CHAPTER three

Entity-Relationship Modeling

chapter OVERVIEW

3.1 *Introduction*

In Chapter 2, we introduced data models. *A data model* is an integrated collection of concepts that represents real world objects, events, and their relationships. We also briefly discussed two types of data models: object-based data models and relation-based data models. It is a common practice in database design to develop an object-based model first and then to systematically convert the model into a relation-based model, which is more suitable for database implementation. In this chapter, we will focus on conceptual database design using object-based models. In Chapter 4, we will discuss the relational data models.

The three most popular object-based models discussed in database literature are the entity-relationship model, the functional model, and the object-oriented model. The entity-relationship model has emerged as one of the most popular techniques in the design of databases due to its inherent advantages. The entity-relationship (E-R) model is easy to learn, yet powerful enough to model complex, real-world scenarios. We have therefore chosen the E-R model for database design discussion.

3.1.1 Topics

This chapter discusses the following topics:

- Entities, attributes, and relationships in entity-relationship diagram.
- Degree of relationships: unary, binary, and ternary relationships.
- Cardinality of relationships: one-to-one relationships, one-to-many relationships, many-to-many relationships.
- Associative and weak entity types.
- Enhanced entity-relationship diagram: generalization and specialization process, and participation and disjoint constraints.

3.2 *The Entity-Relationship Model*

An *entity-relationship model* describes data in terms of the following:

1. Entities
2. Relationship between entities
3. Attributes of entities

We graphically display an E-R model using an ***entity-relationship diagram*** (or E-R diagram) like the sample in Figure 3.1. While this figure may seem to be confusing at first glance, its meaning should become very clear by the end of this chapter.

We will now discuss the components of an E-R diagram in detail.

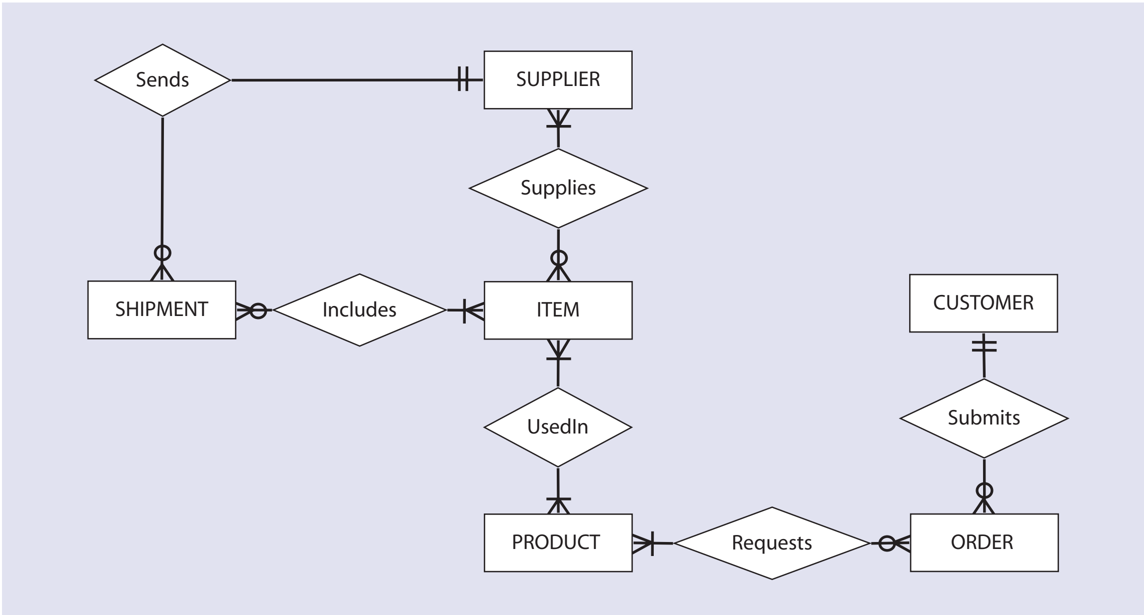


Figure 3.1 Example of an E-R diagram.

3.3 *Entity*

An ***entity*** is an object that exists and which is distinguishable from other objects. An entity can be a person, a place, an object, an event, or a concept about which an organization wishes to maintain data. The following are some examples of entities:

Person: STUDENT, EMPLOYEE, CLIENT
Object: COUCH, AIRPLANE, MACHINE
Place: CITY, NATIONAL PARK, ROOM, WAREHOUSE
Event: WAR, MARRIAGE, LEASE
Concept: PROJECT, ACCOUNT, COURSE

It is important to understand the distinction between an *entity type*, an *entity instance*, and an *entity set*. An ***entity type*** defines a collection of entities that have same attributes. An ***entity instance*** is a single item in this collection. An ***entity set*** is a set of entity instances. The following example will clarify this distinction: STUDENT is an entity type; a student with ID number 555-55-5555 is an entity instance; and a collection of all students is an entity set.

In the E-R diagram, we assign a name to each entity type. When assigning names to entity types, we follow certain naming conventions. An entity name should be a concise singular noun that captures the unique characteristics of the entity type. An E-R diagram depicts an entity type using a rectangle with the name of the entity inside (see Figure 3.2).

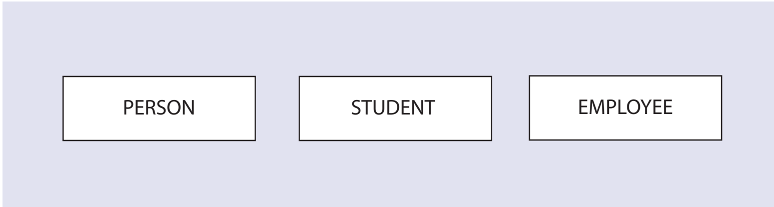


Figure 3.2 The entity representation in an E-R diagram.

3.4 Attributes

We represent an entity with a set of attributes. An ***attribute*** is a property or characteristic of an entity type that is of interest to an organization. Some attributes of common entity types include the following:

STUDENT = {Student ID, SSN, Name, Address, Phone, Email, DOB}
ORDER = {Order ID, Date of Order, Amount of Order}
ACCOUNT = {Account Number, Account Type, Date Opened, Balance}
CITY = {City Name, State, Population}

We use the following conventions while naming attributes:

1. Each word in a name starts with an uppercase letter followed by lower case letters.
2. If an attribute name contains two or more words, the first letter of each subsequent word is also in uppercase, unless it is an article or preposition, such as "a," "the," "of," or "about" (see Figure 3.3).

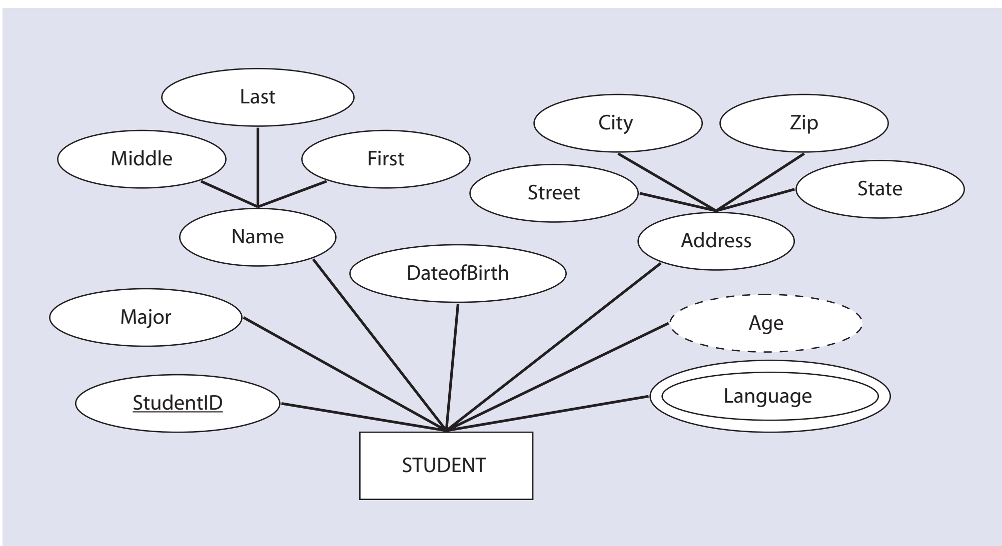


Figure 3.3 Attributes of the STUDENT entity type.

E-R diagrams depict an attribute inside an ellipse and connect the ellipse with a line to the associated entity type. Figure 3.3 illustrates some of the possible attributes in an E-R diagram for the entity STUDENT.

Notice that not all of the attributes in Figure 3.3 are marked in the same way. There are actually several types of attributes featured in this figure. These include: simple, composite, single-valued, multi-valued, stored, and derived attributes. In the following subsections, we discuss the distinctions between these types of attributes.

3.4.1 Simple and Composite Attributes

A ***simple*** or an ***atomic attribute***, such as *City* or *State*, cannot be further divided into smaller components. A ***composite attribute***, however, can be divided into smaller subparts in which each subpart represents an independent attribute. *Name* and *Address* are the only composite attributes in Figure 3.3. All other attributes, even those that are subcategories of *Name* and *Address*, are simple attributes. The figure also presents the notation that depicts a composite attribute.

3.4.2 Single-Valued and Multi-Valued Attributes

Most attributes have a single value for an entity instance; such attributes are called ***single-valued attributes***. A ***multi-valued attribute***, on the other hand, may have more than one value for an entity instance. Figure 3.3 features one multi-valued attribute, *Languages*, which stores the names of the languages that a student speaks. Since a student may speak several languages, it is a multi-valued attribute. All other attributes of the STUDENT entity type are single-valued attributes. For example, a student has only one date of birth and one student identification number. In the E-R diagram, we denote a multi-valued attribute with a double-lined ellipse. Note that in a multi-valued attribute, we always use a double-lined ellipse, regardless of the number of values.

3.4.3 Stored and Derived Attributes

The value of a ***derived attribute*** can be determined by analyzing other attributes. For example, in Figure 3.3 *Age* is a derived attribute because its value can be derived from the current date and the attribute *DateofBirth*. An attribute whose value cannot be derived from the values of other attributes is called a ***stored attribute***. As we will learn, a derived attribute *Age* is not stored in the database. Derived attributes are depicted in the E-R diagram with a dashed ellipse.

3.4.4 Key Attribute

A ***key attribute*** (or identifier) is a single attribute or a combination of attributes that uniquely identify an individual instance of an entity type. No two instances within an entity set can have the same key attribute value. For the STUDENT entity shown in Figure 3.3, *StudentID* is the key attribute since each student identification number is unique. *Name*, by contrast, cannot be an identifier because two students can have the same name. We underline key attributes in an E-R diagram (also see Figure 3.4).

Sometimes no single attribute can uniquely identify an instance of an entity type. However, in these circumstances, we identify a set of attributes that, when combined, is unique for each entity instance. In this case the key attribute, also known as ***composite key***, is not a simple attribute, but a composite attribute that uniquely identifies each entity instance.

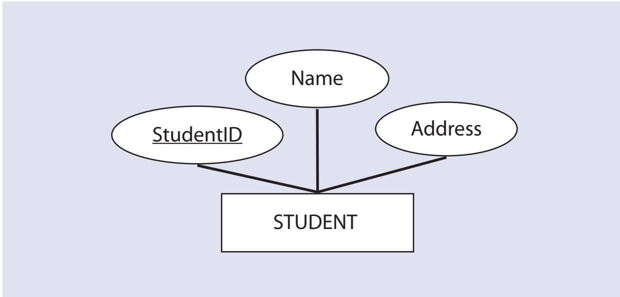


Figure 3.4 The key attribute.

A composite key must be minimal in the sense that no subset of a composite key can form the key of the entity instance. For example, if a composite key has four attributes, A1 to A4, then any subset, say A2, A4 or A2, A3 (or any of 16 combinations), should not form a key for an entity. In other words, we need all attributes, A1–A4, to identify each instance of an entity uniquely. In the E-R diagram, we underline each attribute in the composite key.

For example, consider the CITY entity type (see Figure 3.5). This category includes, potentially, all the cities in the United States. Notice that none of the attributes (i.e. *Name*, *State* or *Population*) can serve as a key attribute since there are many cities in each state and two cities could possibly have the same name or population. However, the composite attribute {*Name*, *State*} is a valid key attribute for the CITY entity as no two cities within a state can have the same name.

An entity can have more than one attribute that qualifies to be an identifier. For the entity shown in Figure 3.6, each of the attributes *Name*, *StateAbbr*, and *UnionOrder* (the order in which the state entered the union of the United States) can be an identifier. In this case, it is a matter of preference as to which attribute is made an identifier or key attribute.

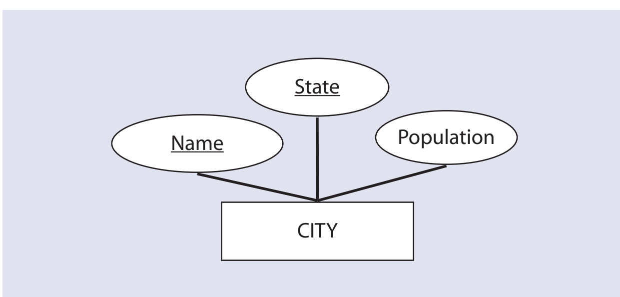


Figure 3.5 The composite key attribute.

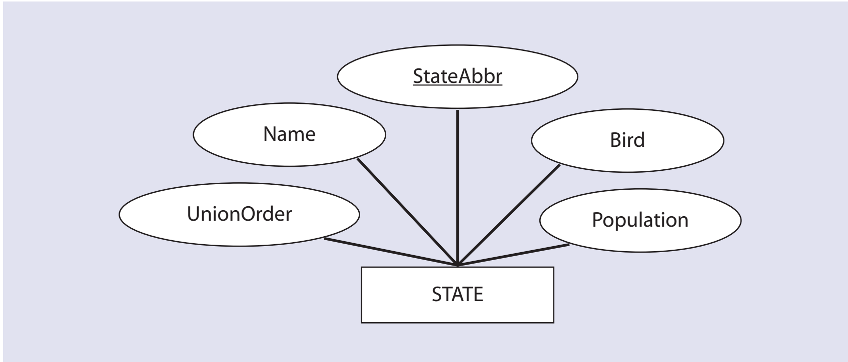


Figure 3.6 An example of more than one key attribute.

3.5 *Relationships*

Entities in an organization do not exist in isolation but are related to each other. Students take courses and each STUDENT entity is related to the COURSE entity. Faculty members teach courses and each FACULTY entity is also related to the COURSE entity. Consequently, the STUDENT entity is related to the FACULTY entity through the COURSE entity. E-R diagrams can also illustrate relationships between entities.

We define a ***relationship*** as an association among several entities. Consider, for example, an association between customers of a bank. If customer Williams has a bank account number 523, then the quality of ownership constitutes a ***relationship instance*** that associates the CUSTOMER instance Williams with the ACCOUNT instance *523*. We can think of the relationship instance as a verb that links a subject and an object: customer Williams *has* an account; student John *registers* for a course; professor Smith *teaches* a course. A ***relationship set*** is a grouping of all matching relationship instances, and the term ***relationship type*** refers to the relationship between entity types. For example, Figure 3.7 illustrates a relationship set between the CUSTOMER and the ACCOUNT instances.

In an E-R diagram, we represent relationship types with diamond-shaped boxes connected by straight lines to the rectangles that represent participating entity types. A relationship type is a given name that is displayed in this diamond-shaped box and typically takes the form of a

Definition

An **entity** is an object that exists and that is distinguishable from other objects.

An **attribute** is a property or characteristic of an entity type that is of interest to an organization.

A **relationship** is an association among several entities.

present tense verb or verb phrase that describes the relationship. An E-R diagram may depict a relationship as the following example of the relationship between the entities CUSTOMER and ACCOUNT does:

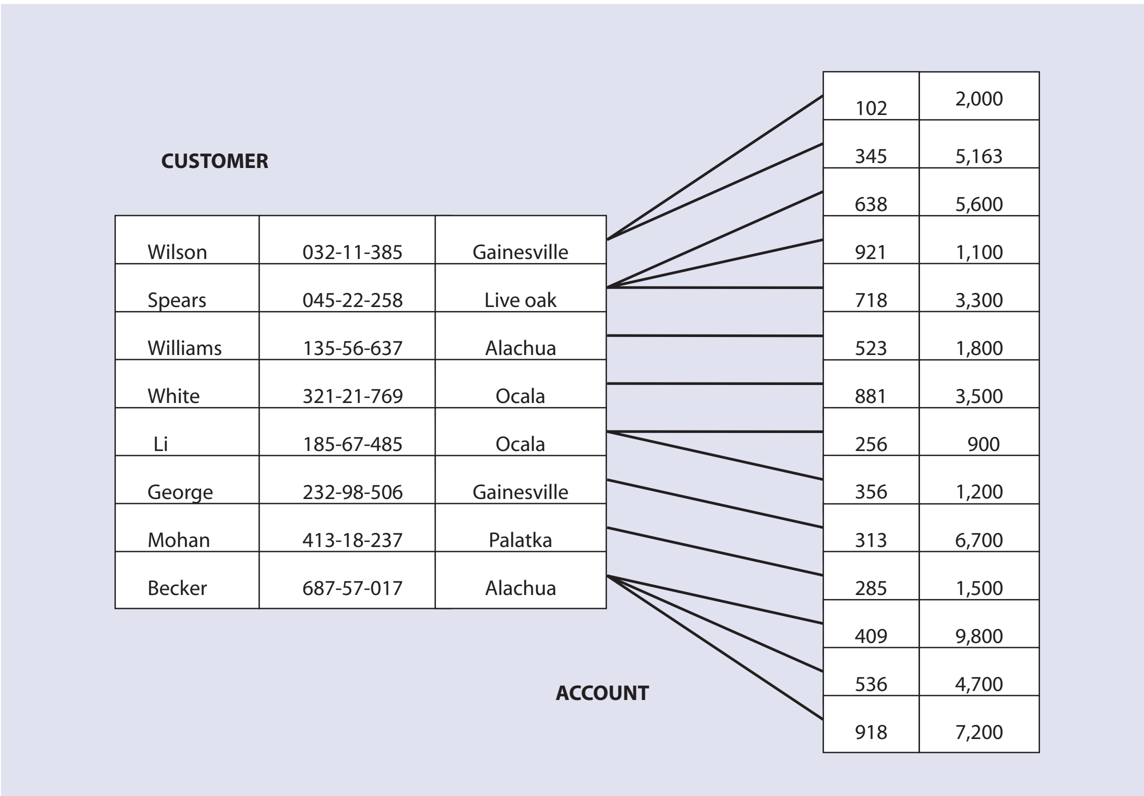


Figure 3.7 The relationship set between the CUSTOMER and ACCOUNT entities.

3.6 *Degree of a Relationship*

The number of entity sets that participate in a relationship is called the ***degree of relationship***. For example, the degree of the relationship featured in Figure 3.8 is two because CUSTOMER and ACCOUNT are two separate entity types that participate in the relationship. The three most common degrees of a relationship in a database are unary (degree 1), binary (degree 2), and ternary (degree 3). We will briefly define these degrees and then explore each kind of relationship in detail in subsequent sections.

Let $E_1, E_2, \ldots, E_n$ denote *n* entity sets and let R be the relationship. The degree of the relationship can also be expressed as follows:

Unary Relationship A unary relationship R is an association between two instances of the same entity type (i.e., $R \in E_1 \times E_1$). For example, two students are roommates and stay together in an apartment. Because they share the same address, a unary relationship exists between them for the attribute *Address* in Figure 3.3.

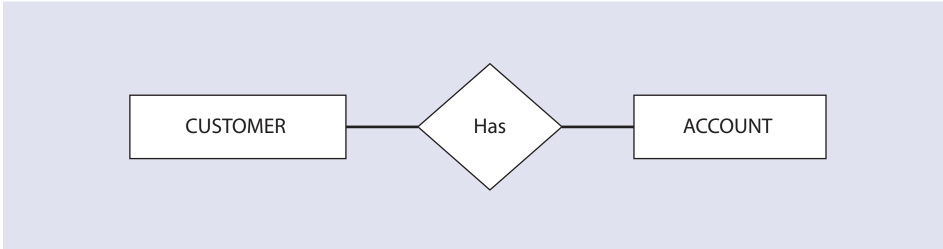


Figure 3.8 The relationship between CUSTOMER and ACCOUNT entities in an E-R diagram.

