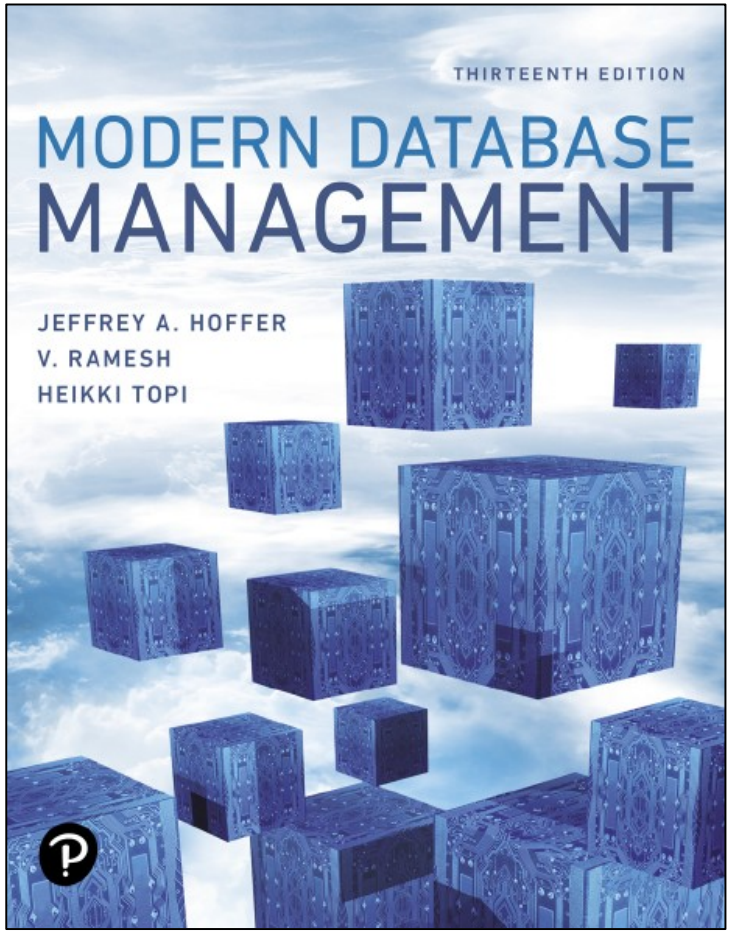
Modern Database Management

Thirteenth Edition



Chapter 2

Modeling Data in the Organization

Learning Objectives

2.1 Define Key terms

2.2 Understand importance of data modeling

2.3 Write good names and definitions for entities, relationships, and attributes

2.4 Distinguish unary, binary, and ternary relationships

2.5 Model different types of attributes, entities, relationships, and cardinalities

2.6 Convert many-to-many relationships to associative entities

2.7 Model time-dependent data using time stamps

E-R Model Constructs

1. **Entities:**
   - Entity instance – person, place, object, event, concept (often corresponds to a row in a table)
   - Entity Type – collection of entities (often corresponds to a table)
2. **Relationships:**
   - Relationship instance – link between entities
   - Relationship type – category of relationship; link between entity types (unary, binary, ternary, & n-nary)
3. **Attributes:**
   - Properties or characteristics of an entity or relationship type (often corresponds to a field in a table)

