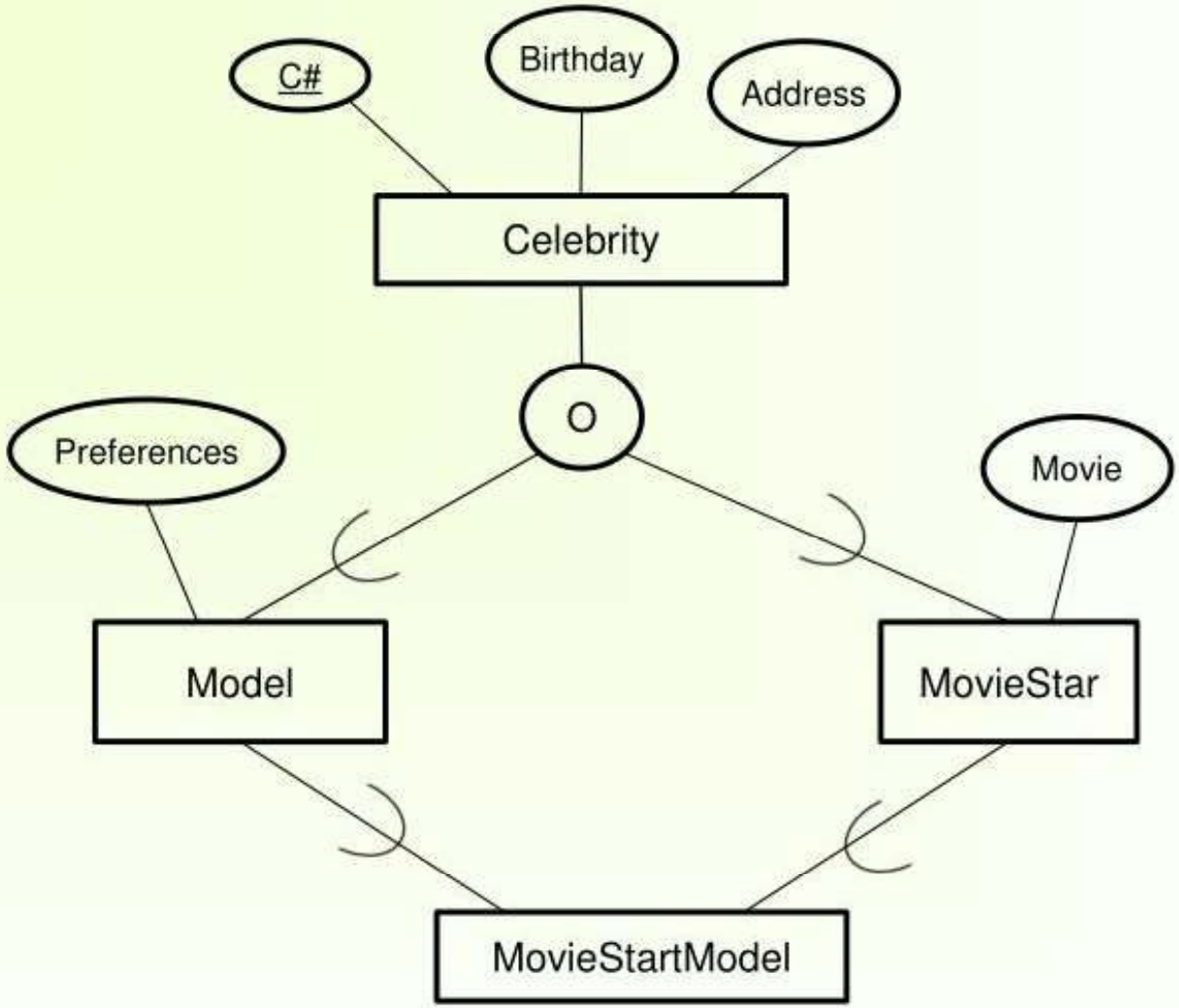
Rules for Superclass/Subclass Hierarchy

- Deleting an entity from a superclass implies automatic deletion of the entity from all subclasses.
- Deleting an entity from a subclass does not imply deleting the entity from its superclass. However, attributed-defined constraints must not be violated.
- At the superclass level, changing the value of an attribute used for attribute-defined specialization requires appropriate changes in subclass membership.

Specialization and Generalization Hierarchies and Lattices (cont'd.)