Binary Relationship A binary relationship R is an association between two instances of two different entity types (i.e., $R \in E_1 \times E_2$). For example, in a university, a binary relationship exists between a student (STUDENT entity) and an instructor (FACULTY entity) of a single class; an instructor *teaches* a student.

Ternary Relationship A ternary relationship R is an association between three instances of three different entity types (i.e., $R \in E_1 \times E_2 \times E_3$). For example, consider a student using certain equipment for a project. In this case, the STUDENT, PROJECT, and EQUIPMENT entity types relate to each other with ternary relationships: a student *checks out* equipment for a project.

3.7 *Cardinality of a Relationship*

The term *cardinal number* refers to the number used in counting. An *ordinal number*, by contrast, emphasizes the order of a number (1st, 7th, etc.). When we say cardinality of a relationship, we mean the ability to count the number of entities involved in that relationship. For example, if the entity types A and B are connected by a relationship, then the ***maximum cardinality*** represents the maximum number of instances of entity B that can be associated with any instance of entity A.

However, we don't need to assign a number value for every level of connection in a relationship. In fact, the term *maximum cardinality* refers to only two possible values: one or many. While this may seem to be too simple, the division between one and many allows us to categorize all of the permutations possible in any relationship. The maximum cardinality value of a relationship, then, allows us to define the four types of relationships possible between entity types A and B. Figure 3.9 illustrates these types of relationships.

One-to-One Relationship In a one-to-one relationship, at most one instance of entity B can be associated with a given instance of entity A and vice versa.

One-to-Many Relationship In a one-to-many relationship, many instances of entity B can be associated with a given instance of entity A. However, only one instance of entity A can be associated with a given instance of entity B. For example, while a customer of a company can make many orders, an order can only be related to a single customer.

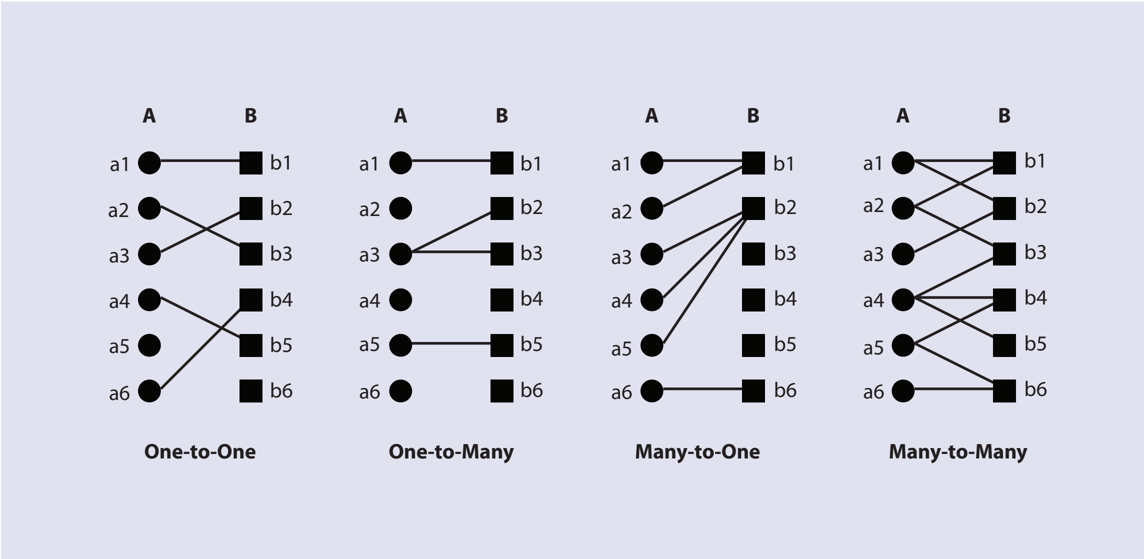


Figure 3.9 The four types of relationships between entity types A and B.

Many-to-Many Relationship In a many-to-many relationship, many instances of entity A can be associated with a given instance of entity B, and, likewise, many instances of entity B can be associated with a given instance of entity A. For example, a machine may have different parts, while each individual part may be used in different machines.

Representing Relationship Types Figure 3.10 displays how we represent different relationship types in an E-R diagram. An entity on the *one* side of the relationship is represented by a vertical line, "I," which intersects the line connecting the entity and the relationship. Entities on the *many* side of a relationship are designated by a crowfoot as depicted in Figure 3.10.

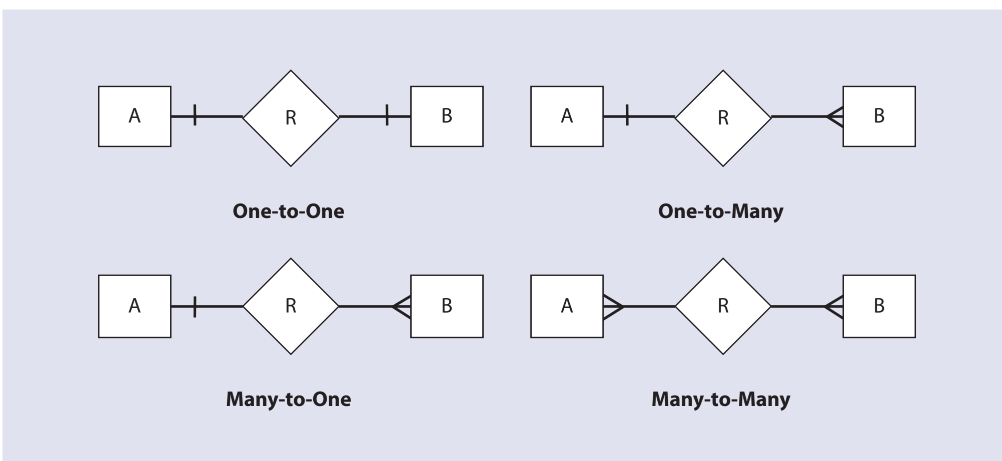


Figure 3.10 The relationship types based on maximum cardinality.

We will now discuss the minimum cardinality of a relationship. The ***minimum cardinality*** between two entity types A and B is defined as the minimum number of instances of entity B that must be associated with each instance of entity A. In an E-R diagram, we allow the minimum cardinality to take two values: zero or one. If the minimum cardinality is zero, we say that entity type B is an *optional* participant in the relationship; otherwise, it is a *mandatory* participant. An optional relationship is represented by an "O" and mandatory relationship is represented by "|" in an E-R diagram.

Figure 3.11 shows the four possibilities of the minimum cardinality of a relationship between two entity types A and B. Figure 3.11(a) depicts a situation in which no minimum cardinality constraints exist between the instances of entities A and B, meaning both entities A and B are optional participants in the relationship. Figure 3.11(b) illustrates a situation in which each instance of entity B must be associated with at least one instance of entity A, but no association is required for an instance of entity A. Figure 3.11(c) illustrates a situation in which each instance of entity A must be associated with at least one instance of entity B, but no association is required for an instance of entity B. Finally, Figure 3.11(d) illustrates a situation in which each instance of entity A and B must be associated with at least one instance of entity B and A, respectively.

An E-R diagram displays both the maximum and the minimum cardinalities of the relationships between two entities. Since there are four basic possibilities of maximum cardinalities and four possibilities of minimum cardinalities between two entities, there are 16 types of relationships possible between two entities in terms of cardinality. We will see several examples of these relationships while studying unary, binary, and ternary relationships.

Definition

The number (unary, binary, or ternary) of entity sets that participate in a relationship is called the **degree of relationship**.

The **cardinality of relationship** represents the minimum/maximum number of instances of entity B that must/can be associated with any instance of entity A.

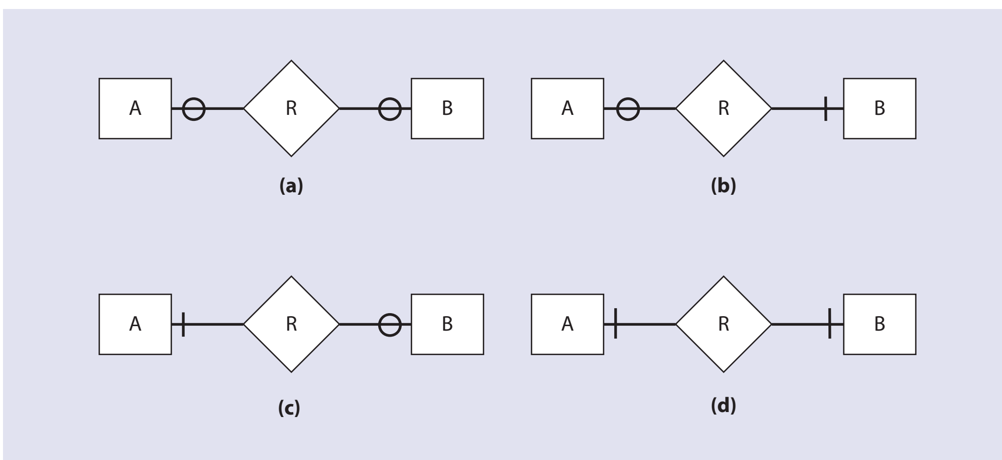
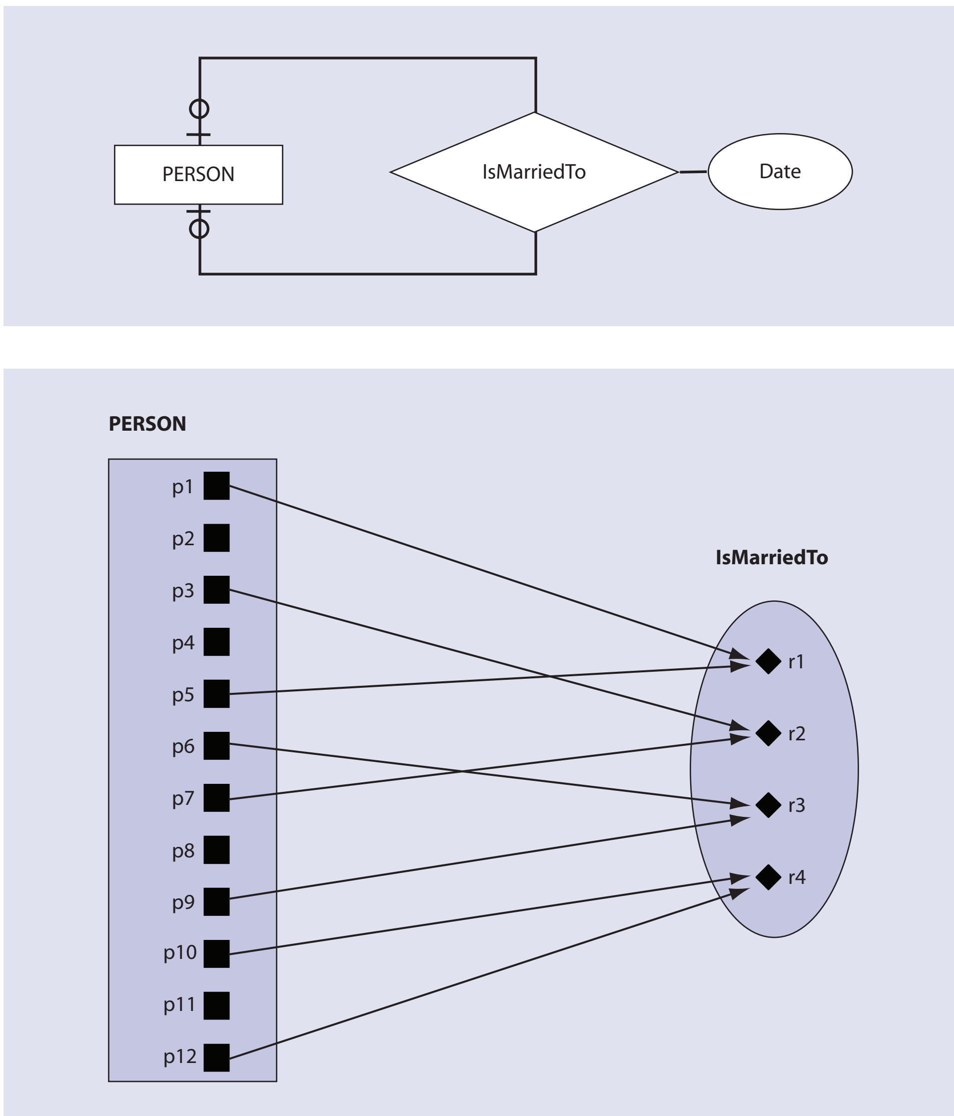


Figure 3.11 The relationship types based on minimum cardinality.

3.8 *Unary Relationship*

As we mentioned in Section 3.5, a unary relationship ($R \in E_1 \times E_1$) is an association between two entities of the same entity type. Figure 3.12(a) displays the E-R diagram of a unary relationship *IsMarriedTo*. Whenever two people in the entity type PERSON get married, the relationship instance *IsMarriedTo* is created. The *Date* of marriage is an attribute of this relationship. Since a person can only be married to one other person, marriage is a one-to-one relationship. Furthermore, since a person can be unmarried, the minimum cardinality of the *IsMarriedTo* relationship is zero.

Figure 3.12(b) depicts several relationship instances of this relationship type. Each relationship instance (r1, r2, r3, and r4) connects two instances in PERSON. The lines allow us to



(a)

(b)

Figure 3.12 The unary one-to-one relationship.

read relationships between entity instances. For example, `r1` suggests that person `p1` is married to person `p5`, and so forth.

Figure 3.13(a) illustrates another example of a unary relationship, expressed by the relationship instance *Supervises*. This relationship instance exists whenever an employee supervises another employee. The relationship *Supervises* is a one-to-many relationship since an employer can supervise many employees but a supervisee can have only one supervisor. The minimum cardinality for supervising is zero (an employee may not supervise anyone) but the minimum cardinality of being supervised is one (every employee must be supervised).

Figure 3.13(b) shows several relationship instances of this relationship type. Each relationship instance (`r1`, `r2`, `r3`, and `r4`) connects two entity instances in EMPLOYEE; these are the supervisor instance and the supervisee instance. For example, `r2` suggests that employee `e3` supervises employee `e7`, and so forth.

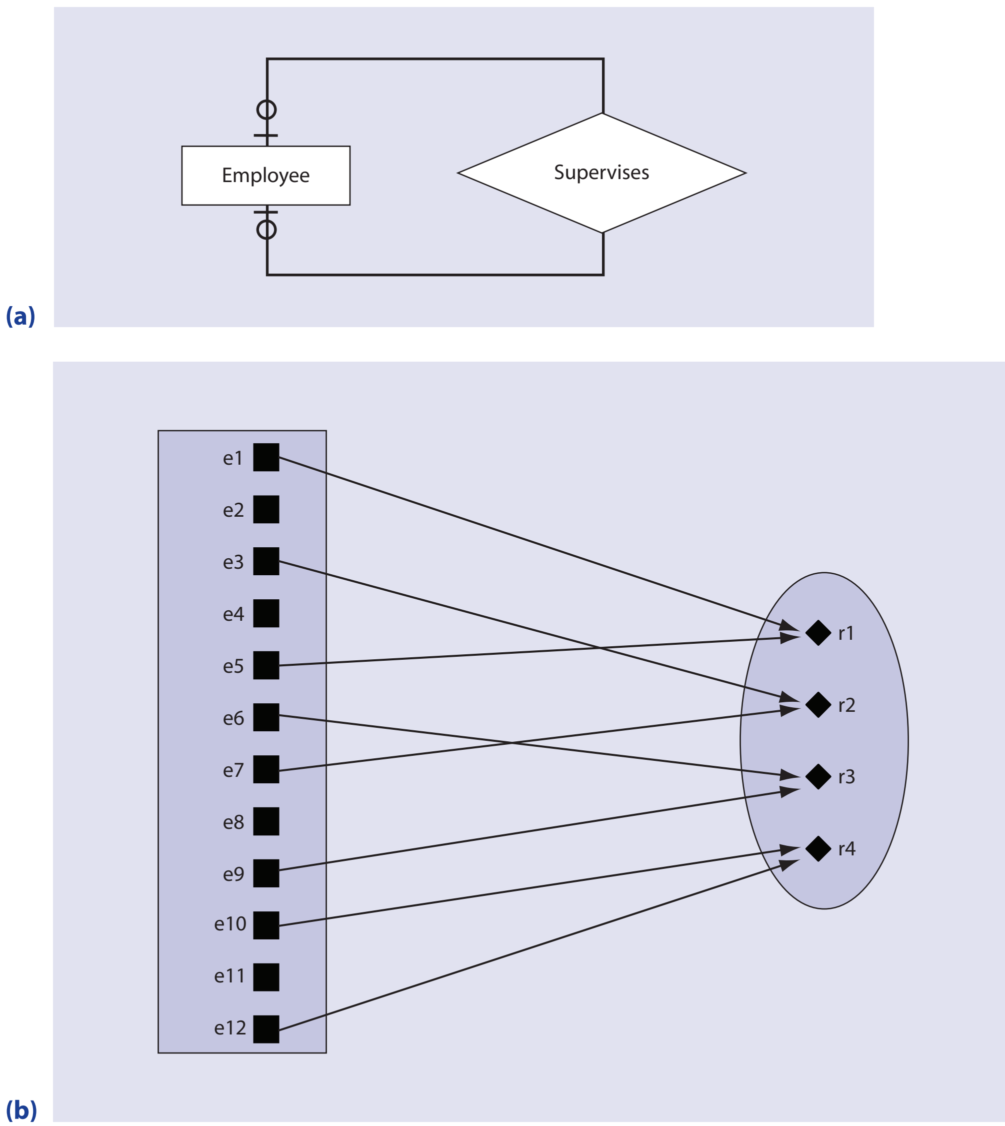


Figure 3.13 The unary one-to-many relationship.