Figure 2-1 Sample E-R Diagram

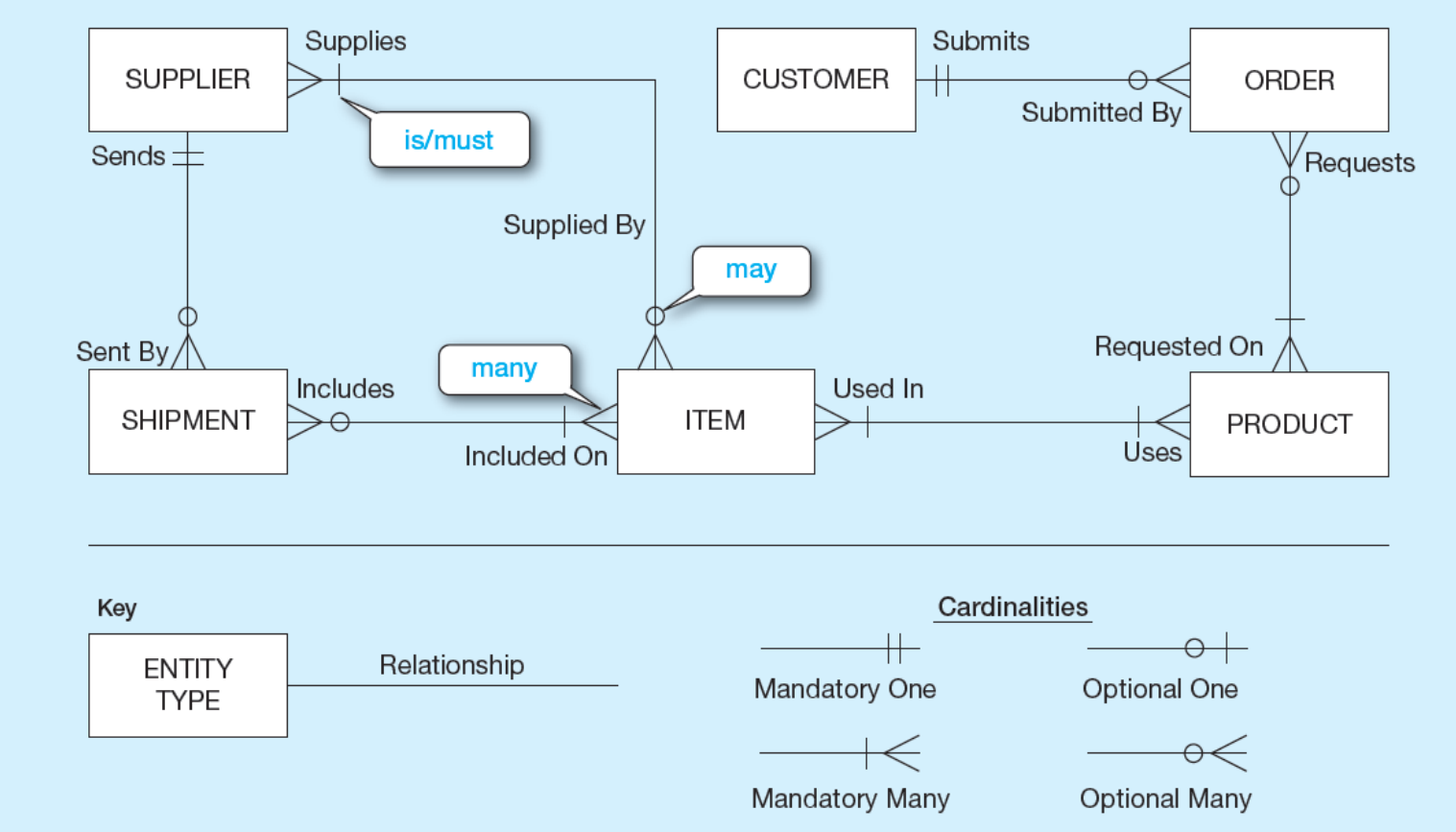