- **Multiple inheritance**
  - Subclass with more than one superclass
  - If attribute (or relationship) originating in the same superclass inherited more than once via different paths in lattice
    - Included only once in shared subclass
  - A shared subclass must satisfy the multiple inheritance intersection constraint, where each instance of the shared subclass is an instance of all of its superclasses
- **Single inheritance**
  - Some models and languages limited to single inheritance

Specification Lattice



MovieStarModel = Model ^ MovieStar

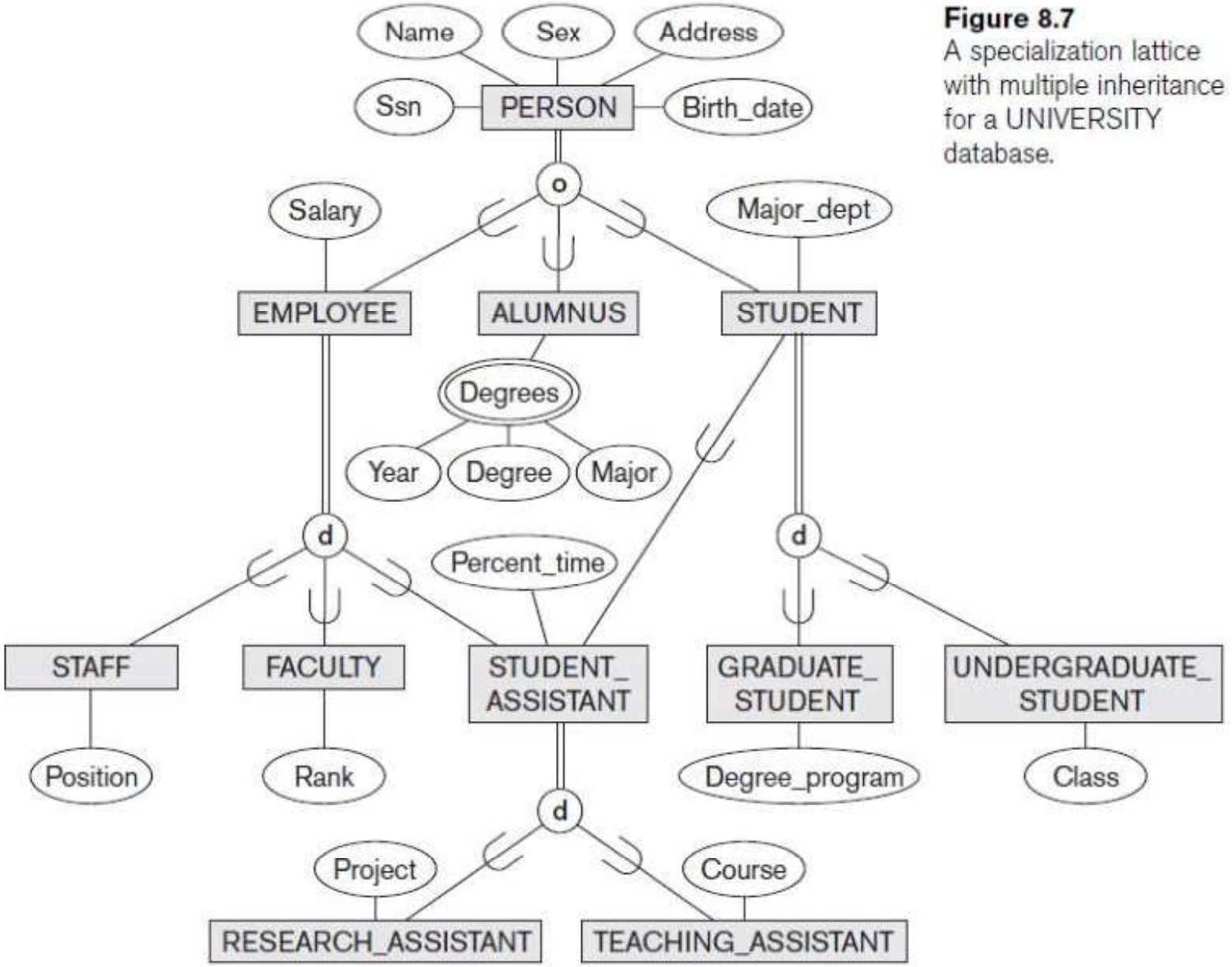


Figure 8.7
A specialization lattice with multiple inheritance for a UNIVERSITY database.