(a)

(b)

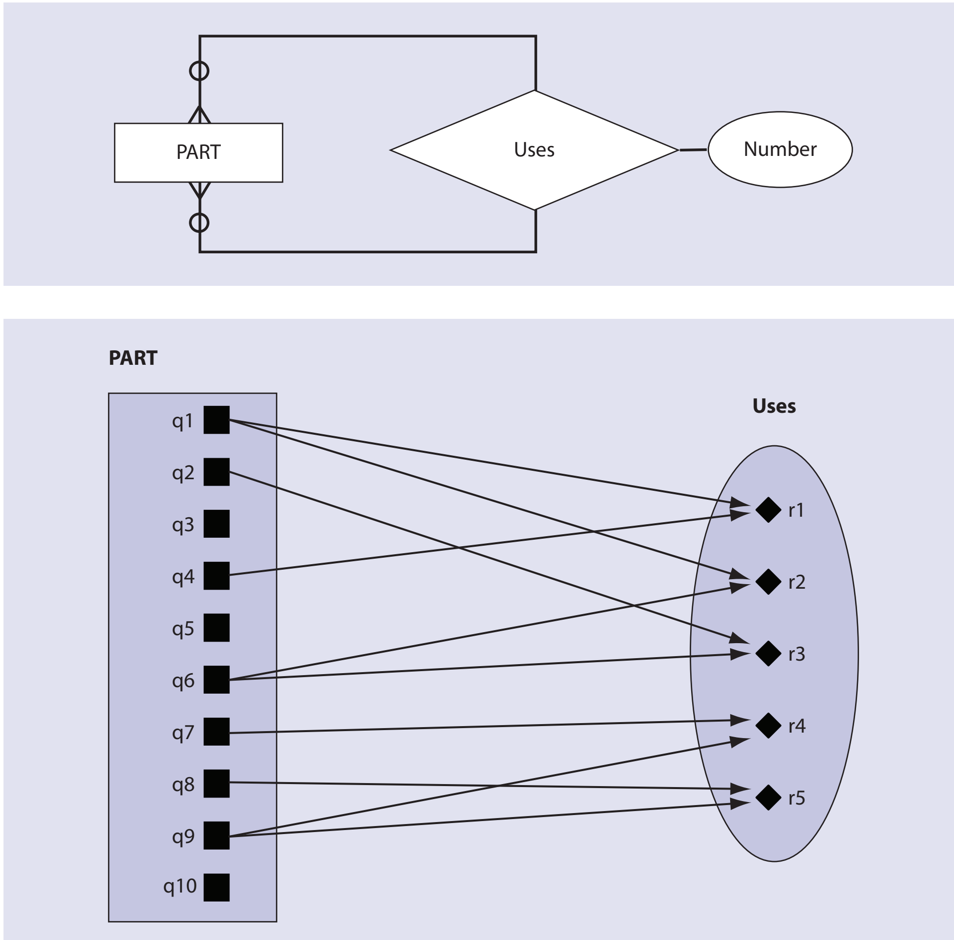


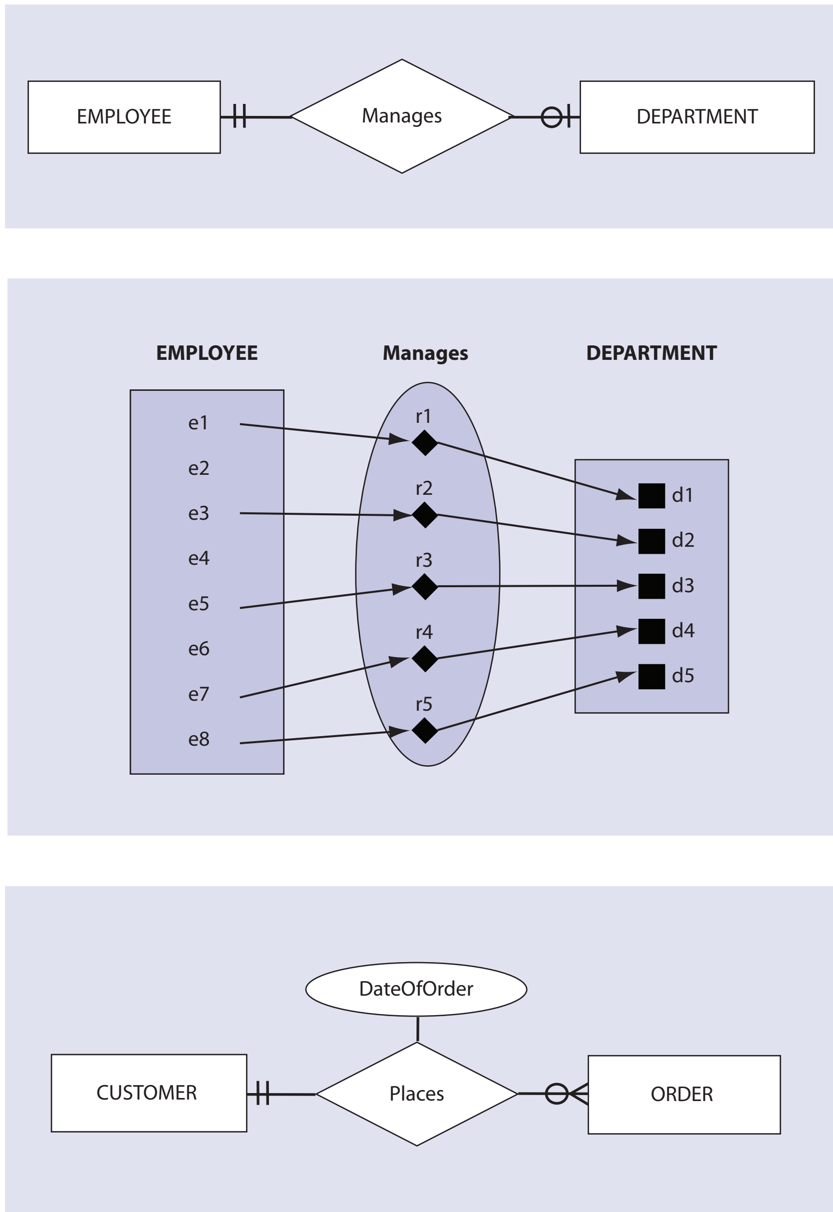
Figure 3.14 The unary many-to-many relationship.

Figure 3.14(a) provides an example of a unary relationship in which certain parts are used to make other parts. The attribute *Number* stores the number of parts used. The relationship *Uses* quantifies the number of parts used to make any other part. The *Uses* relationship is a many-to-many relationship since a part can use many parts and can be used in the making of many parts. The minimum cardinality of this relationship is zero because a part may not use any other part and may not be used by any other part. Figure 3.14(b) displays five instances of this relationship. Observe that several lines can emanate from a part instance and several lines terminate at a part instance.

3.9 *Binary Relationship*

A binary relationship, or an association between two entity types, is the most common form of a relationship expressed by an E-R diagram. Recall that a binary relationship is $R \in E_1 \times E_2$ in which E_1 and E_2 are two different entity types. The examples of binary relationships with the relationship instances are presented in Figure 3.15. Notice that each relationship instance obeys a basic characteristic of the binary relationships; in other words, each relationship instance is connected to exactly two entity instances of different entity types.

By applying what we have learned earlier in this chapter, we should be able to determine the cardinalities of these relationships quite easily. The relationship in Figure 3.15(a) is a one-to-one relationship since we assume that an employee can manage at most one department and each department is managed by at most one employee. The minimum cardinality can be determined



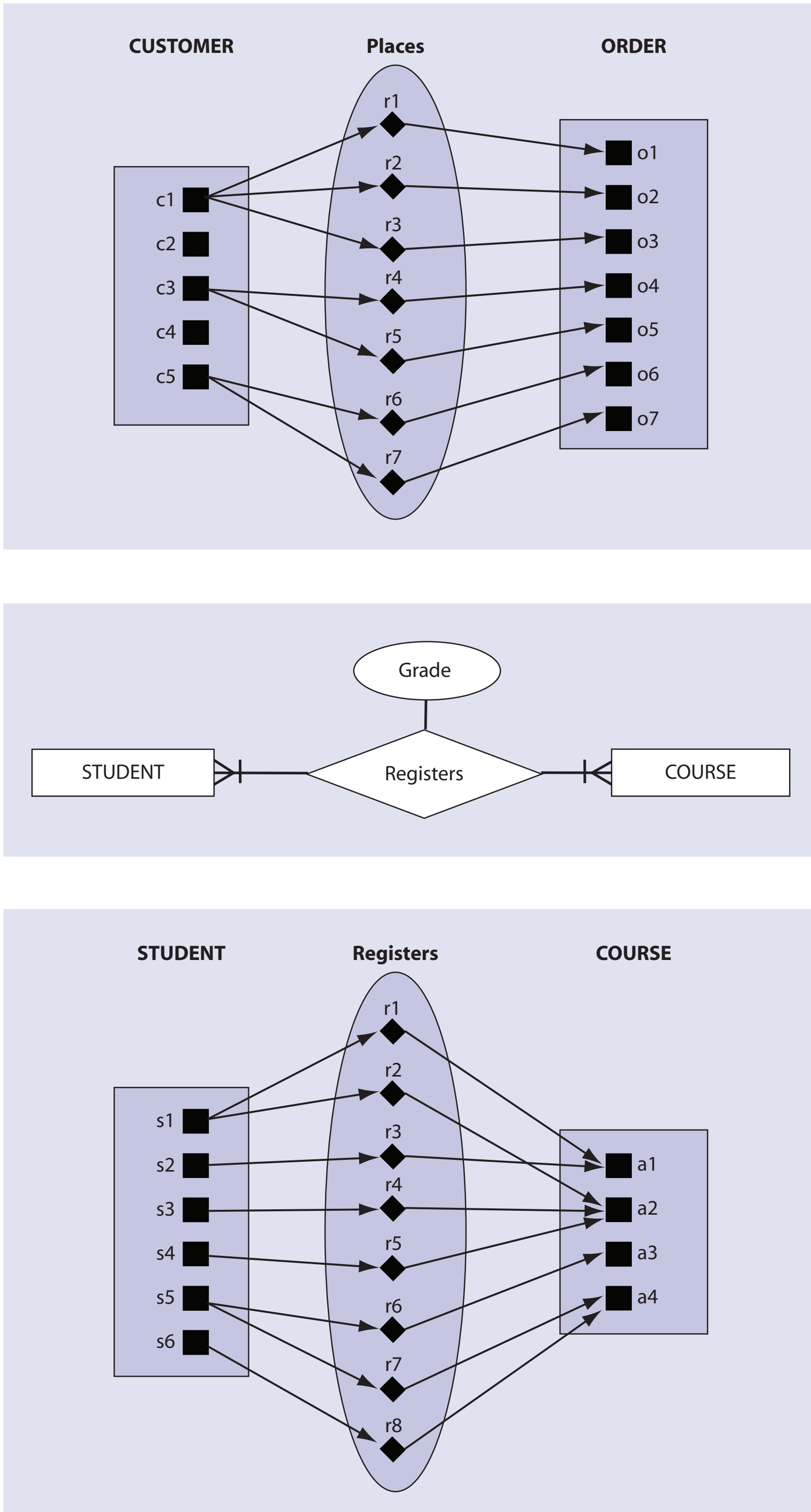
(a)

(b)

(c)

(*continues*)

Figure 3.15 (a) and (b): Binary one-to-one relationships; (c) and (d): Binary one-to-many relationships; (e) and (f): Binary many-many relationships.



(d) (e) (f)

Figure 3.15 (*continued*)

since each department must be managed by an employee, but not all employees manage departments. For example, while some of the employees in Figure 3.15(b), such as e2, e4, and e6, do not manage a department, every department is managed by an employee.

The binary one-to-many relationship represented in Figure 3.15(c) features a slightly different arrangement. While a customer can place several orders or may choose not to order at all, each order must be placed by exactly one customer. Figure 3.15(d) shows us that, while each order is made by a single customer, not all customers place orders. However, Figure 3.15(d) differs from Figure 3.15(b) in that a single customer may place any number of orders; while the first customer places three orders, and the fifth customer places only two, both transactions represent a maximum cardinality of many.

Finally, Figure 3.15(e) and (f) illustrates a many-to-many relationship, featuring a minimum cardinality of zero and a maximum cardinality of many. In other words, each student can participate in many activities, a single activity, or no activity at all. Conversely, any number of students, from many to none, can participate in a given activity.

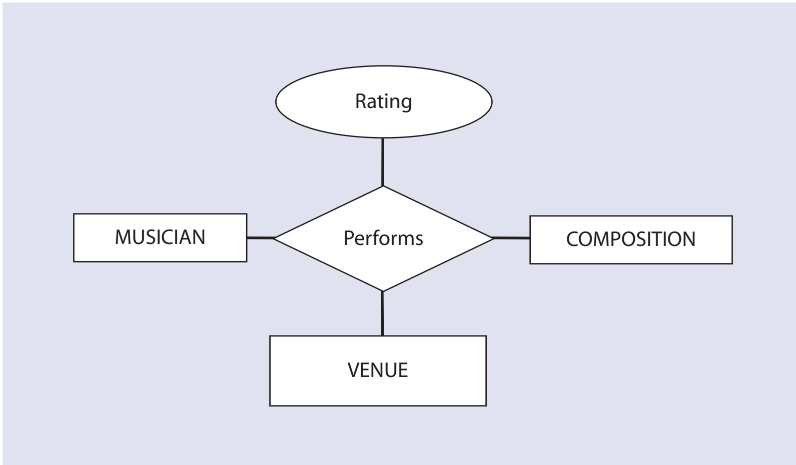
3.10 *Ternary Relationships*

Recall that a ternary relationship R is a relationship among instances of three different entity types, E_1, E_2, and E_3 ($R \in E_1 \times E_2 \times E_3$). Each instance of the ternary relationship R requires the participation of an instance from each of the entity types E_1, E_2, and E_3. See Figure 3.16 for examples of ternary relationships.

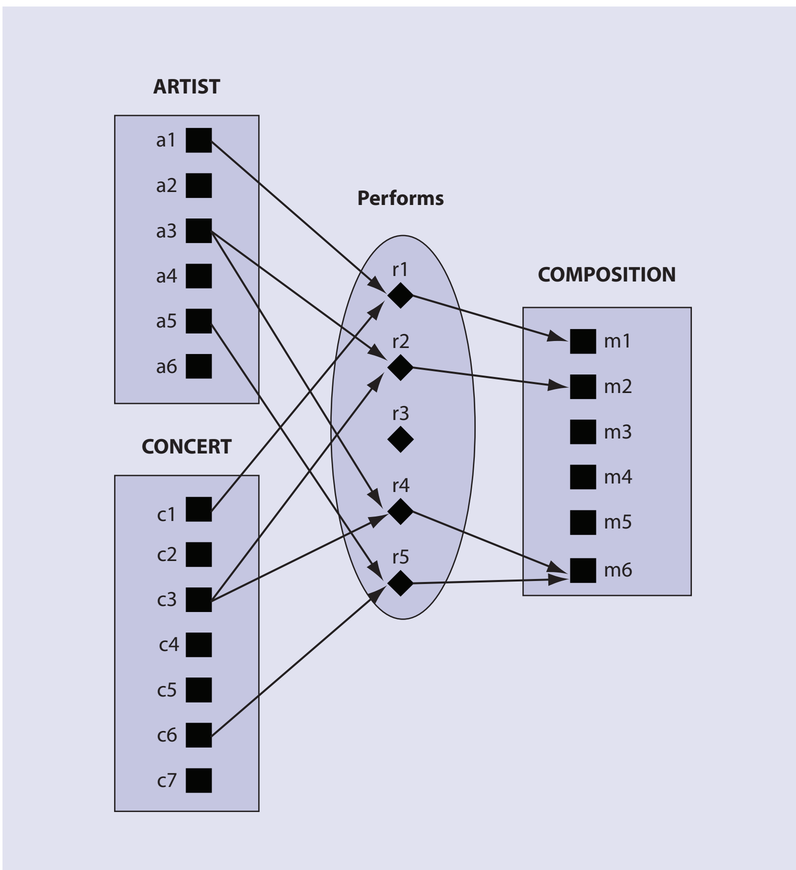
The examples in Figure 3.16(a) and (b) reveal the possible relationships associated with a competition for classical musicians. Let's suppose that at this competition, musicians perform individually and in small groups for judges who rate the performances. Each performance requires an artist, a composition, and a venue. Observe that each relationship instance, *Performs*, connects an entity instance of MUSICIAN, COMPOSITION, and VENUE. Each relationship instance also has an attribute, *Rating*, that stores the average rating of the performance by a panel of judges. Similarly, in Figure 3.16(c), students use equipment to work on projects; each instance of *Uses* involves an instance of STUDENT, PROJECT, and EQUIPMENT. If a student uses two pieces of equipment to work on a project, there are two instances of the relationship *Uses*. A campus lab may use the attribute in this ternary relationship, the *Date* of use, to log the equipment usage.

Ternary relationships differ significantly from the other kinds of relationships that we have examined so far. It is important to remember that a ternary relationship is not equivalent to two binary relationships. Suppose that we recognize this ternary relationship as a series of binary relationships, such as ARTIST-CONCERT, ARTIST-COMPOSITION, and CONCERT-COMPOSITION. These three binary relationships can store data about the artists who performed at different concerts, the compositions performed by artists, and the compositions performed at different concerts; however, the relationships cannot store the compositions performed by an artist at a particular concert. Therefore, three binary relationships cannot encapsulate the range of data stored by a single ternary relationship. A ternary relationship, though, can capture the data of three binary relationships.

Observe that the cardinalities of the relationships in Figure 3.16 are not expressed in the E-R diagrams. The cardinalities of a relationship are defined for a pair of entities; in a ternary relationship, there are three pairs of ternary relationships. We cannot express the cardinalities of these types of relationships in E-R diagrams. Instead, we can turn the ternary relationship into an associative entity, a process that we will discuss in Section 3.12. This technique allows us to demonstrate the cardinalities of the entities within the associative entity.



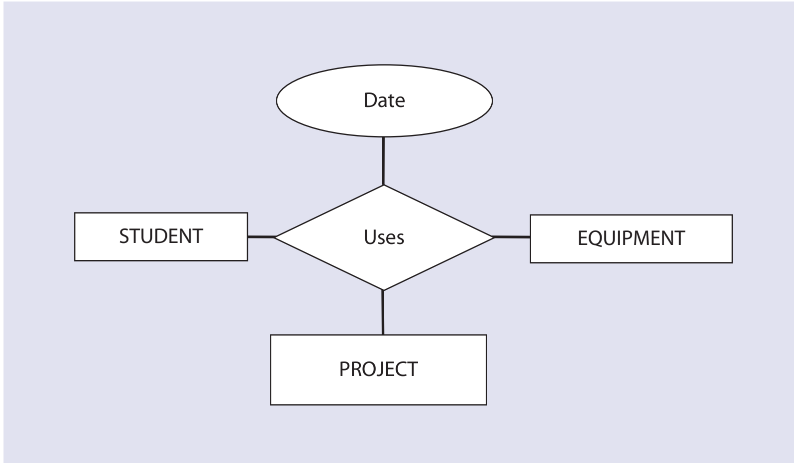
(a)



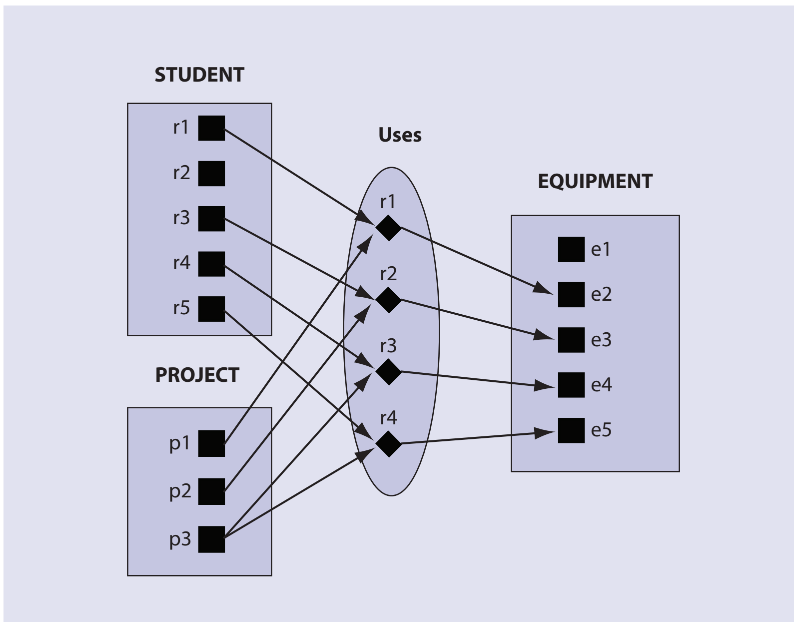
(b)

(continues)

Figure 3.16 Examples of ternary relationships.



(c)



(d)

Figure 3.16 (*continued*)

3.11 *Attributes of Relationships*

We have already discussed examples of E-R diagrams that reveal attributes stemming off relationships. Recall Figure 3.15(c), for example. In that E-R diagram, the attribute *DateofOrder* collects data for the relationship CUSTOMER *places* ORDER. Attributes on relationships are like attributes on entity types we have seen so far. An attribute on a relationship stores information related to the relationship. In Figure 3.17, the attribute *Quantity* stores the number of components that make up an entity type ITEM. Note that the attribute *Quantity* stems from the relationship and not from the entity.

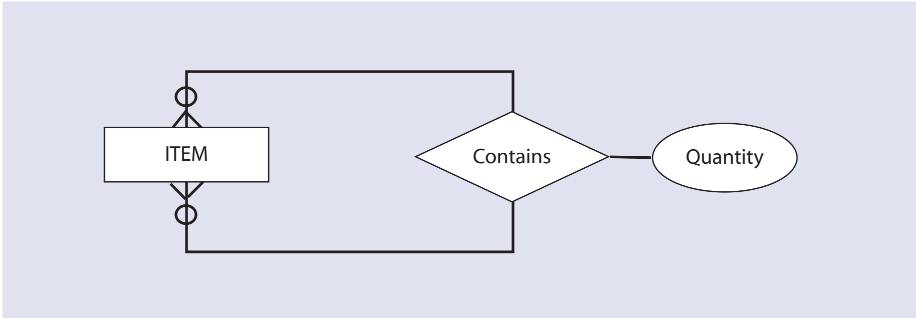


Figure 3.17 An example of attributes of relationships.