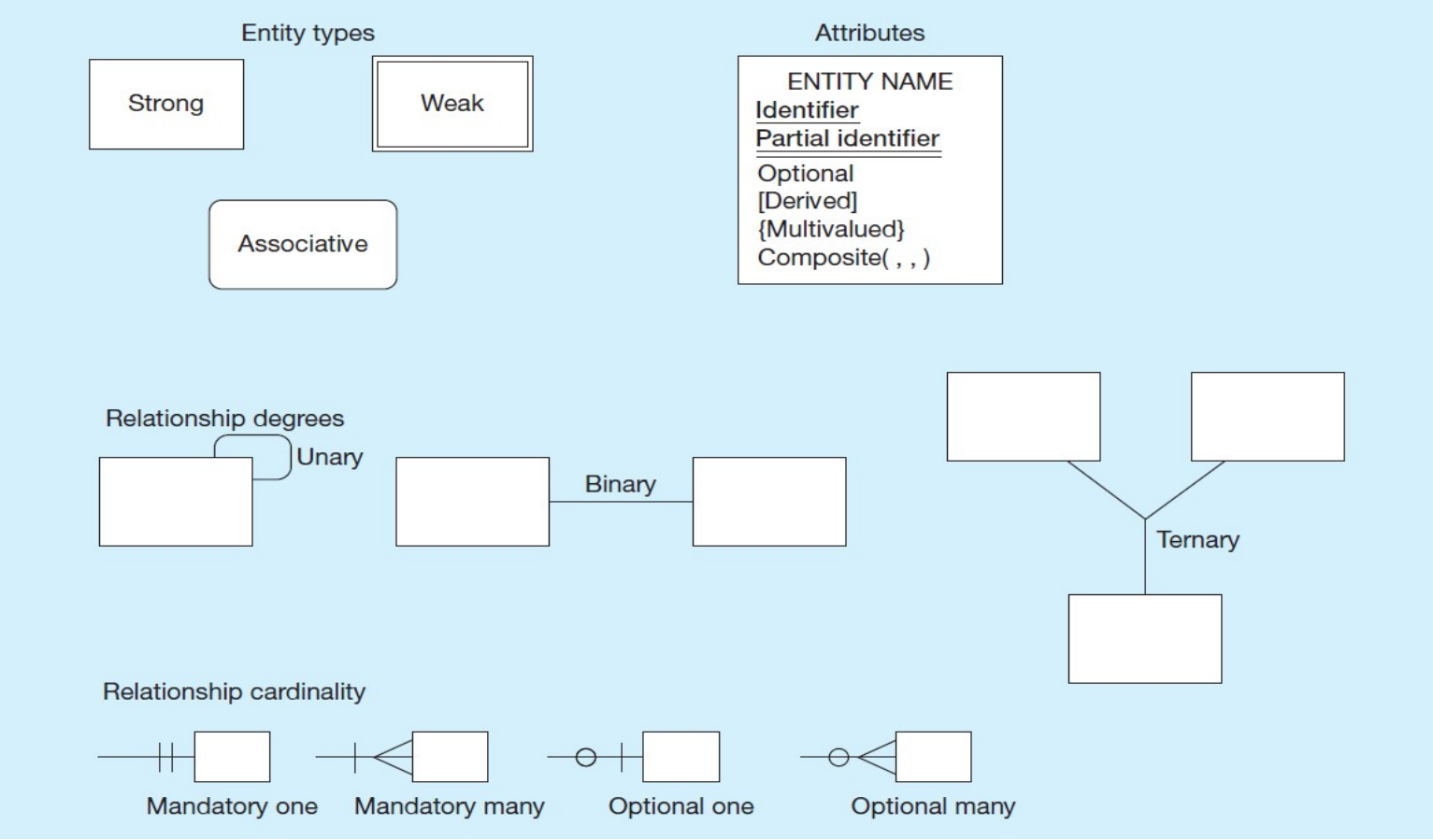