Utilizing Specialization and Generalization in Refining Conceptual Schemas

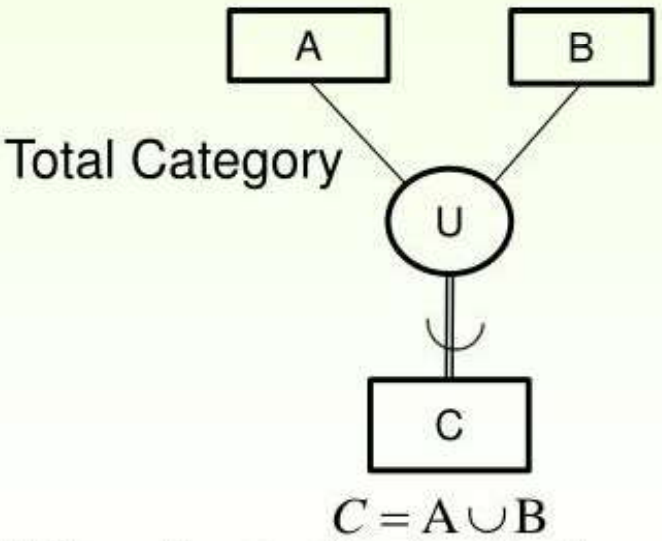
- Specialization process
  - Start with entity type then define subclasses by successive specialization
  - **Top-down conceptual refinement process**
- **Bottom-up conceptual synthesis**
  - Involves generalization rather than specialization

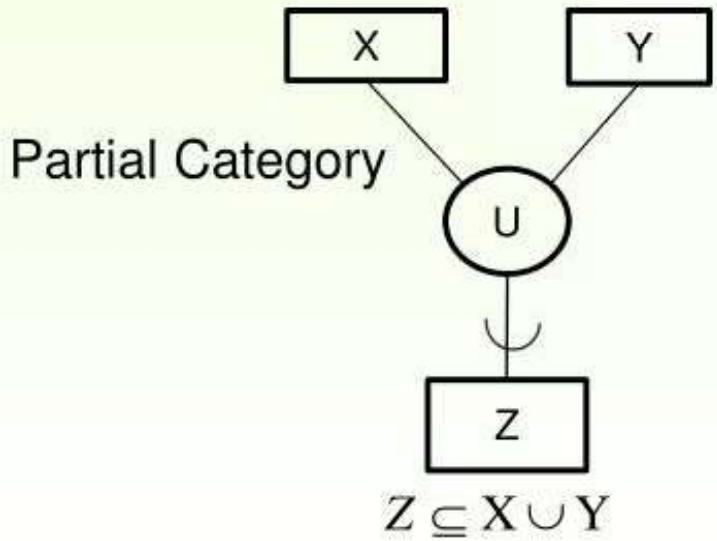
Modeling of UNION Types Using Categories

- **Union type** or a **category**
  - Represents a single superclass/subclass relationship with more than one superclass
  - Subclass represents a collection of objects that is a subset of the UNION of distinct entity types
  - Attribute inheritance works more selectively
  - A category represents a union of its superclasses, where an instance of a category subclass must be an instance of at least one superclass, but is not necessarily a member of all superclasses.

Categories and Categorization

- A category can be total or partial. A total category holds the union of all entities inits superclasses, whereas a partial category can hold a subset of the union.





A Sample UNIVERSITY EER Schema, Design Choices, and Formal Definitions

- The UNIVERSITY Database Example
  - UNIVERSITY database
    - Students and their majors
    - Transcripts, and registration
    - University's course offerings