3.12 *Associative Entities*

An ***associative entity*** is an entity type that connects the instances of one or more entity types and contains attributes particular to this association. Basically, an associative entity is a relationship that has been turned into an entity because it meets one of the following conditions:

1. It is a many-to-many binary relationship
2. It is a ternary relationship or a relationship of an even higher degree

The associative entity in an E-R diagram is represented by an entity box enclosing the diamond relationship symbol (see Figure 3.18). This symbol demonstrates that the entity is generated from a relationship.

Consider, for example, the E-R diagram of the binary relationship illustrated in Figure 3.18(a). When we convert this relationship into an associated entity, we get the E-R diagram in Figure 3.18(b). In the example, the relationship *Participates* is converted into an associated entity, ENROLLMENT. If the relationship has attributes, they become the attributes of the corresponding associative entity. *DateJoined* is an example of an attribute of a relationship turned into an attribute of an associative entity in Figure 3.18(b).

Furthermore, recall that every entity in an E-R diagram must have an identifier. The identifiers of two original entities together serve as a composite identifier of the associative entity. In our example, the identifier of the entity ENROLLMENT is a composite attribute comprised of the identifiers of the STUDENT and ACTIVITY entities. Since a student can enroll in every activity but can only enroll once in any given activity, the composite attribute {*StudentID, ActivityID*} is a valid identifier for the associated entity ENROLLMENT.

Now that ENROLLMENT is a new entity in Figure 3.18(b), note that there is no relationship diamond on the line between the entity STUDENT and the associative entity ENROLLMENT because the associative entity represents a relationship. Furthermore, the E-R diagram in Figure 3.18 (b) depicts the cardinalities. Each instance of the STUDENT entity is related to several instances of the ENROLLMENT entity; in fact, each instance of the STUDENT entity is related to as many instances of the ENROLLMENT entity as the number of activities in which the student participates. Therefore, there is a one-to-many relationship between the STUDENT entity type and the ENROLLMENT entity type. Similarly, there is a one-to-many relationship between the entity types ACTIVITY and ENROLLMENT because several students may enroll in one activity.

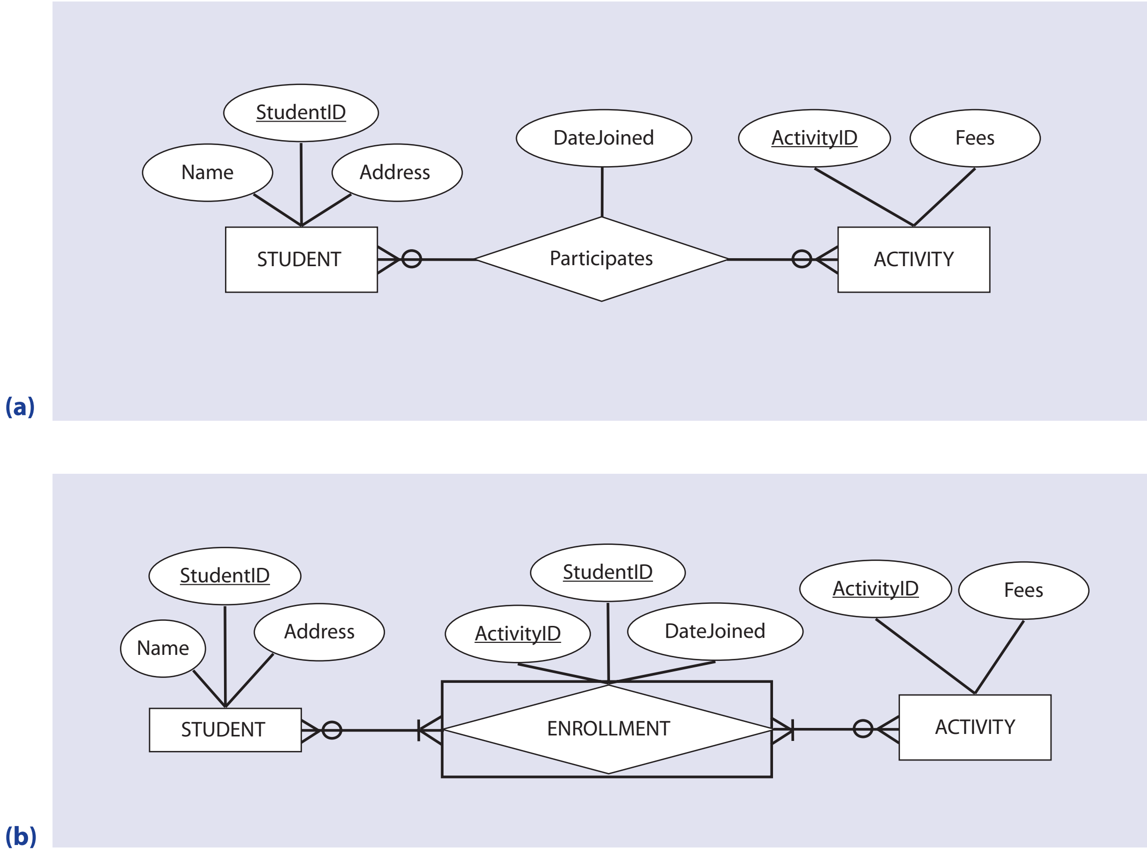
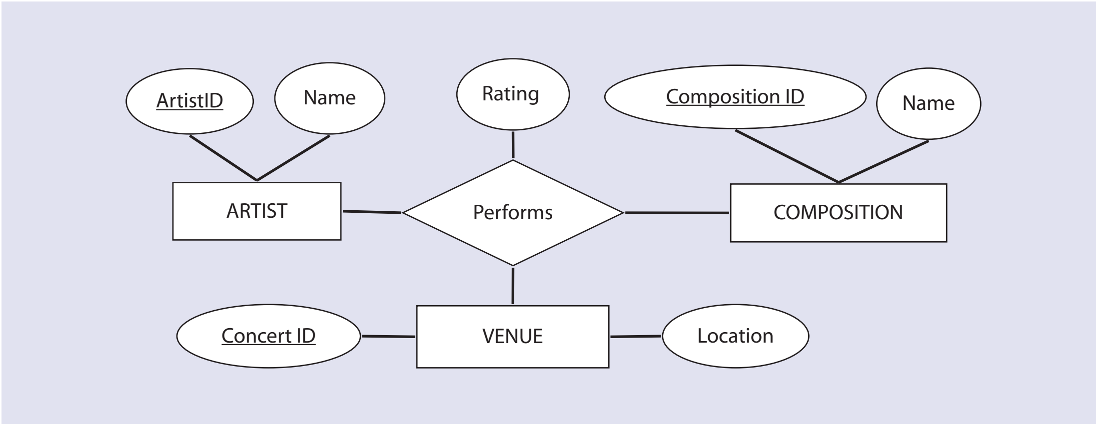


Figure 3.18 (a) A many-to-many binary relationship; (b) A many-to-many binary relationship converted to an associated entity.

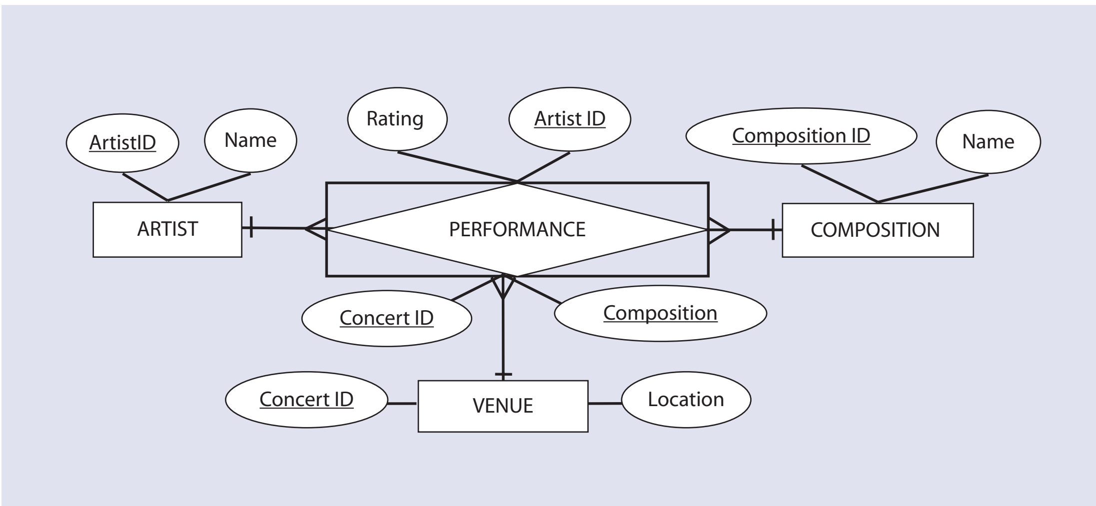
Notice that when we convert a relationship into an associative entity, we change its name from a verb to a noun. Therefore, the relationship expressed by the verb participates in Figure 3.18(a) becomes the associative entity represented by the noun enrollment in Figure 3.18(b). This grammatical distinction parallels the main reason for turning a relationship into an associative entity. A verb expresses an action, a process. However, once we want to count, record, or analyze an instance of that process, we end up with a person, place, or thing. In short, by transforming a relationship into an associative entity, we can analyze and record a process in the same way that we analyze and record a noun.

In Chapter 4, we make it clear why we can only turn a many-to-many relationship into an associative entity. We convert a ternary relationship into an associative entity as it allows us to express the cardinalities of a relationship correctly. We will now consider another example of associative entity converted from a ternary relationship.

The E-R diagram shown in Figure 3.19 is similar to the one we first observed in Figure 3.16. Here, we convert the ternary relationship instance *Performs* into the associative entity PERFORMANCE. We change the verb performs, an action not necessarily related to a specific event, to performance, a noun that expresses an event that can be recorded, counted, and analyzed. Observe that the identifier of the associated entity PERFORMANCE is comprised of the union of the identifiers of the entities involved in the ternary relationship. One of the most important



(a)



(b)

Figure 3.19 (a) A ternary relationship; (b) A ternary relationship converting to an associated entity.

characteristics of this arrangement is that each entity, ARTIST, VENUE, and COMPOSITION, has a one-to-many relationship (minimum cardinalities are omitted for clarity) with the associated entity PERFORMANCE.

3.13 *Weak Entity Types*

Entity types can be classified into two categories: *strong entity types* and *weak entity types*. A ***strong entity type*** exists independent of other entity types, while a ***weak entity type*** depends on another entity type. Consider a student database that includes an entity STUDENT. Suppose that we also record data about each student's dependents, such as a spouse or children, in this database. To do so, we must create the entity type DEPENDENT. The entity type DEPENDENT does not exist on its own and owes its existence to the STUDENT entity type. When students graduate and their records are removed from the STUDENT entity set, the records of their dependents are also removed from the DEPENDENT entity set. In the E-R diagram, a weak entity is indicated by a double-lined rectangle. The corresponding relationship diamond is also double-lined.

The entity type on which a weak entity type depends is called the *identifying owner* (or simply owner), and the relationship between a weak entity type and its owner is called an *identifying relationship*. In Figure 3.20, STUDENT is the owner and *Has* is the identifying relationship. For a weak entity, we define a *partial key attribute* that, when combined with the key attribute of its owner, provides the full identifier for the weak entity. In Figure 3.20, for example, *DependentName* is the partial identifying attribute. When we combine it with the *StudentID*, it uniquely identifies the dependent. Of course, in this example, we make an implicit assumption that people do not give the same name to more than one of their children.

Definition

An **associative entity** is an entity type that connects the instances of one or more entity types and contains attributes particular to this association.

A **strong entity type** exists independent of other entity types, while a **weak entity type** depends on another entity type.

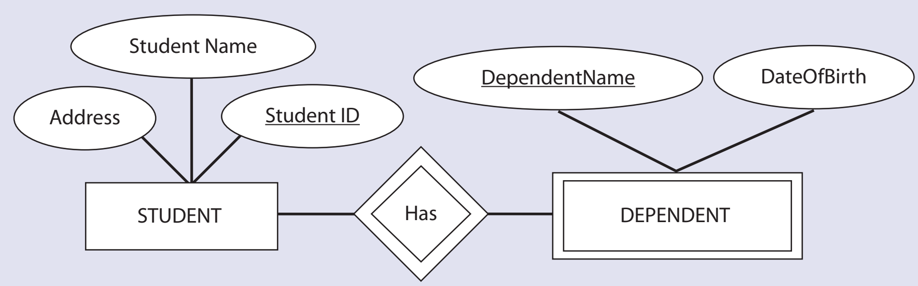


Figure 3.20 The weak entity in an E-R diagram.

3.14 *Enhanced Entity-Relationship Modeling*

We have illustrated the basic concepts of the entity-relationship model. These concepts are generally adequate to design solutions for common business problems. However, since the early 1980s, business database applications and database requirements have changed dramatically. Business relationships and business data have become much more complex as applications such as multimedia systems, geographical information systems, and computer-aided engineering have gained prominence. These applications demand more complex database requirements than traditional applications, rendering the basic concepts of E-R modeling inadequate. As a result, researchers in the '80s enhanced the original E-R model to form the *enhanced entity-relationship (EE-R) model*.

In the remaining sections of this chapter, we will discuss the EE-R model by detailing superclass and subclass entities, the processes of specialization and generalization, and various constraints on these processes.

3.15 *Superclass, Subclass, and Relationships*

Recall that we have defined an *entity type* as a collection of entities that share the same attributes. In this section, we will introduce two special types of entities: superclass and subclass. We will also discuss the notations that allow us to include these special entities in a regular E-R diagram to create an enhanced entity-relationship (EE-R) diagram. We will also study superclass and subclass relationships as well as the process of attribute inheritance associated with superclass and subclass entity types.

3.15.1 Superclass and Subclass

All the instances of an entity type share the same set of attributes, but each single attribute may not be a required attribute of for each instance. For example, consider the PERSON entity type. The entity PERSON includes all personnel in a university. Specifically, it includes students, staff, and faculty members. The attributes of this entity, *SSN*, *Address*, *Email*, *Salary*, *Class*, *GPA*, and *Office Phone* are depicted in Figure 3.21. Although all these attributes are common to the PERSON entity, attributes such as *Class* and *GPA* are not required for faculty instances of PERSON. At the same time, the attribute *Salary* is required for staff and faculty but may not be relevant for students.

This distinction is important when we have a large number of instances for an entity type. In such a scenario, one or more attributes for each instance won't have any valid value. In our example, each faculty instance will have no valid value for *GPA* and *Class*, and many student instances won't have value for *Salary*, *Rank*, and *Designation*. As we will see in Chapter 4, having many redundant fields results in a data redundancy problem and degrades the database performance.

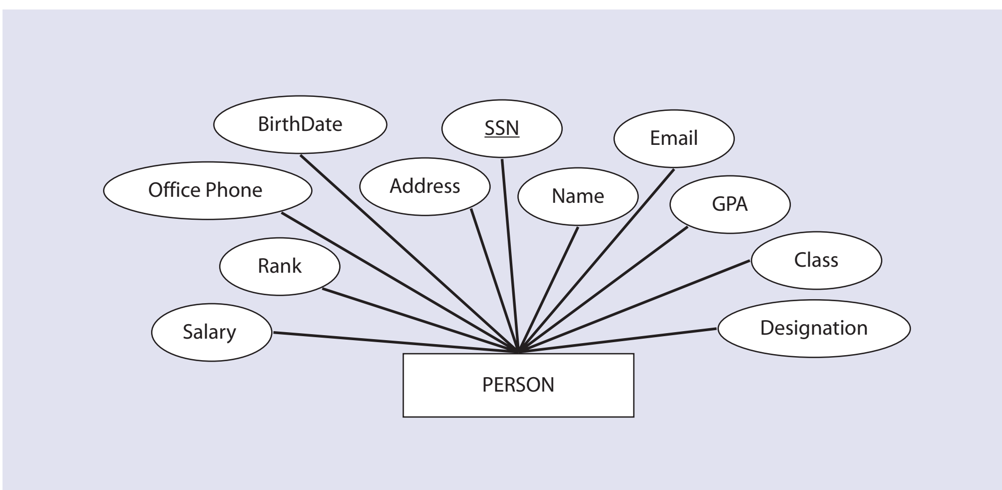


Figure 3.21 The E-R diagram for entity type PERSON.

One possible solution to this problem is to extend the E-R model to accurately represent the complex groupings that we encounter in everyday life. Let us begin our discussion by formally defining superclass and subclass entity types.

A ***superclass*** is an entity type that has one or more distinct subgroups with unique attributes. For example, the entity type PERSON in Figure 3.22 is a superclass that includes faculty, staff, and students as its subgroups. The superclass features only those attributes that are common for all its subgroups. For example, attributes of PERSON such as *SSN*, *Name*, *Address*, and *Email* are shared by all its subgroups regardless of an individual's position as student, faculty, or staff within the university. The subgroups with unique attributes are defined as ***subclasses***. The PERSON superclass thus has three subclasses: STUDENT, STAFF, and FACULTY. A subclass entity type STUDENT has attributes of its superclass along with its own attributes such as *Major*, *GPA*, and *Class* that uniquely identify the subclass. Figure 3.22 depicts a superclass and subclasses.

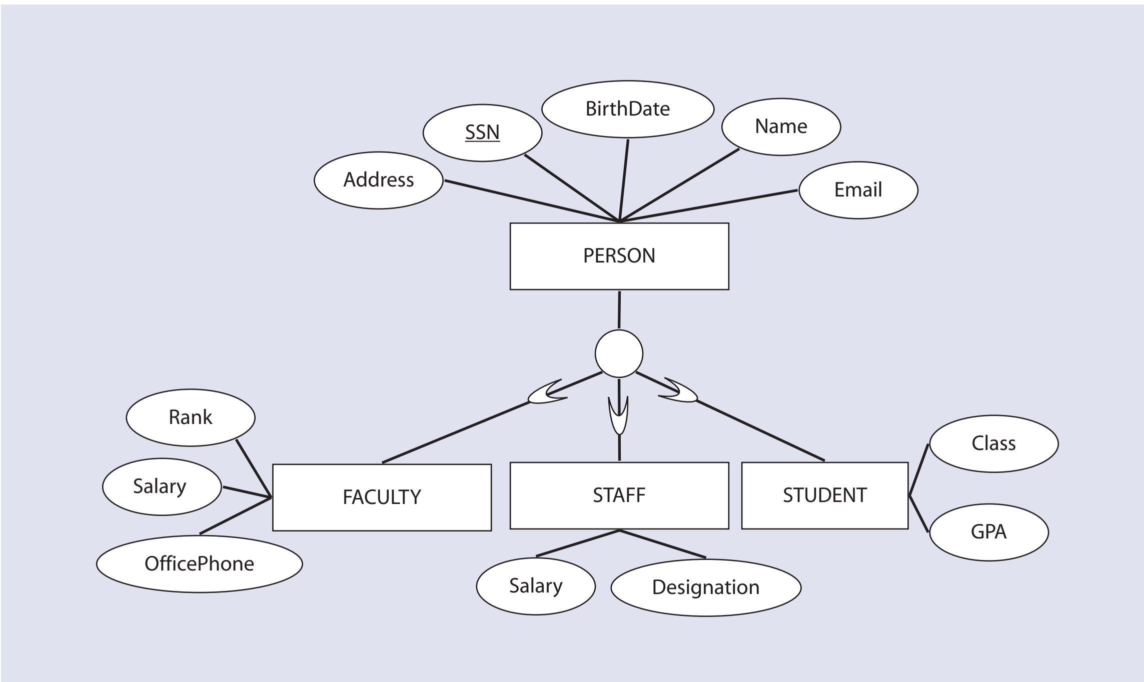


Figure 3.22 The enhanced E-R diagram for entity type PERSON.

Let us examine the formal notations used to represent a superclass and subclass. We represent both superclass and subclass entity types with the usual rectangle notation of an entity type. A superclass is connected to a circle by a short line. This circle, in turn, is connected to each of the subclass entity types with another line. An optional U-shaped symbol on each line that connects subclass entity types to the circle indicates that a subclass is a subset of a superclass entity type. This U-shaped symbol also indicates the direction of the superclass/subclass relationship. In that sense, it operates like an arrow pointing from the superclass to a subclass. Other notations of basic E-R diagrams prevail (see Figure 3.23 for general layout).

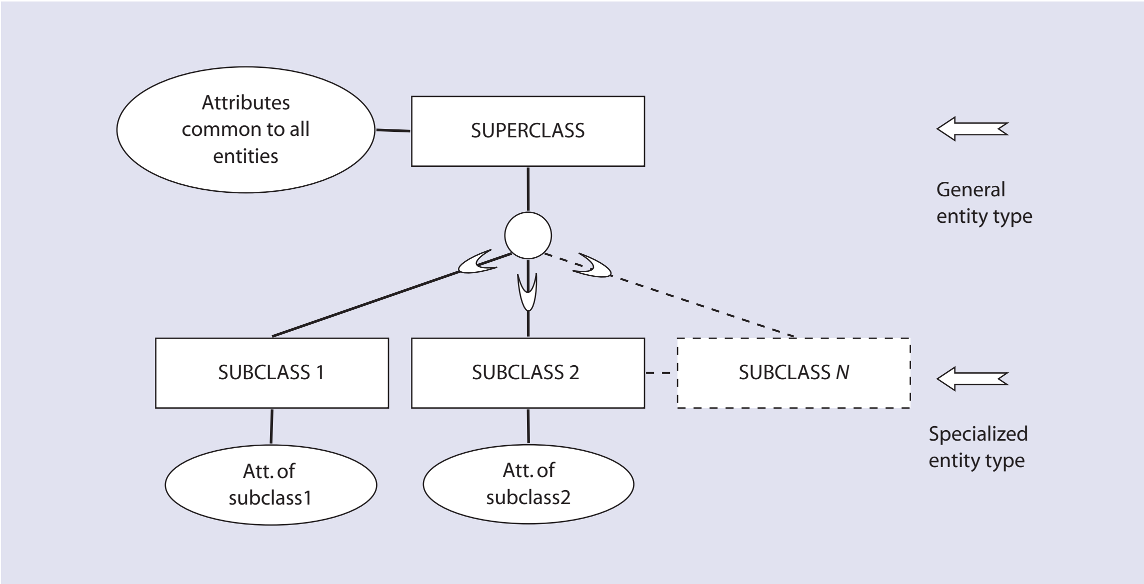


Figure 3.23 The notations for superclass and subclass relationships.

There are several reasons for introducing superclass and subclasses into an E-R model. First, incorporating a superclass and subclasses maintains the cleanliness of the structure and improves the conceptual flow. Second, it adds familiar semantic information to an E-R diagram. We consider using subtypes if the following conditions apply:

1. Certain attributes apply to some, but not all, instances of an entity type.
2. An instance of a subtype participates in a relationship unique to that relationship (see Figure 3.24).

Definition

The **Enhanced Entity-Relationships (EE-R) model** is a revised E-R model that extends the original E-R model and supports additional semantic concepts by providing new modeling constructs.

A **superclass** is an entity type that has one or more distinct sub groups with unique attributes.

A **subclass** is an entity type that shares common attributes or relationships distinct from other subclasses.

3.15.2 Attribute Inheritance and Subclass Relationships

Attribute Inheritance is the property by which subclass entities inherit attributes of the superclass. The subclass entity type is an entity type in itself with a set of attributes and relationships. In addition to its own attributes, a subclass entity inherits all the attributes of its superclass. For example, the STUDENT subclass has attributes such as *MajorDept*, *Class*, and *GPA*. It also inherits all the attributes of the PERSON entity type. In other words, if Chris Alto is an instance of the STUDENT subclass, then he is necessarily an instance of the PERSON superclass as well. Chris Alto is a value for the *Name* attribute for the entity PERSON, and the STUDENT subclass inherits it from the PERSON. Note that Chris's 4.0 *GPA* is an attribute of the STUDENT subclass only. Thus, the identifier of a superclass is also inherited by all its subclasses and serves as an identifier for subclasses. Subclasses don't need their own identifier. Thus, the relationship between a superclass and subclass is always a one-to-one relationship (notations are usually ignored in an E-R diagram).

Although, each subclass also possesses attributes of its superclass, each subclass entity expresses a distinct role in an E-R diagram. It can have its own attributes and relationships that distinguish one subclass from another. Figure 3.24 depicts relationships for the STUDENT and FACULTY subclasses.

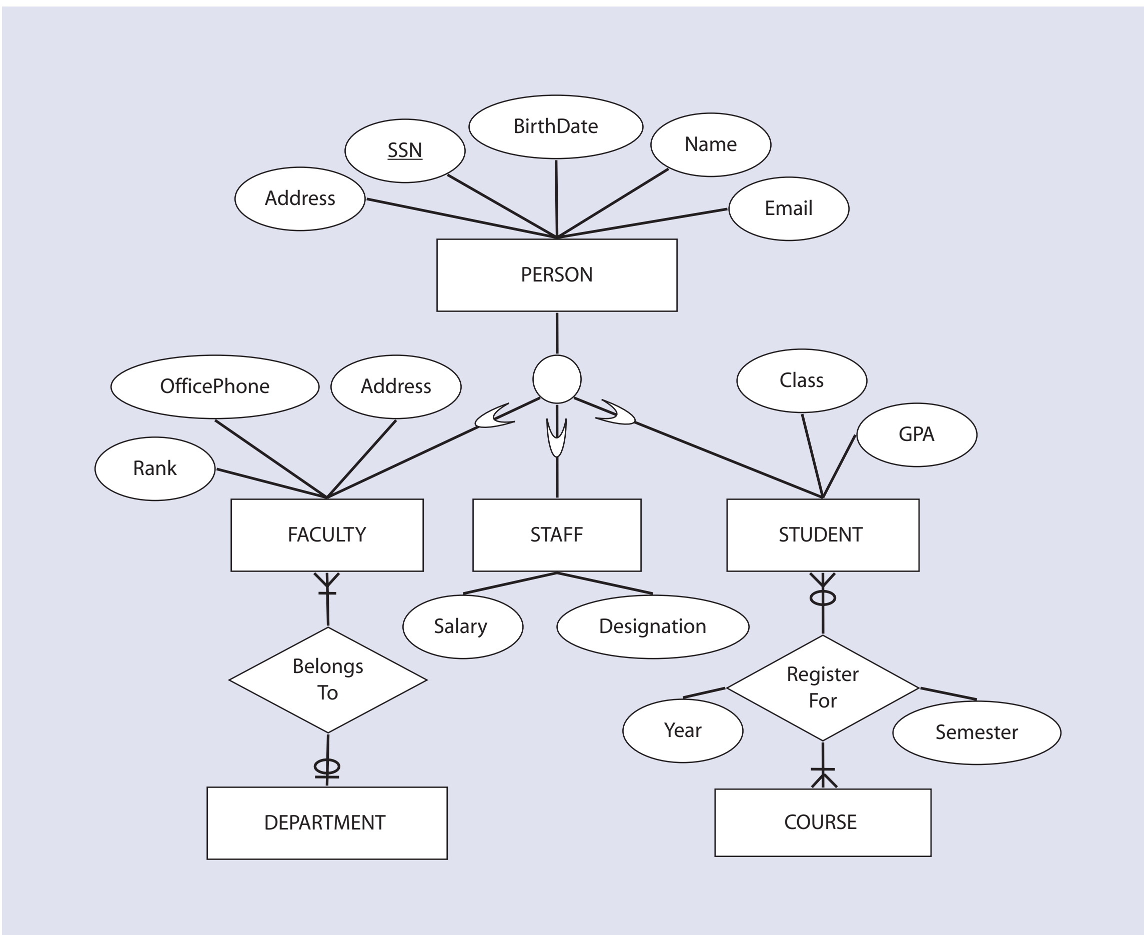


Figure 3.24 An example of the superclass/subclass relationships.

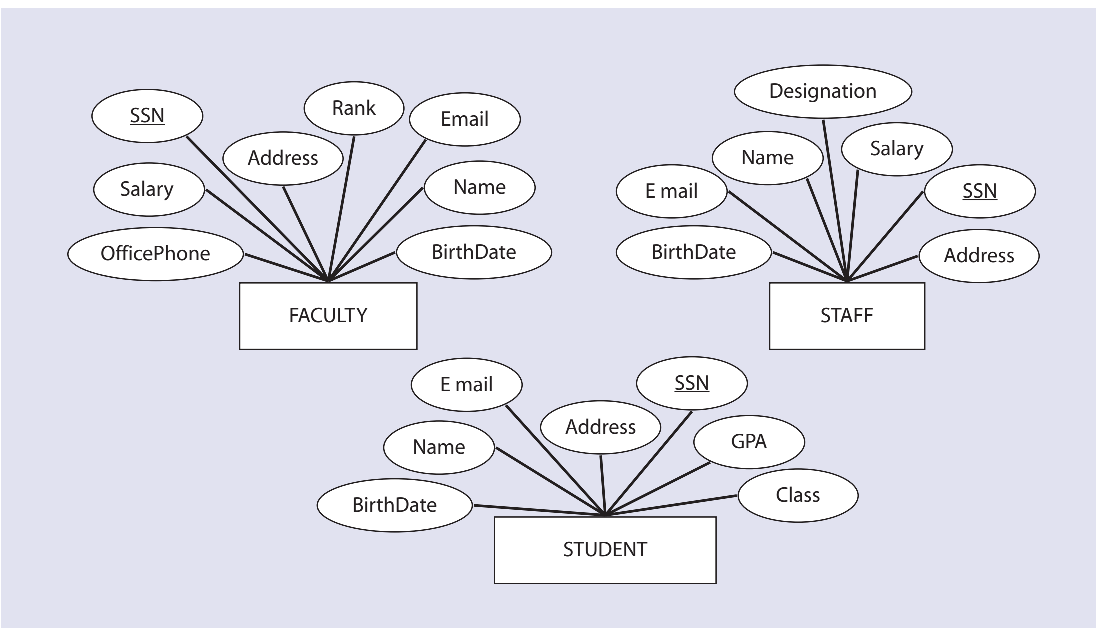
3.16 *Generalization and Specialization Process*

In the previous section we described the superclass/subclass model. However, in the process of modeling a complex real-word scenario, it is important to recognize when superclass and subclass entity types can be generated. The processes of *specialization* and *generalization* are used to find such opportunities. These processes serve as conceptual models for the development of superclass/subclass relationships.

3.16.1 Generalization

Generalization is the process of defining general entity types from a set of specialized entity types by identifying their common characteristics. In other words, this process minimizes the differences between entities by *identifying a general entity type* that features the common attributes of specialized entities. Generalization is a bottom-up approach as it starts with the specialized entity types (subclasses) and forms a generalized entity type (superclass).

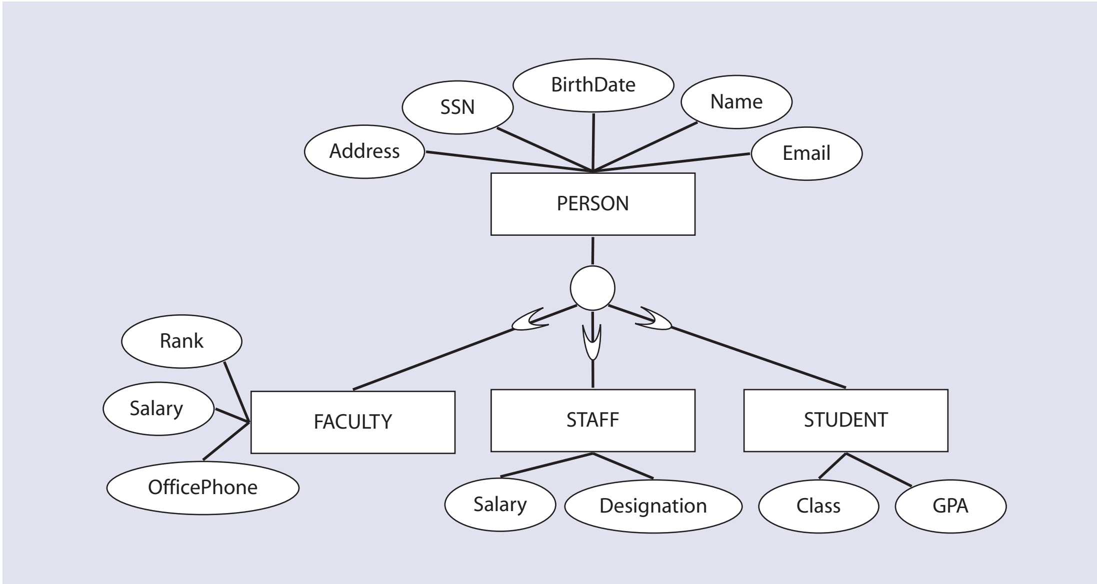
For example, suppose that someone has given us the specialized entity types FACULTY, STAFF, and STUDENT, and we want to represent these entity types separately in the E-R model as depicted in Figure 3.25(a). However, if we examine them closely, we can observe that a number of attributes are common to all entity types, while others are specific to a particular entity. For example, FACULTY, STAFF, and STUDENT all share the attributes *Name*, *SSN*, *Birth Date*, *Address*, and *Email*. On the other hand, attributes such as *GPA*, *Class*, and *MajorDept* are specific to the STUDENTS; *OfficePhone* is specific to FACULTY, and *Designation* is specific to STAFF. Common attributes suggest that each of these three entity types is a form of a more gen-



(a)

(*continues*)

Figure 3.25 (a) STAFF, FACULTY, and STUDENT entities before generalization; (b) PERSON superclass and FACULTY, STAFF, and STUDENT subclasses after generalization.



(b)

Figure 3.25 (*continued*)

eral entity type. This general entity type is simply a PERSON superclass entity with common attributes of three subclasses (see Figure 3.25(b)).

Thus, in the generalization process, we group specialized entity types to form one general entity type and identify common attributes of specialized entities as attributes of a general entity type. The general entity type is a superclass of specialized entity types or subclasses.

3.16.2 Specialization

Specialization is the process of defining one or more subclasses of a superclass by identifying its distinguishing characteristics. Unlike generalization, specialization is thus a top-down approach. It starts with the general entity (superclass) and forms specialized entity types (subclasses) based on specialized attributes or relationships specific to a subclass.

For example, consider Figure 3.26(a). LIBRARY ITEM is an entity type with several attributes such as *IdentificationNo*, *RecordingDate*, *Frequency*, and *Edition*. After careful review of these items, it should become clear that some items such as books do not have values for attributes such as *Frequency*, *RecordingDate*, and *CourseNo*, while Video CDs do not have an *Author* or an *Edition*.

In addition, all items have common attributes such as *IdentificationNo*, *Location*, and *Subject*. Someone creating a library database, then, could use the specialization process to identify superclass and subclass relationships. In this case, the original entity LIBRARY ITEM forms a superclass entity type made up of attributes shared by all items, while specialized items with distinguishing attributes, such as BOOK, JOURNALS, and VIDEOCD, form subclasses.

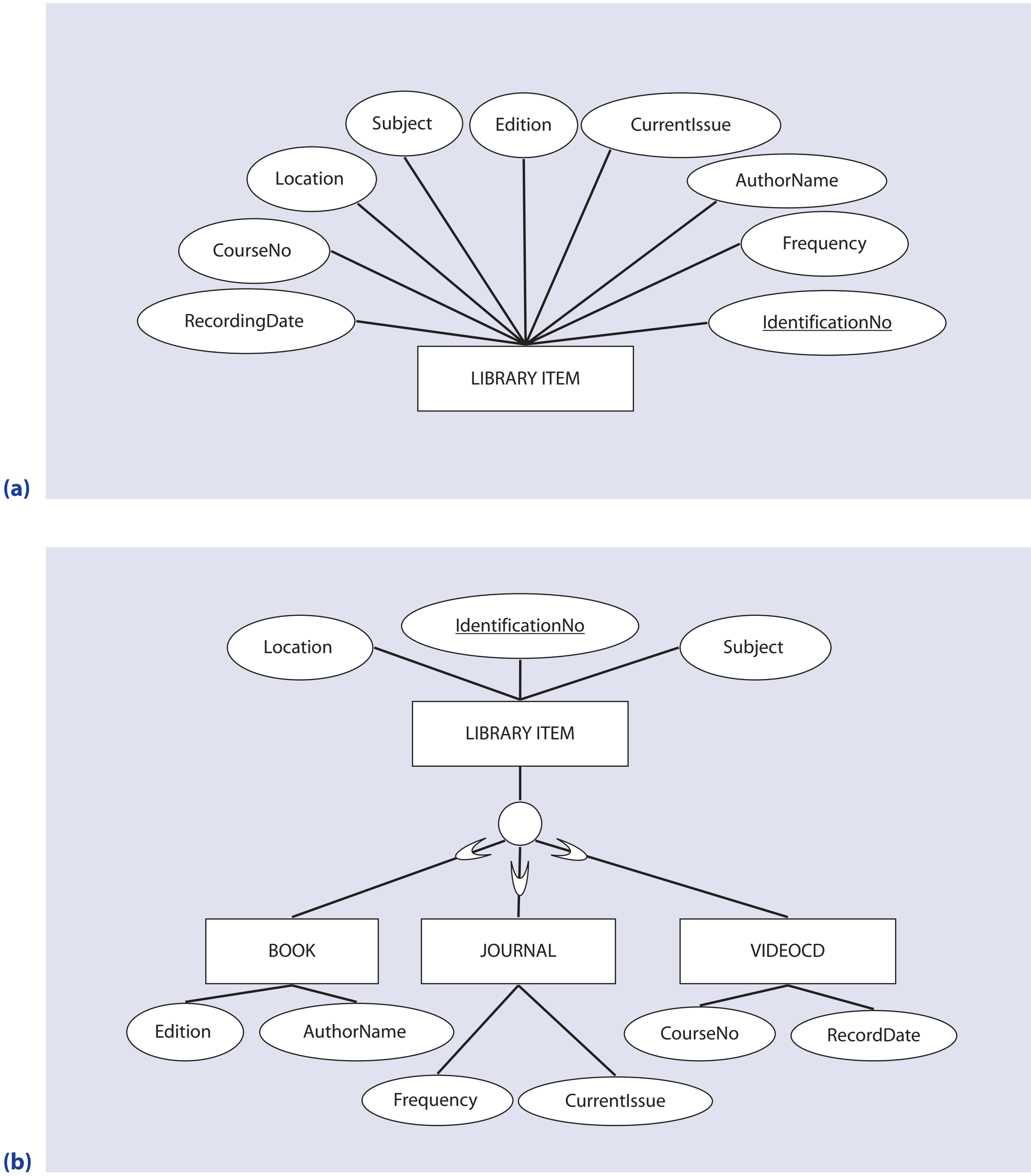


Figure 3.26 (a) LIBRARY ITEM entity before specialization; (b) LIBRARY ITEM superclass and BOOK, JOURNAL, and VIDEOCD subclasses after specialization.

Definition

Attribute inheritance is the property by which subclass entities inherit values for all attributes of the superclass.

Generalization is the process of defining a general entity type from a set of specialized entity types by identifying their common characteristics.

Specialization is a process of defining one or more subclasses of a superclass by identifying their distinguishing characteristics.

3.17 *Participation and Disjoint Constraints*

So far we have discussed superclass/subclass entity types, their relationships, and two processes to identify them. In this section, we will discuss constraints on superclass/subclass relationships. Specifically, we will introduce a *participation* constraint and a *disjoint* constraint. These constraints are intuitive and help us manifest business rules and incorporate them into the design of an EE-R.

3.17.1 Participation Constraints

Participation constraints dictate whether each instance (member) of a superclass must participate as an instance (member) of a subclass. A participation of superclass instance may be mandatory or optional in one or more subclasses. The mandatory constraint is also known as a total participation (constraint) or total specialization rule, while an optional constraint is known as a partial participation (constraint) or partial specialization rule.