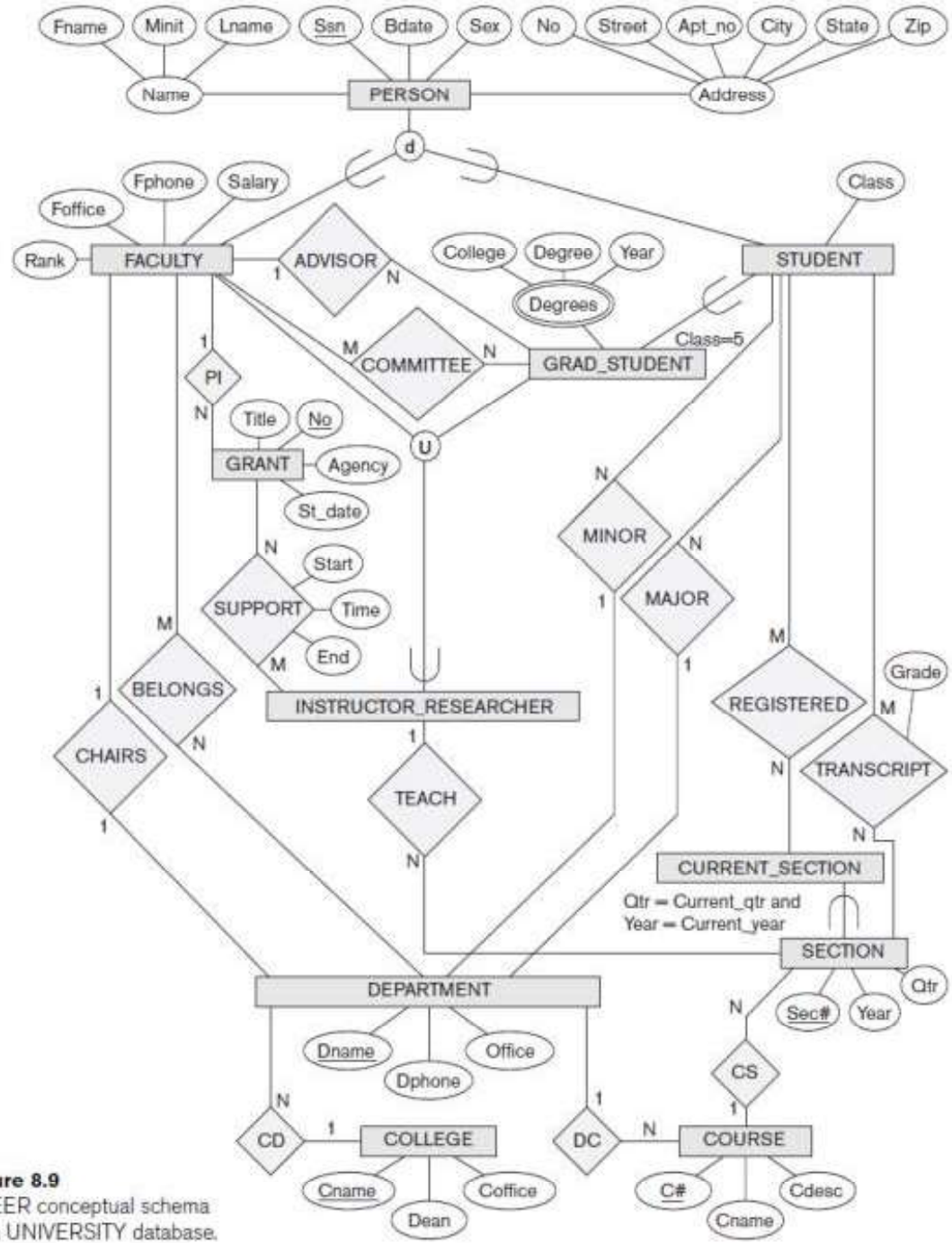


Figure 8.9
An EER conceptual schema for a UNIVERSITY database.

Design Choices for Specialization/Generalization

- Many specializations and subclasses can be defined to make the conceptual model accurate
- If subclass has few specific attributes and no specific relationships
  - Can be merged into the superclass
- If all the subclasses of a specialization/generalization have few specific attributes and no specific relationships
  - Can be merged into the superclass
  - Replace with one or more type attributes that specify the subclass or subclasses that each entity belongs to

Design Choices for Specialization/Generalization (cont'd.)

- Union types and categories should generally be avoided
- Choice of disjoint/overlapping and total/partial constraints on specialization/generalization
  - Driven by rules in miniworld being modeled

Formal Definitions for the EER Model Concepts

- **Class:** Set or collection of entities
- **Subclass:** Class whose entities must always be a subset of the entities in another class
- For such a superclass/subclass relationship, we must always have

$$S \subseteq C$$

Formal Definitions for the EER Model Concepts

- A specialization Z = {S1, S2, ..., Sn} is a set of subclasses that have the same superclass G; that is, G/Si is a superclass/subclass relationship for i = 1, 2, ..., n. G is called a generalized entity type (or the superclass of the specialization, or a generalization of the subclasses {S1, S2, ..., Sn}).
  - Z is said to be total if we always (at any point in time) have

$$\bigcup_{i=1}^{n} S_i = G$$

  - Otherwise, Z is said to be partial.
  - Zis said to be disjoint ifwe always have

    Si ∩ Sj = ∅ (empty set) for i ≠ j
  - Otherwise, Z is said to be overlapping.

Formal Definitions for the EER Model Concepts (cont'd.)

- A subclass S of C is said to be **predicate-defined** if a predicate p on the attributes of C is used to specify which entities in C are members of S; that is, S = C[p], where C[p] is the set of entities in C that satisfy p.
- A subclass that is not defined by a predicate is called **user-defined**.
- A specialization Z(or generalization G) is said to be **attribute-defined** if a predicate (A = ci), where A is an attribute of G and ci is a constant value from the domain of A, A specialization Z(or generalization G) is said to be attribute-defined if a predicate(A = ci), where A is an attribute of G and ci is a constant value from the domain of A,

Formal Definitions for the EER Model Concepts (cont'd.)

- **A category T is a class that is a subset of the union of n defining superclasses D1, D2,..., Dn, n > 1, and is formally specified as follows:**

$$T \subseteq (D_1 \cup D_2 \ldots \cup D_n)$$

- **A predicate pi on the attributes of Di can be used to specify the members of each Di that are members of T. If a predicate is specified on every Di, we get**

$$T = (D_1[p_1] \cup D_2[p_2] \ldots \cup D_n[p_n])$$