Total Participation Rule In total participation, membership is mandatory. Each instance of a superclass must be an instance of at least one subclass. For example, consider Figure 3.27.

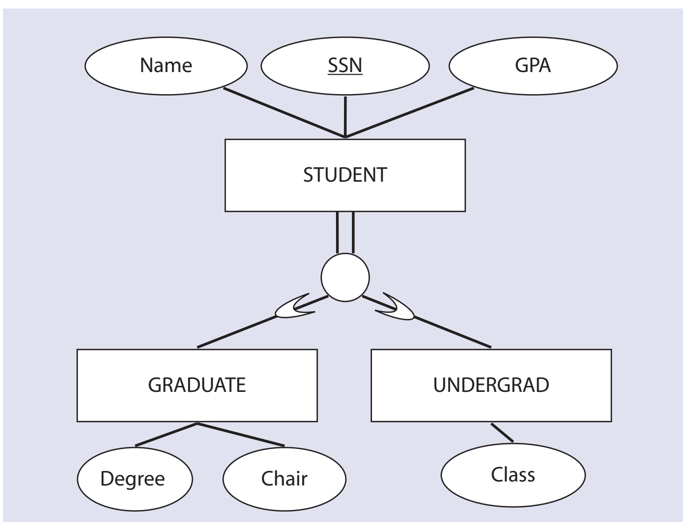


Figure 3.27 An example of the total participation rule.

Every instance of the superclass STUDENT must be an instance of either the GRADUATE student or UNDERGRAD student subclass. That is, if John Doe is an instance of a STUDENT, then John must be a graduate or undergraduate student. However, whether John Doe belongs to the graduate, undergraduate, or both entity types is answered by the disjoint rule, which we will define in a moment. We use a double line between the superclass entity type and the circle to represent the total participation.

Partial Participation Rule Membership is optional in a partial participation. An instance of a superclass does not have to be an instance of any of the subclasses. For example, consider Figure 3.26. An instance of the LIBRARY ITEM superclass can be a member of BOOK, VIDEO CD, or JOURNALS; however it is not mandatory for an instance to belong to any of these subclasses. If the library item *Newspaper* is an instance of a superclass, it does not have to be included in one of the subclasses; it can stay at the superclass level without having values for any subclass attributes. We use a single line between the superclass entity type and the circle (the default notation) to represent partial participation.

3.17.2 Disjoint Constraints

Disjoint constraints define whether it is possible for an instance of a superclass to simultaneously be a member of one or more subclasses. Disjoint constraints indicate whether a superclass instance can be disjointed or overlap more than one subclass.

Disjoint Rule The disjoint rule states that if an instance of a superclass is a member of any subclass, then it cannot be a member of more than one subtype (note that the participation rule will dictate whether the instance is a member of a subclass or not). For example, consider Figure 3.28. We put a constraint of disjoint rule to indicate that a student must be either a graduate or an undergraduate student but cannot belong to both subclasses simultaneously. We indicate the disjoint rule by putting a letter "D" in the joining circle of the superclass/subclass relationship (see Figure 3.28).

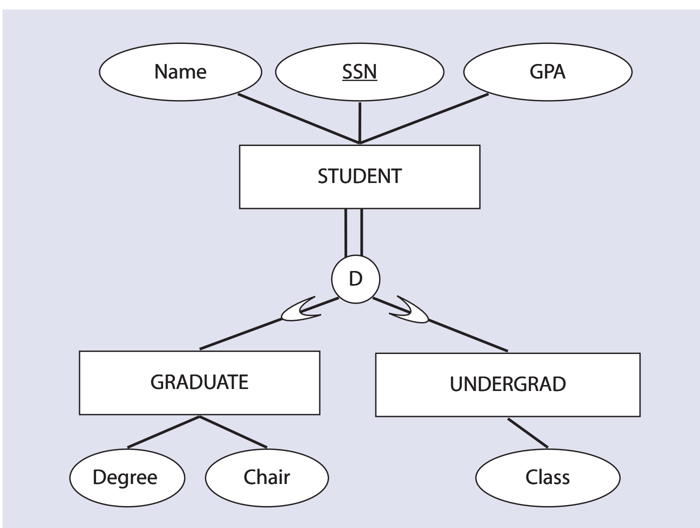


Figure 3.28 An example of the disjoint rule.

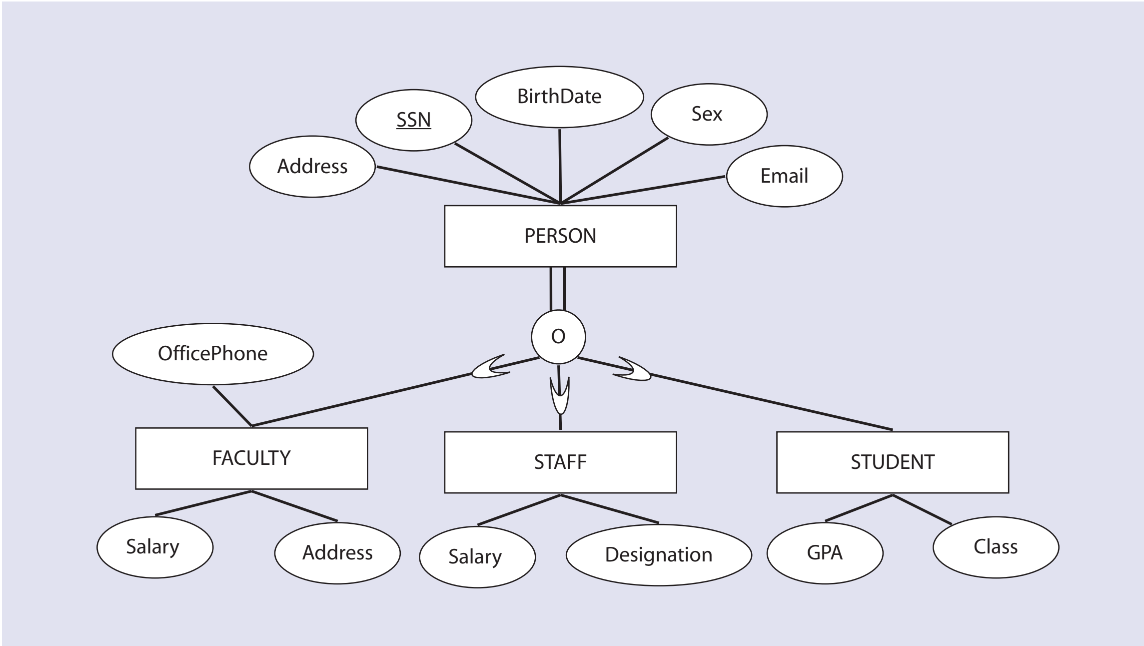


Figure 3.29 An example of the overlap rule.

Overlap Rule The overlap rule states that if an instance of a superclass is a member of any subclass, then it can be a member (overlap) of more than one subtype. For example, consider Figure 3.29. It reveals that an instance of a PERSON must be FACULTY, STAFF, a STUDENT, or any combination of these three. A work-study person would overlap, for example, the STAFF and STUDENT entities. We indicate the overlap rule by putting a letter "O" in the joining circle of the superclass/subclass relationship.

3.17.3 Subclass Discriminators

The participation and disjoint constraints allow an instance of a superclass to be a member of one or more subclasses. We keep track of membership using a special attribute of a superclass, a ***subclass discriminator***. In other words, we put a discriminating attribute on a superclass that can record membership of superclass instances into one or more subclasses. Figure 3.30 provides an example of total participation with the disjoint rule. An instance of STUDENT item must be a member of exactly one subclass. For every instance of a superclass when it is assigned to either subclass, we record the membership in the *StudentType* attribute. For example, if John Doe is a graduate student, we assign a value, say "G," to *StudentType*. If Monica Adams is an undergraduate student, we assign *StudentType* "U," and so on.

There are two special cases with partial participation and overlap rules using a discriminating attribute. If the participation is partial and an instance of a superclass does not participate in any subclasses, we assign NULL value to the discriminating attribute. If an overlap rule is in effect and if an instance of a superclass participates in more than one subclass, we assign multiple values to the discriminating attribute. Thus, the discriminating attribute becomes a multi-valued attribute under the overlap rule.

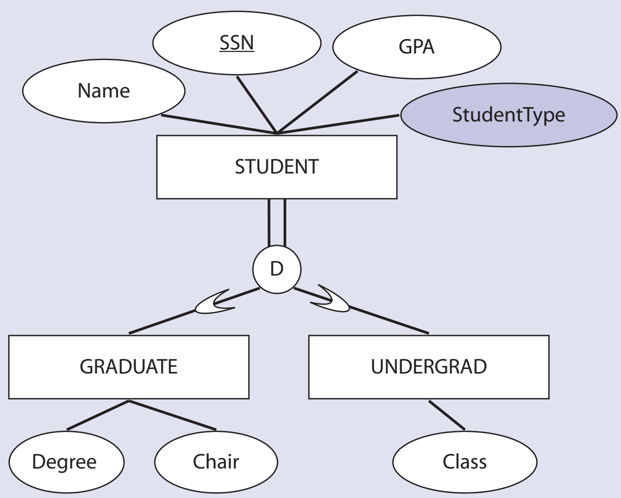


Figure 3.30 *StudentType* subclass discriminator.

Definition

Participation constraints dictate whether each instance (member) of a superclass must participate as an instance (member) of a subclass.

Disjoint constraints define whether it is possible for an instance of a superclass to simultaneously be a member of one or more subclasses.

A superclass/subclass hierarchy is a hierarchical structure of a superclass and its various subclasses in which each subclass has exactly one superclass.

3.18 *Superclass/Subclass Hierarchy*

A subclass type can have one or more subclasses. In other words, a subclass can act as superclass for other entity types, resulting in a hierarchy of superclass and subclasses. A ***superclass/subclass hierarchy*** is a hierarchical structure of a superclass and its various subclasses in which each subclass has exactly one superclass. For example, we have discussed PERSON and STUDENT superclass entity types. We put these two examples together in Figure 3.31 to create a hierarchical structure.

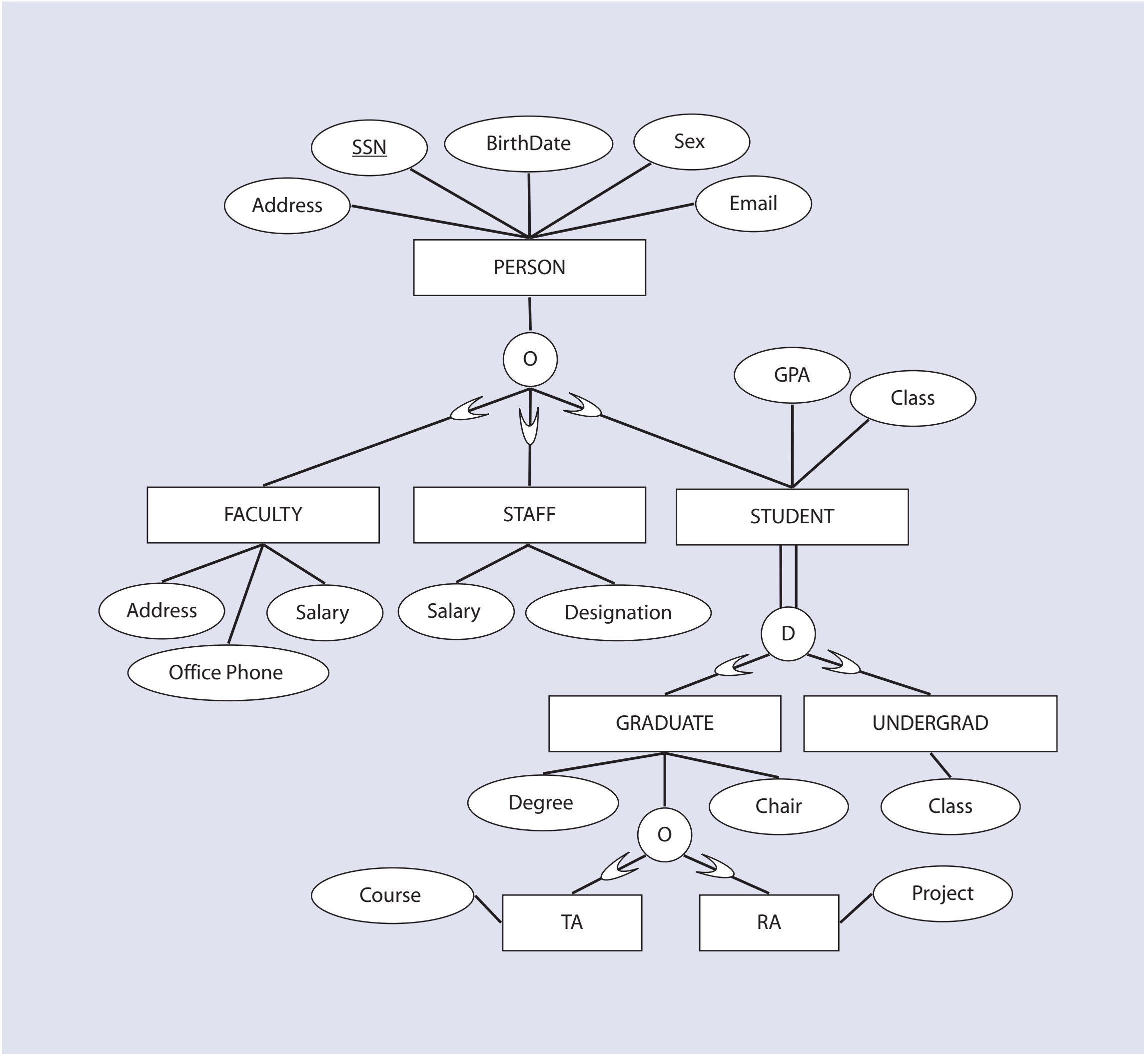


Figure 3.31 An example of a superclass/subclass hierarchy.

PERSON is the superclass with FACULTY, STAFF, and STUDENT as its subclasses. The relationship is partial participation with overlap rules. The STUDENT subclass has its own subclasses: GRADUATE student and UNDERGRAD student. Further, the GRADUATE subclass acts as a superclass for TA and RA subclasses; a graduate student can be a teaching assistant or a research assistant or both. Thus each level of a hierarchy can have its own set of entity types and relationships to depict the organizational requirements.

3.19 *Case Study: Conceptual Design for University Database*

In this section, we illustrate the conceptual design for a university database in Figure 3.32

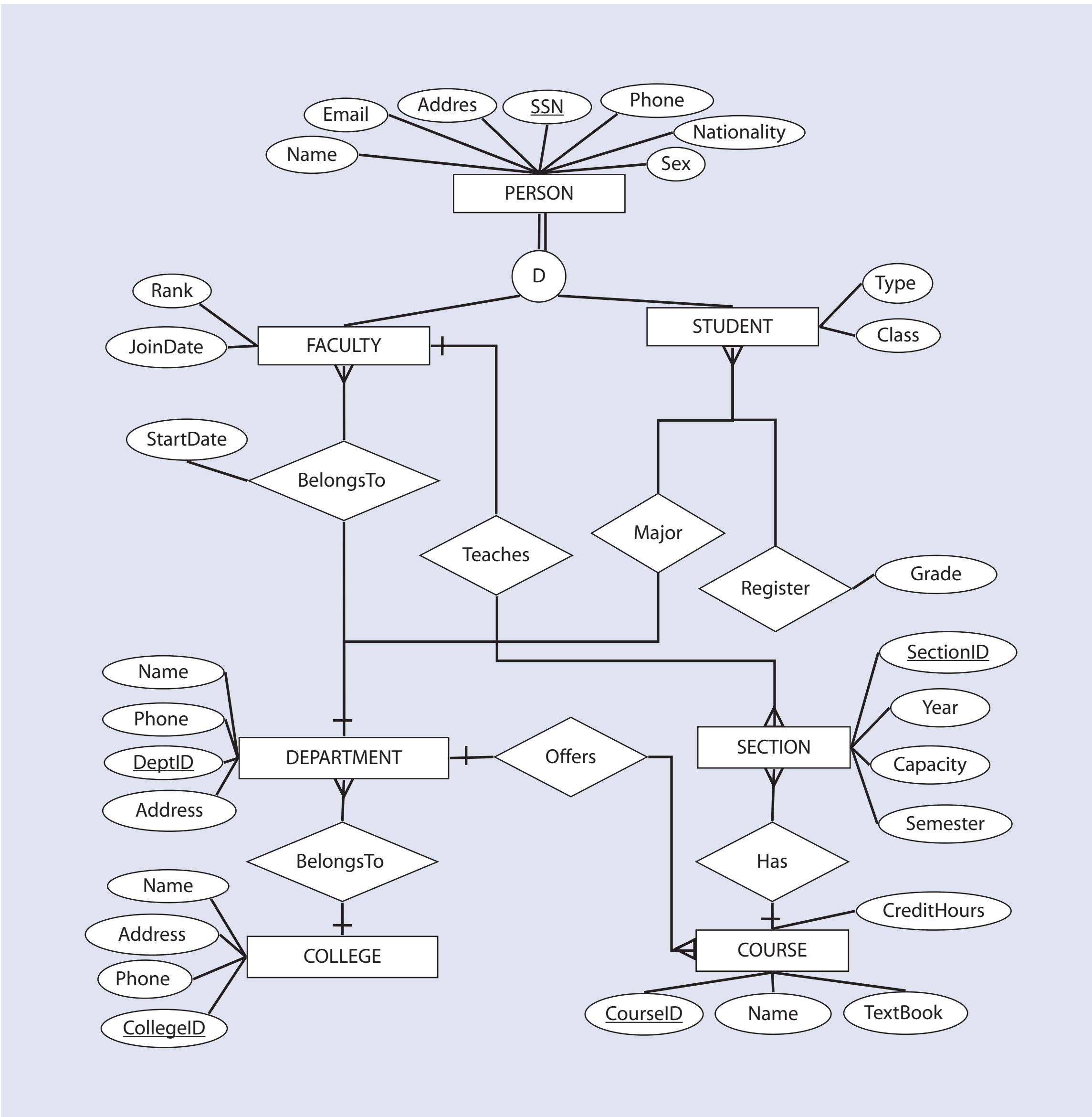


Figure 3.32 An EE-R diagram for the university database.

3.20 In-Class Assignment

Consider the following problem statement to draw an E-R diagram:

> The annual Bolder Boulder is one of America's top 10 km races. The race is held each Memorial Day in Boulder, Colorado. This race attracts world-class runners as well as casual joggers. It is a point-to-point race beginning at the Bank of Boulder at the northeast corner of the city, winding throughout the city streets, and ending near the town center in the University of Colorado's football stadium.

The organizers record the following information for each race: the date of the race; the total number of runners registered for the race (on-line pre-registration is possible); the actual number of participants; the number of male runners; the number of female runners; the name of the male winner; the name of the female winner; the name of the male master (runner over the age of 40) winner; and the name of the female master winner. In addition, the following information about each participant is recorded: social security number, name, birth date, gender, address, age, and certified personal record (PR) running times for a 10 km race.

3.21 Summary

- An *entity type* defines a collection of entities that have the same attributes. An *entity instance* is a single item in this collection, and an *entity set* is a set of entity instances.
- An *attribute* is a property or characteristic of an entity type that is of interest to an organization.
- Most attributes are *single-valued*, meaning they have a single value for an entity instance. A *multi-valued* attribute is an attribute that can take more than one value for an entity instance.
- A *derived attribute* is one whose value can be derived from other attributes.
- A *key attribute* is an attribute or a combination of several attributes that uniquely identify an individual instance of an entity type.
- The *degree of the relationship* is the number of entity sets that participate in a relationship.
- A *unary relationship* R is an association between two instances of the same entity types.
- A *binary relationship* R is an association between two instances of two different entity types.
- A *ternary relationship* R is an association between three instances of three different entity types.
- The *maximum cardinality* represents the maximum number of instances of any entity B that can be associated with any instance of any entity A.
- The *minimum cardinality* of a relationship between the entity types A and B is the minimum number of instances of entity B that must be associated with each instance of entity A.
- An *associative entity* is an entity type that associates the instances of one or more entity types and contains attributes particular to this association.
- A *strong entity type* exists independent of other entity types. A *weak entity type's* existence depends on another entity type.
- The *Enhanced Entity-Relationships (EE-R) model* is a revised E-R model that extends the original E-R model and supports additional semantic concepts by providing new modeling constructs.
- A *superclass* is an entity type that has one or more distinct sub groups with unique attributes. The subgroups should be of importance to the organization and are therefore necessarily represented in a data model.
- A *subclass* is an entity type that shares common attributes or relationships distinct from other subclasses.
- *Attribute inheritance* is the property by which subclass entities inherit values for all attributes of the superclass.
- *Generalization* is the process of defining a general entity type from a set of specialized entity types by identifying their common characteristics.
- *Specialization* is a process of defining one or more subclasses of a superclass by identifying their distinguishing characteristics.
- *Participation constraints* dictate whether every instance of a superclass must participate as an instance of a subclass. If it is mandatory for a

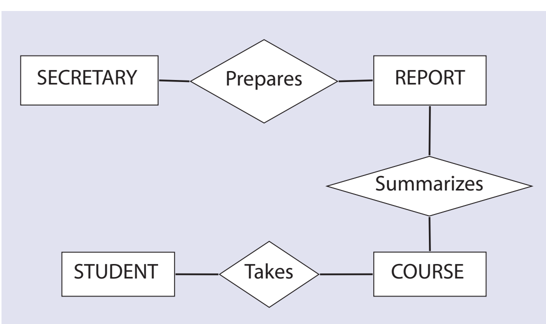
superclass instance to be a member of at least one subclass, it is known as the total participation rule. When some instances of a superclass are free to not participate in any of the subclasses, it is referred to as the partial participation rule.

- *Disjoint constraints* dictate whether it is possible for an instance of a superclass to be a member of one or more subclasses simultaneously. The *disjoint rule* states that if the instance of superclass is a member of a subclass, then it must belong to exactly one subclass. The *overlap rule* states that members of a subclass can belong to more than one subclass.
- A *subclass discriminator* is an attribute of a superclass that classifies its instance into the appropriate subclass. The value of a discriminator determines the target subclass for an instance of a superclass.
- A superclass/subclass hierarchy is a hierarchical structure of a superclass and subclasses wherein each subclass has exactly one superclass.

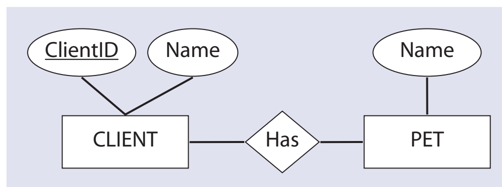
3.22 Exercises

3.22.1 Review Questions

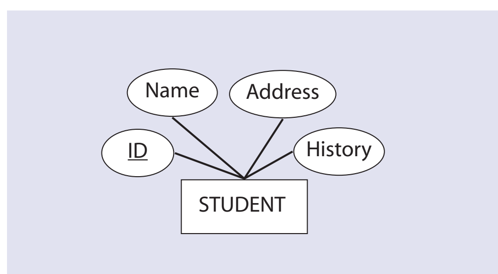
1. Explain the following terms using an example: entity-relationship model, entity type, weak entity, attribute, key attribute, derived attribute, multi-valued attribute.
2. Explain, using an example (other than the ones discussed in the book), the contrast between the following terms:
   a. entity type; entity instance
   b. strong entity type; weak entity type
   c. simple attribute; composite attribute
   d. stored attribute; derived attribute
3. Provide an example of multiple relationships between entities. Draw the E-R diagram.
4. Under what conditions is a relationship converted to an associative entity type? Give an example.
5. Explain why we study the E-R model of a database.
6. A department in a university stores the information about its students and courses in a database. The administrative assistant manages the database. At the end of the semester, he prepares a report about each course. Is the E-R diagram correct? If not, explain why and draw the correct diagram.



7. A veterinary doctor stores the information about his clients and their pets in a database. Complete the following E-R diagram.



8. Explain the term "degree of a relationship." Provide a simple example for each of the three types of relationships described in this chapter and draw the corresponding E-R diagrams.
9. A university maintains a detailed database of its student information. The following is the E-R diagram of entity STUDENT. Draw an alternative E-R diagram for this database.



10. Discuss the reasons for choosing one of the attributes as an identifier of an entity type.
11. "Enhanced E-R diagrams provide more flexibility in designing a database than E-R diagrams." Support this statement with an example.

12. Using an example, define the following terms: subtype, subtype discriminator, disjoint rule, overlapping rule. Demonstrate how they relate to one another.
13. Using an example, explain the reasons why a database designer considers using subtype/supertype relationships.
14. The Admissions office in a university stores the information about the university entrance applications in a database. In order to facilitate processing the applications, the office classifies the applicants as freshman, transfer, graduate, and international students.
    a. Draw the EE-R diagram of this database. Identify a unique attribute for each entity subtype.
    b. Add a subtype discriminator for each of the subtypes of the entity APPLICANT.
15. Using an example, explain the difference between the generalization rule and the specialization rule.
16. The Accounting department in a company keeps the information about its employees in a database. The employees are classified as part-time employees, full-time employees, and interns. Interns are usually students who work with the company during the summers. Complete the EE-R diagram below by:
    a. Adding a unique attribute (relationship) for each entity subtype
    b. Adding an attribute shared by all entities
    c. Identifying whether the disjoint or overlap rule applies
    d. Identifying whether the total specialization or partial specialization rule applies

3.22.2 Hands-On Exercises

1. Savannah's family has owned and operated a 640-acre farm for several generations. Since the business is growing, Savannah is considering building a database that would make it easier to manage the farm's activities. She is considering the following requirements for the database:
    a. For each livestock classification group (for example: cow, horse, etc.), Savannah keeps track of each animal's identification number and classification.
    b. For each crop, she records the crop identification number and the classification.
    c. Savannah has recorded the yield of each crop classification group during the last ten years. The records consist of the year, yield, sales, crop price, and amount of money earned.
    d. Savannah has also recorded the yield of each livestock classification group during the last ten years. The records consist of the following historical data: the year, the (historical) selling price per head, the number of livestock in the end of the year, the number of livestock sold during a one-year period, and the total amount of money earned.

    Draw an E-R diagram for this application. Specify the key attribute of each entity type.
2. The Coca Cola Company in Atlanta, Georgia produces a wide range of products that are delivered to its worldwide clientele once a week. The company stores information about its employees, products, and customers in a database that includes the following set of tables:
    a. The company records the following information about its customers: customer identification number, name, address, X (longitude) and Y (latitude) coordinates of their location, and the amount of time (in fractions of an hour) required making a stop at that location.
    b. Each employee has an employee identification number, name, address (which consists of a city, state, and zip code), gender, birth date, position in the company, wage earned per hour of regular time work, wage earned per hour of overtime work, number of dependents, and number of years worked for the Coca Cola Company.
    c. Each product has a product identification number, price, and number of units produced per day.

    Products may be ordered by one or more customers, and a customer may order one or more products. Furthermore, employees produce one or more products, and a product may be produced by exactly one employee. Draw an E-R diagram for the above description and add the minimum and maximum relationship cardinalities. Identify the key attributes for each entity and the composite attributes.
3. Consider the database in Hands-On Exercise 2. Update the E-R diagram for the cases described below. For each entity, identify the key attribute(s).

a. For tax purposes, the Coca Cola Company extends the data kept for each employee to include additional information about their dependents. This information consists of each dependent's name, birth date, and age.
b. The company has decided to record information about its suppliers and the raw material(s) supplied by them. The following information is recorded for each supplier: the supplier identification number, the address, and the name of the contact person. The raw material information consists of the raw material identification number and the material's name.
c. The company keeps a fleet of vehicles to facilitate the distribution of the products to the customers. Each vehicle has an identification number, model, and capacity.

It should also be noted that employees use raw materials to produce one or more final products, which are delivered to the customer. A supplier supplies many raw materials, but a particular raw material is purchased from only one supplier. It is not necessary to include all the attributes and cardinalities from the previous question. Just be sure to include the entities to demonstrate the relationships and attributes for the new entities.

4. Again consider the database in Hands-On Exercise 2. The Coca Cola Company has a few plants that are distributed throughout Georgia. For each plant, the following information is stored in the database: a plant identification number, the address, and the X and Y coordinates of the plant location. Note that the inventory level, inventory capacity, and quantity produced of a particular product differ by plant.

   The systems of relationships that exist are as follows. A plant may have one or more employees, and an employee works in exactly one plant. Additionally, a plant produces one or more products, and a product may be produced in one or more plants. Also, an employee may contribute to one or more products, and a product is produced by exactly one employee. It is not necessary to include all the attributes and cardinalities from the previous question. Just be sure to include the entities to demonstrate the relationships and attributes for the new entities.

5. Major airline companies that provide passenger services in Taiwan are UniAir, TransAsia Airways, Far Eastern Transport, and Great China Airlines. Taiwan's Federal Aviation Administration (TFAA) maintains a database with information about all the airlines. This information is made accessible to all airlines in Taiwan with the intention of helping the companies assess their competitive position in the domestic market. The information consists of the following:
   a. Each airline has an identification number, a name and address, a contact person's name, and a telephone number.
   b. Each aircraft has an aircraft identification number, a capacity, and a model.
   c. Each employee has an employee identification number, a name, an address, a birth date, a gender, a position within the company, and a qualification.
   d. Each route has a route identification number, an origin, a destination, a classification (domestic or international route), a distance of the route, and a price charged per passenger.
   e. Each airline records information about its buy/sell transactions. (For example, selling an airplane ticket is a sell transaction; paying for maintenance is a buy transaction.) Each transaction has a transaction identification number, a date, a description, and an amount of money paid/received.

   The above information is related as follows. Each employee works for exactly one airline. Airlines assign different aircraft on different routes based on the availability. Furthermore, each airline makes one or more transactions; however, each transaction is associated with exactly one airline. Make the necessary assumptions about the other relationships and draw an E-R diagram.

6. Consider the database described in Hands-On Exercise 5. Each airline owns different aircraft models, each with a different capacity. Depending on the length of the route and flight classification (domestic or international) the aircraft are assigned to different routes. The relationship between the airlines, aircraft, and routes is a ternary relationship. Each flight carries a number of passengers, has a particular time length (which depends on the distance of the route and the model of the aircraft), and has a departure and arrival time. Draw the E-R diagram for this database.

7. "Electronic commerce" is one of the most common terms in the business world. It is the buying and selling of goods and services on the Internet. One of the most popular products in e-commerce, as it is more commonly known, is the compact disc. This exercise describes the database of a CD warehouse.

The company's customers and employees access the database. Assume that customers have access to a company's Web site and that they are able to open an account by providing their social security number, name, address, and music preferences. For each employee, the following information is recorded: an employee identification number, a name, an address, a birth date, and the title of the position within the company.

The products for the CD warehouse are the albums. The database records the following information about each album: an album identification number, the name, the group name, the musical category, the name of the vocalist, the names of the other band members, and the number of CDs in stock. The database also keeps the following information about suppliers: a supplier identification number, the address, the name of the company, and the name of the contact person.

An album may be bought by one or more customers, and one customer may buy one or more albums. Additionally, an employee may monitor one or more albums; however, an album is monitored by exactly one employee. Suppliers may provide one or more albums; however, an album is supplied by exactly one supplier. Using the above information, draw an E-R diagram for the CD warehouse database. Identify the relationship cardinalities and the key attribute of each entity.

8. One of the three-star hotels in the Miami area is in the process of updating its database. The hotel has various room types on each of its floors. The rooms may be regular, deluxe, or a suite and each can be either a single, double, or triple. The suites have ocean views and are bigger than the regular rooms. The deluxe rooms are as big as suites, but they do not have an ocean view. All the rooms have air conditioning. Most of the rooms are non-smoking, but the hotel offers some smoking rooms as well. Each floor has a different number of a particular room type. The price of each room differs by the size of the room, the view, and the room's location (first floor, second floor, etc.).

The customers are charged on a per-day basis. The number of days is computed based on the check-in time and the check-out time. The following details are stored for each customer: name, address, check-in date, check-out date, payment method, and final bill amount. In addition to the room charges, there may be extra fees, such as telephone usage, fax services, extra beds, and room service. For this application, we assume that a room can be booked by more than one customer as long as there is no overlap and that a customer can be assigned to any available room. Draw an E-R diagram for the hotel database and state any assumptions made in order to develop a complete diagram.

9. The Newark divisional office of the Life Insurance Corporation of America stores all the necessary information about its policyholders in a database. A policyholder pays a premium until the maturity of the policy or until his or her death, at which time the sum assured and the bonus are paid to the beneficiary. The premium to be paid is determined based on the age of the person proposed and the term of the policy. The Newark division records the following information about each policyholder: social security number, name, address, date of birth, maturity amount, and annual premium.

The corporation has divided its Newark division into 15 zones for its convenience. For each zone, they store the Zone ID and the location. Each zone also has a manager. Every zone has a number of agents allotted, typically ranging from 10 to 20 and every agent must procure a minimum of 10 customers. Each policyholder in a particular zone is assigned to only one agent affiliated with that zone; this agent puts forth the terms of the policy. Additionally, an agent in a particular zone can serve multiple policyholders. Make assumptions for the remaining relationships and draw an E-R diagram for the corporation's database.

10. Ford distribution centers provide automotive parts to authorized dealers. The dealers then distribute the parts to customers throughout North America. Ford is under pressure to provide excellent customer service at a minimum cost. Maintaining a well-organized database of information will contribute to achieving this goal.

Ford stores the following information about each of its distribution centers: the identification number, the location (X longitude coordinate and Y latitude coordinate), the address (city, state, zip code) and the name of the contact person. The following information is kept about each dealer: the identification number, the dealer's location (X longitude coordinate and Y latitude coordinate), the address (city, state, zip code) and the name of the contact person.

The following information is kept about each product: a product identification number, its name, its price, its weight and its value. Ford also records the following information about the flow and cost data for all the distribution centers to dealer channels: distribution center identification number; dealer identification number; product identification number; the number of miles between each distribution center and its dealers (determined using the road network); the quantity of products being shipped; and the dollar value of the shipment.

A distribution center can ship many products to different dealers. A dealer, meanwhile, can receive several products from different distribution centers. Every week, the distribution center sends to the dealer a shipment for which the following information is recorded: distance (using the road network); the quantity of products being shipped; the monetary value of the shipment; and the date of the shipment. Using this information, draw an E-R diagram for the Ford database. State any assumptions made in order to develop a complete diagram.

11. The University Housing Office receives many applications from graduate and married students requesting apartments on campus. The housing villages are located at five different locations, and each village has about 500 apartments. Each apartment falls into one of the apartment categories, which are determined based on the following criteria: village location; if the apartment has a dishwasher; whether it is a one- or two-bedroom; whether it has central air conditioning or a window unit; and if it is furnished.

The Housing Office records the following information about the head of each household: social security number, name, telephone number, marital status, and college and department in which he or she is enrolled. The Housing Office also keeps the following information about the students who have applied for on-campus housing but have not yet been assigned to an apartment: social security number, name, telephone number, marital status, college and department in which he or she is enrolled, and his or her apartment preference.

An applicant applies to one or more villages, and a resident resides in exactly one apartment. Make suitable assumptions for the other relationships, and draw an E-R diagram for the University Housing Office database.

12. A database is being developed to manage the course timetable of an academic institution. For each course, the following information is recorded: an identification number, the name of the course, the number of students attending the course, and the number of credits.

For each teacher, the following information is recorded: social security number, name, department, skills, and yearly salary. For each class period, the following information is recorded: period number, starting time, ending time. For each room, the following information is recorded: room number, room type (classroom, office, auditorium, or computer lab), and capacity.

Staff members refer to the above information in order to make appropriate assignments of a particular course to a specific time period, classroom, and professor. A professor teaches one or more courses; however, a course is taught by exactly one professor. Using this information, draw an E-R diagram for this database. Clearly state any assumptions made about other relationships in the diagram.

13. A database is being constructed to monitor the teams, players, and games of the national intercollegiate football championship. For each player, the following information is recorded: social security number, name, address, birth date, team position, and the number of years the player has been with the team.

For each team that participates in the football championship, the following information is recorded: the name of the team; the name of the university it represents; the rank of the team in the current season; the number of games that the team has won in the current season; and the number of games that the team has lost in the current season. The database also keeps information about the team coach. This information includes the following: social security number, name, age, number of years coaching the current team, total number of years coaching, and number of times that the team he or she has coached won the championship.

The following information is recorded about each game: the game identification number, the date and place of the event, the start time, the end time, and the winner. A coach can lead exactly one team, and a team can have exactly one coach. Each team may play one or more games in a season, and a game is played by one or more teams. Additionally, a team may have one or more players, and a player can play for exactly one team. A team may win more than one game in a season, and a game is won by exactly one team. Draw an E-R diagram for this database. Clearly state any assumptions made.

14. Consider the database described in Hands-On Exercise 13. Suppose that we want to design an attribute for the COACH entity type to keep track of previous employment. Such an attribute should have one entity for each college he or she worked for. Each entry should be composed of a college name, a start and end date of employment, and the title of the position (coach, assistant coach, etc.). Update the E-R diagram from Hands-On Exercise 13 to account for this modification.

15. Again consider the database described in Hands-On Exercise 13. Each player of a football team not only plays for a college, but, at the same time, is enrolled at that college. Each college has one or more players enrolled, and a player is enrolled in exactly one college. Additionally, each team belongs to exactly one college. Update the E-R diagram to account for this information. Include only the name of the university.

16. JobSearch.com is an Internet-based company. It provides information about open positions to students looking for a job as well as information about candidates to the companies. Both the companies and the students can access the database. Students access the database to post their resumes and to look for open positions in their area of interest. Companies access the database to post its job openings and to look for the candidates who best fit its needs.

For every student, the following information is recorded in the database: student identification number, name, birth date, address, gender, country of citizenship, immigration status, university, major department, degree program, and skills. For every company, the following information is recorded: identification number, name, address (city, state, and zip code), telephone number, Web site address, and industry. When a job opening is submitted, the companies specify the following: posted date, job description, type of job (full-time, part-time, co-op, etc.). A contact person, who works for the company then looks into the database, selects candidates, schedules the interviews, and conducts the interviews.

A particular job posting can belong to exactly one company; however, a company may post multiple jobs. Also, a student may be pursuing one or more jobs, and a job may be pursued by one or more students. Using the above information, draw an E-R diagram for the JobSearch.com database. Clearly state any assumptions made.

17. The Florida Bus Traveling Agency needs to computerize its reservation database systems. The corporation has 18 buses spread over 20 routes to various destinations in Florida. There are two types of buses: regular and super deluxe. It has 10 regular buses with a seating capacity of 48 and eight super deluxe buses with a seating capacity of 36.

For each bus the following information is stored in the database: bus number, capacity, and type. The buses travel certain routes. For each route, the following information is recorded: route identification number, city of origin, and city of destination. Customers usually book trips that do not necessarily have to correspond to bus routes. A trip begins in one of the cities that is visited by a route (not necessarily the city of the origin of the route) and ends at another city visited by that route (not necessarily the city of the destination of the route). For every trip a customer books his or her ticket for, the following information is recorded: trip identification number, city of origin, city of destination, departure time, and arrival time.

A bus is assigned to a particular route, which passes through the origin and destination cities of one or more trips. Many buses can pass through a particular route. Draw an E-R diagram for the

Florida Bus Traveling Agency database. Clearly state any assumptions made.

18. SunRise hotel is located in Palm Beach. The hotel keeps a detailed database of the rooms and special services offered, as well as a database of employees and customers. Keeping a detailed database of the rooms facilitates the management of the hotel's everyday activities.

The hotel keeps the following information about each customer: social security number, name, and address. For every room, the following information is recorded: room identification number, location (first floor, second floor, etc.), status (available or not available), rate, and room type (regular or luxurious). The hotel offers special services to customers, if requested. For the special services, the following information is recorded: identification number, rate, and service type.

A customer may occupy exactly one room, and a room may be occupied by more than one customer as long as there is no overlap. Additionally, a customer may use more than one special service. Make suitable assumptions for the remaining relationships, and draw an E-R diagram for the SunRise hotel database.

19. GERU is a regional multi-service utility providing electric (E), natural gas (NG), water (W), and telecommunications (T) services to its customers. GERU is interested in developing a database of customers, services provided, and rates. This database will help GERU to maintain its operations and also to enable customers to track their energy consumption, check their payment history, report power failures, and tap into an array of services and useful information.

The customers are classified into four major groups: domestic (D), commercial (C), agricultural (A), and industrial (I). Currently, GERU has 4,500 domestic connections, 1,200 commercial connections, 100 agricultural connections, and 500 industrial connections. For each customer, the following information is recorded: identification number, name, address, classification, and sign-in date.

Each connection offered by GERU has associated characteristics and rates that depend on the type of service and the customer classification. (For example, domestic rates differ from the industrial rates.) For each service, the following information is recorded: service identification, type, and rate. GERU also sends a bill to its customers every month for using its services. This bill includes the consumption total, the money due, and the due date.

Customers may use more than one service, and service may be requested by more than one customer. Using this information, draw an E-R diagram for the GERU database. Clearly state any assumptions made.

20. VedMed is a veterinary hospital. The hospital keeps a database of its clients, pets, employees, and inventory. This information is used to provide better customer service and to manage everyday operations.

The database includes the following information about each of the customers: customer identification number, name, address, and e-mail address. The database records the following information about each pet that visits the hospital: name, species, and birth date. In addition, for each pet, a history of the visits to the doctor is maintained. For each visit, the date, type of service offered, additional comments, and payment amount are recorded.

Detailed records about the doctors working for the hospital are also stored in the database. Part of this information is made available to the customers in order to help them choose the doctor who best fits their needs. The doctors' database includes the following: identification number, name, address, gender, area of specialization, and degree earned. The hospital has a pharmacy where the customers can purchase medications. For every item in the inventory, the following information is recorded: identification number, name, description, price, quantity on hand, and safety stock level.

A pet may visit multiple doctors, and a doctor receives visits from one or more pets. Additionally, a customer may purchase one or more medications. Make suitable assumptions for the remaining relationships and draw an E-R diagram.

21. Memorabilia is an on-line company that buys sports products from various producers around the country and sells them to on-line customers. Customers visit Memorabilia's Web site, select an item, and make an order. As soon as the customer's order is received, the product is delivered to the customer, and the inventory level is updated. The company orders a particular product from a supplier when the inventory level drops below a certain level. The company has decided to maintain a detailed database of the customers, suppliers, and products to manage the operations. The following information is stored in the database:

a. Customer: name, address, gender, and preferred sport.

b. Supplier: supplier identification number, name of the company, and address.

c. Product: product identification number, price per unit, amount in the inventory, amount requested of the suppliers but not yet received, amount ordered by the customers but not yet shipped.

A customer may order one or more products, and a product may be ordered by one or more customers. Additionally, products may be provided by one or more suppliers, and a supplier may provide more than one product. Using this information, draw an E-R diagram for Memorabilia's database. Clearly state any assumptions made.

22. A small bookstore has been keeping track of its business mainly on paper. The owner is planning to expand the business and would like to have a state-of-the-art database system to improve bookkeeping and customer service.

As a caring bookstore owner, she would like to send information about new books, new editions of a book, and deals to the customers based on their profiles. If the customer is a faculty member at a university, then she wants to offer free copies of a new textbook or a new edition of an existing textbook. If the customer is a student who likes reading science fiction, she wants to send monthly notices about new releases.

The system will help the store maintain details about books, publishers, and customers. A book may be a textbook, a novel, a comic, a children's book, or a cookbook. Publishers are the suppliers of the books. The bookstore buys books from many publishers. Typical customers of the store are libraries, institutions, and individuals such as students, faculty, and others. If they wish, customers can open an account with the store and be given a customer number. With their customer number and a password that they set, the customers are able to login to the database from their own PC. They are able to search books, place orders, check their account status, and also submit reviews about books they have read. The database also maintains a record of the transactions. For example, when a customer places an order, a payment is made. In the case that the inventory level for a particular book drops below a certain limit, the bookstore places an order to the publishers for new copies.

A customer may buy one or more books, and a book may be bought by one or more customers. Additionally, each book is provided by exactly one publisher; however, a publisher may provide one or more books. For each entity type, identify corresponding attributes and draw an E-R diagram.

23. Medicare is a medical service program that provides acute care for hospitalization, visits to a doctor's office, medical tests, and a limited amount of skilled nursing care for patients recuperating from an acute illness. Medicare covers 12 federally mandated services and several optional services. Medicare is developing a database management system that will perform the following functions: confirm patient eligibility, assign doctors, and pay doctors, pharmacists and hospitals promptly. The system should be designed to structure, store, retrieve, and analyze such critical Medicare management information as information about patients, doctors, pharmacies, and prescription drugs. The system stores the following information:

a. Patients: identification number, name, address, birth date, gender, identification number, sign-up date, and annual income.

b. Patient history: date of the visit, duration of the visit, diagnosis, and medication prescribed.

c. Doctors: identification number, name, address, gender, birth date, specialization.

d. Pharmacy: identification number, address, telephone number, name of contact person.

e. Pharmacy inventory (for every drug kept in the inventory): identification number, name, price, date of the last purchase, amount in the inventory, and amount ordered (not yet received).

f. Sales at a pharmacy: identification number, date of the transaction, and quantity purchased.

A patient may visit multiple doctors, and a doctor is visited by one or more patients. Additionally, a patient may buy his or her medication from the pharmacy. Using this information, draw an E-R diagram for the Medicare database management system.

24. Wood paneling manufacturers face a number of complex decisions in their daily proceedings. For example, allocating production resources and combining various raw materials to meet production goals require real-time decision making. Due to changing supplies and costs, the management of a wood paneling manufacturer has decided to build a database system to fine-tune their production processes.

Consider a wood paneling manufacturer that produces a furniture-grade particleboard. Each of the panels consists of a middle layer and two surface layers that are symmetrical. To enhance its mechanical properties, each panel has several strata of different materials, compositions, and specific gravity. A panel's quality can be controlled by specifying different density profiles and raw material requirements. There are eight different types of raw materials, and the database records information about the manufacturer's suppliers, the quantities available, and the maximum capacity of the bottleneck equipment. The raw material needs can be supplied by six different sources, including sawdust, shavings, sawmill, residual, chips, and short or long logs (softwoods or hardwoods). The database should also keep a detailed matrix of specifications that reveals the quantity of each individual raw material allowable in various layers of the different products.

Another variable that affects the production schedule is the production capacity. The database keeps track of each piece of equipment's production capacity, the equipment type (name), the maintenance date, and a description of its activities.

Given that each piece of equipment is assigned to produce a product from raw materials and that each raw material has exactly one supplier, draw an E-R diagram of the manufacturer database. Make suitable assumptions for the remaining relationships.

25. The housing office in a university is facing difficulties in assigning undergraduate students to dormitories. One problem it is facing is that the current system does not consider student preferences; therefore, once the assignment is completed and announced, students ask to be reassigned to a different room or even a different building. Also, the number of undergraduate students entering the university is increasing every year, yet the amount of dormitories remains the same.

The housing office has decided to build a database management system to assign students to dormitories. In particular, the new system will review students' preferences, including air conditioning, dorm location, and room location. It will also review each student's record from previous years, including payment history and history of rule violations or damages, to determine placement. The database will record these factors as student information and residential history.

There are ten dormitories on campus. Each dormitory is classified into the four following categories: married couples, males only, females only, and co-ed. Each dormitory has up to four floors and each floor has up to ten apartments. An apartment is classified as a single, double, or triple room. Some apartments have air conditioning. Each apartment has a pre-set yearly cost, and students make one payment for their one-year stay. A student can apply to exactly one dorm; however, a dorm houses more than one student. Draw an E-R diagram and state any assumptions made.

26. People have tried to earn large profits from betting on horse races by finding ways of outwitting other betters. A large variety of books has been published on horse races, and each author claims that following his or her methods will lead to profits. So far, Bert has not been able to make big profits using any of the existing methods. Therefore, he has decided to build his own system to place sophisticated bets.

Horse racing is one of the most thoroughly documented sports. There are thousands of pieces of data published on each upcoming race. This data is available to Bert, and he needs to build a database to organize and store all this information.

a. For each horse, the following information is recorded: name, owner, height, breed, and parents.

b. For each jockey, the following information is recorded: social security number, name, weight, height, number of years the jockey has been racing, and age.

c. For each racetrack: name, location (address), type (dirt, turf, etc.), and distance.

d. For each race: name, attendance, prize amount, winner, distance, and location.

A jockey rides exactly one horse, and a horse is ridden by exactly one jockey. Additionally, a horse may participate in one or more races on a racetrack. Draw an E-R diagram of this database.

27. The United States Tennis Association (USTA) is concerned about developing a ranking system of the tennis players that is objective, consistent, and broad-based. These qualities are important in a ranking system since rankings directly affect the acceptance of a tennis player's entry into a tournament and his or her placement in the draw.

USTA use a particular formula to determine the ranking points for each player. The formula uses the following information: number of tournaments played by the player; tournament points earned; number of matches played; and number of wins of the player (say player "*i*") over player *j*. The new system's performance relies on the efficiency of the database, which records the following information:

a. For each player: social security number, name, tournament points earned, current ranking, weight, height, and birth date.

b. For each court: name, type (grass, clay, and hard surface), location.

c. For each tournament: name, location, and tournament strength.

The strength of a tournament is a function of the quality of the players and the size of the tournament. A player may play one or more matches in a tournament on one or more courts. Also, a tournament involves many matches with different players who may play on the same court. Using this information, build an E-R diagram for the database described above.

28. The traditional MBA program has been receiving criticism because it is focused on analytical training. However, potential employers are looking for executives with a broader education. The American Assembly of Collegiate Schools of Business (AACSB) requires the following components in an MBA curriculum: a common body of knowledge in management; a field of specialization; and general competence for overall management.

Due to the business world's growing interest in a broader curriculum, most of the schools are trying to improve their MBA programs, which is not an easy task. As a first step, AACSB has decided to build a database that contains information about students and alumni/alumnae, schools, and courses offered. This information will be useful in preparing the new curriculum for the MBA program. The database consists of the following entities:

a. For each school: name; budget allocated for the MBA program; location (suburban, urban, rural); and whether it is AACSB accredited or not. (AACSB accredits those schools that meet certain requirements. Being AACSB-accredited is important for schools as it indicates their quality.)

b. For each alumnus or alumna: name, social security number, gender, current position, current salary, and GPA at graduation.

c. For each course: name, code, type of course (foundational, functional, general, sectoral, or institutional), and topics covered.

d. For each current student: name, social security number, gender, current GPA, courses completed, and the date she or he began the program.

Each school has one or more students and alumni/alumnae, whereas each student or alumnus/alumna can belong to exactly one business school. Additionally, each school offers one or more courses, which are taken by one or more students. Make suitable assumptions for the remaining relationships and draw an E-R diagram for this database.

29. Blue Bell is a large apparel manufacturer. It produces jeans and several other lines of sports and casual apparel for men, women, and children in the USA and has a substantial international business. Each of the above production lines breaks down into styles, lots, sizes, and stock keeping units (SKU). With time, the management began to realize that the number of product lines, styles, and SKUs had grown tremendously. The production process had incorporated several new automated manufacturing operations in response to a change in customer tastes. The management was concerned about the high investment in working capital. A large part of the working capital goes toward inventories. In this process of extending the business, Blue Bell is eager to effectively coordinate its activities. A well-organized database management system will help Blue Bell to do so. The database includes the following entities:

a. Production line: name (for example, jeans, t-shirts, swimwear, etc.), identification number, and classification (for example, men's-wear, women's-wear, boy's-wear, and girl's-wear).

b. Lot: identification number and name (for example blue, green, etc.).

c. Product: identification number, name, size, and inventory level.

d. Customer: identification number, name, address, and telephone number.

The relationships are explained as follows. A customer may buy one or more products, and a product may be bought by one or more customers. For each transaction, the date and value are recorded. Additionally, each product comes from exactly one lot, and a lot contains one or more products. Each lot comes from exactly one production line, and a production line produces one or more lots. Given this information, draw an E-R diagram for the Blue Bell database. Clearly state any assumptions made.

30. The main library of a university is interested in developing a database management system. The library carries the following items: books, journals, conference proceedings, reference textbooks, and copies of some recorded lectures on CD. These items are loaned to members in various categories.

The library has a total of 250,000 items. The library uses a coding scheme for classifying the items. Using this code, the librarian is able to identify the type of the item (book, journal, CD), the subject, the title, the author(s), and whether the item is on loan, overdue, or on the shelf.

Members of the library are undergraduate and graduate students, faculty members, staff, part-time students, and visiting scholars. The university issues to all members an ID card that can also be used at the library. Students and visiting scholars renew their ID card every semester. Faculty and staff ID are valid as long as they are still employed. The database records the social security number, name, and address of each member.

A member who borrows a CD cannot take it outside the library and has to use the computers in the library to watch it. Journals and conference proceedings can be borrowed for 2 days, and reference textbooks can be taken on loan overnight only. Faculty and graduate students can borrow books for a period of 3 weeks, other students and visiting scholars can borrow them for 2 weeks, and staff can borrow books for 1 week only. Note that journals, unlike books, do not have a single author. Journals are usually published a few times a year. An item may be borrowed by one or more members, and a member may borrow more than one item. Using this information, draw an EE-R diagram for the university library database. Clearly state any assumptions made.

31. A financial institution offers to its customers various investment schemes such as shares and debentures. The company also accepts fixed deposits from the general public, institutions, and its employees. The company maintains a database with valuable information about its customers (such as ID number, name, and address) and its financial investments.

Fixed deposits have varying terms of 1, 2, and 3 years. For fixed deposits, there are currently two payment schemes. Under the first scheme, the investors get the principal in addition to the interest on maturity. Under the second scheme, they get the principal on maturity, but interest is paid periodically. The interest rates for fixed deposits under the first scheme may be paid quarterly, semi-annually, or annually. Investors have the option to renew their deposits on maturity. The company also raises debentures periodically. The debentures may be either convertible or non-convertible. Convertible debentures can be converted to equity shares on completion of the period. Debentures are issued for periods of one, two, or three years. The company also calls for shares periodically. The company issues two kinds of shares: equity shares and preferred shares. The dividend is declared at the end of the year. The system keeps track of the dividend rates and the dividends issued to the shareholders.

The company records the name of the department for the employees and the name of the contact person for institutions. Draw the EE-R diagram of this database. Identify the subtypes (if any) of the entities FINANCIAL INVESTMENT and MEMBER. Identify one (or more) unique attribute (relationship) for each subtype as well as one (or more) attribute that is shared by all entity subtypes.

32. Great Marbles is a small company in Venezuela that produces two main types of products: gems and marbles. They produce two sizes of gems and six sizes of marbles. Both products are offered in 25 colors. The different kinds of decoration marbles are distinguished by color, size, and shape. Each gem is classified as a matte gem or a luster gem.

The database system keeps track of the inventory levels, the backorder level (orders not satisfied yet), and the number of outstanding orders (orders that are not yet received). The management refers to this information to decide when, which, and how many products to produce. Other than the information about the products, the database records information about the raw materials, suppliers, and customers. For each type of raw material, the database keeps the following information: identification number, name, and price per unit. For each supplier, the database records the supplier identification number, name, and address.

The company sells its products in Venezuela as well as in the US and other Latin American countries. For the international customers, the company provides special packaging and charges increased rates. For each customer, the following information is recorded: identification number, name, and address. A customer may order one or more products, and a product may be ordered by one or more customers. Additionally, each product is produced from one or more raw materials, and a particular raw material may be used in multiple products. Each raw material is supplied by exactly one supplier, and a supplier may supply one or more raw materials. Draw an EE-R diagram for the database described above. Clearly state any assumptions made.

33. The medical school at the University of Florida serves UF students as well as the general public as a moderately sized hospital. The hospital stores information about patients, including name, address, date of visit, and doctor's name, in a database. The hospital does not charge the students for its services and charges reduced rates if the patient is a UF faculty or staff member.

Data about wards, equipment, and operating rooms are also recorded. The hospital has three types of operating rooms used for major, minor, and small operations, respectively. There are two types of wards: general and special. The hospital has 55 general wards and 35 special wards. The general wards have a capacity of eight beds each. The special wards have one or two beds. The hospital also has an intensive care unit with a capacity of four beds. The patients are charged on a per day basis, and the rates depend on the type of the wards.

The hospital uses the following equipment to examine patients: an X-ray machine, a CT-Scan machine, and an ultrasonic imager. If any of this equipment is needed to examine the patient, the patient is charged extra. The charges are based on the number of hours that the machine is in use. The X-ray machine costs $350 an hour, the CT-Scan machine costs $750 an hour, and the ultrasonic imager costs $150 an hour. Patients may require one or more wards, pieces of equipment, and/or operating rooms. Make suitable assumptions for the remaining relationships and draw an EE-R diagram for this database.

34. National Car Rental maintains a detailed database of its inventory (cars to rent) and customers. Customers rent a car mainly for two purposes: business and leisure. For each customer, National records the social security number, name, and address. If the customer rents the car for business purposes, in addition to the above information, National records the name of the company and the work phone number of the customer. In order to provide better service to customers, National prioritizes bookings based on length-of-rent (LOR).

National offers the following types of cars: luxury, midsize, and economy. For each type, the company calculates protection level. A protection level is the number of cars that should be reserved for the demand in the current class. The company monitors the number of cars of a particular type available, as well. National charges customers a daily rate depending on LOR and the type of car they rent. A customer may be assigned exactly one car, and a car may be assigned to one or more customers as long as there is no overlap. Draw an EE-R diagram for the above circumstances.

35. All academic departments in a university maintain a database of its students. Students are classified into undergraduate, graduate, and international students. There are a few reasons for grouping the students into these three categories. For example, the department's administrative assistant informs the undergraduate students about undergraduate courses offered, the graduate students about graduate courses and professional conferences, and international students about new immigration laws. Identify the subtypes (if any) of the entity STUDENTS. Also identify a unique attribute (relationship) for each subtype. Draw an EE-R diagram for a department's database.

36. A blood bank serves a critical purpose in providing a required type of blood to patients at critical times. A blood bank's database monitors the inventory of the blood together with relevant information such as blood type, date received, location, date of expiry, and donor.

The database stores information such as name, address, and telephone number for a blood bank. Supplementary information about the donors is recorded as well. Donors are classified into occasional and regular donors. For the regular donors, the database keeps information such as identification number, blood type, and history of donations. The database also keeps a list of healthcare providers in the area along with their addresses and telephone numbers. The healthcare providers are the customers of the blood bank. They keep track of the blood transactions performed. These transactions are classified into normal transactions and unexpected transactions (for example, due to car accidents during the holiday season). The reason for keeping track of the unexpected transactions is to use this information to estimate the extra amount of blood needed in the inventory for each age group during the next holiday season. A blood bank receives a particular bag of blood from exactly one donor. The blood bank then distributes the blood to health care providers. Draw an EE-R diagram for this database.

37. YXZ is a construction company. The company keeps a list of employees as well as a list of jobs that are scheduled in a particular day. Every day, the management gets a list of required jobs and a list of employees available. A job is then assigned to the employee who has the skills needed to do the job. (In other words, an employee should have enough skills to perform the job assigned.) We want to build a database that will facilitate the process of assigning employees to jobs.

Employees are classified into three main groups: managers, engineers, and workers. Managers take care of managerial issues, engineers direct production processes, and workers perform labor-intensive jobs that require a certain level of technical skill. Jobs are classified into those that require a high level of technical skill, a moderate level of technical skill, and managerial skills. The classification of employees and jobs into groups facilitates the process of assigning an employee to a job.

An employee may perform one or more jobs, and a job is performed by exactly one employee. Draw an EE-R diagram for this database. Identify the subtypes (if any) of the entity EMPLOYEE. Identify one (or more) unique attribute (relationship) for each subtype as well as one (or more) attribute that is shared by all entity subtypes.