Figure 2-2 Basic E-R Notation



Business Rules

- Are statements that define or constrain some aspect of the business
- Are derived from policies, procedures, events, functions
- Assert business structure
- Control/influence business behavior
- Are expressed in terms familiar to end users
- Are automated through DBMS software

A Good Business Rule Is:

- Declarative – what, not how
- Precise – clear, agreed-upon meaning
- Atomic – one statement
- Consistent – internally and externally
- Expressible – structured, natural language
- Distinct – non-redundant
- Business-oriented – understood by business people

A Good Data Name is:

- Related to business, not technical, characteristics
- Meaningful and self-documenting
- Unique
- Readable
- Composed of words from an approved list
- Repeatable
- Written in standard syntax

Entities

- **Entity**
  - a person, a place, an object, an event, or a concept in the user environment about which the organization wishes to maintain data
- **Entity type**
  - a collection of entities that share common properties or characteristics
- **Entity instance**
  - a single occurrence of an entity type

Entity Type and Entity Instances

FIGURE 2-3 Entity type EMPLOYEE with two instances

Entity type: EMPLOYEE

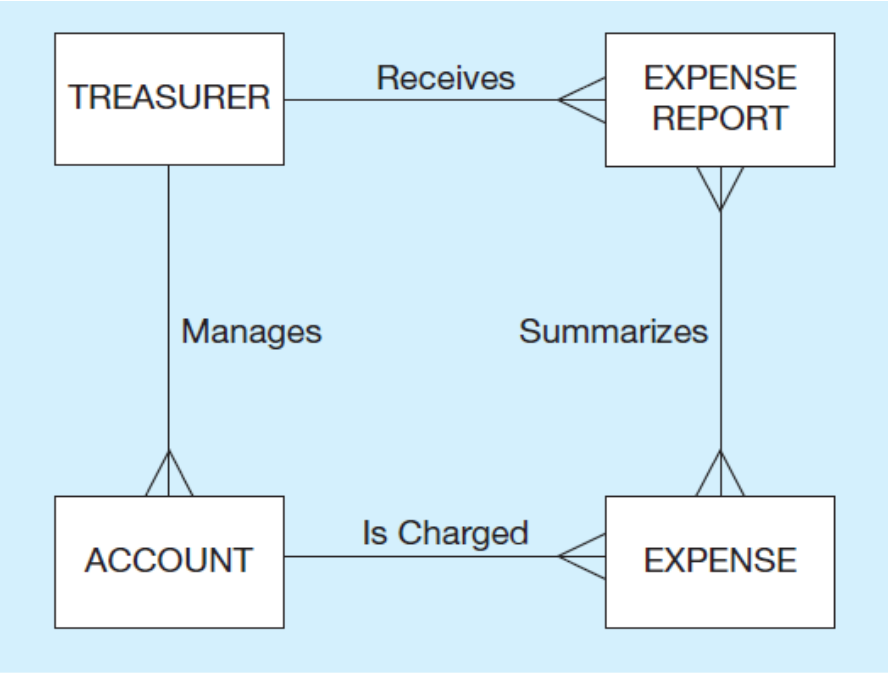
Attributes	Attribute Data Type	Example Instance	Example Instance
Employee Number	CHAR (10)	64217836	53410197
Name	CHAR (25)	Michelle Brady	David Johnson
Address	CHAR (30)	100 Pacific Avenue	450 Redwood Drive
City	CHAR (20)	San Francisco	Redwood City
State	CHAR (2)	CA	CA
Zip Code	CHAR (9)	98173	97142
Date Hired	DATE	03-21-1992	08-16-1994
Birth Date	DATE	06-19-1968	09-04-1975

An Entity…

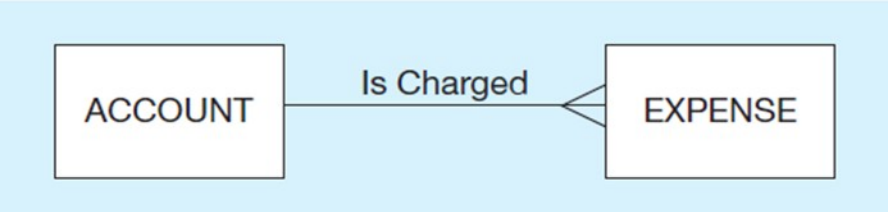
- **Should Be**:
  - An object that will have many instances in the database
  - An object that will be composed of multiple attributes
  - An object that we are trying to model
- **Should Not Be**:
  - A user of the database system
  - An output of the database system (e.g., a report)

Figure 2-4 Example of Inappropriate Entities

(a) System user (Treasurer) and output (Expense Report) shown as entities



b) E-R diagram with only the necessary entities



Problem 1 (Identify Components):

A. A book is identified by its I S B N, and it has a title, a price, and a date of publication. It is published by a publisher, each of which has its own I D number and a name. Each book has exactly one publisher, but one publisher typically publishes multiple books over time.

Note: This solution assumes that we have a reason to track a Publisher even if it does not yet have a Book published.

Problem 2 (Identify Components):

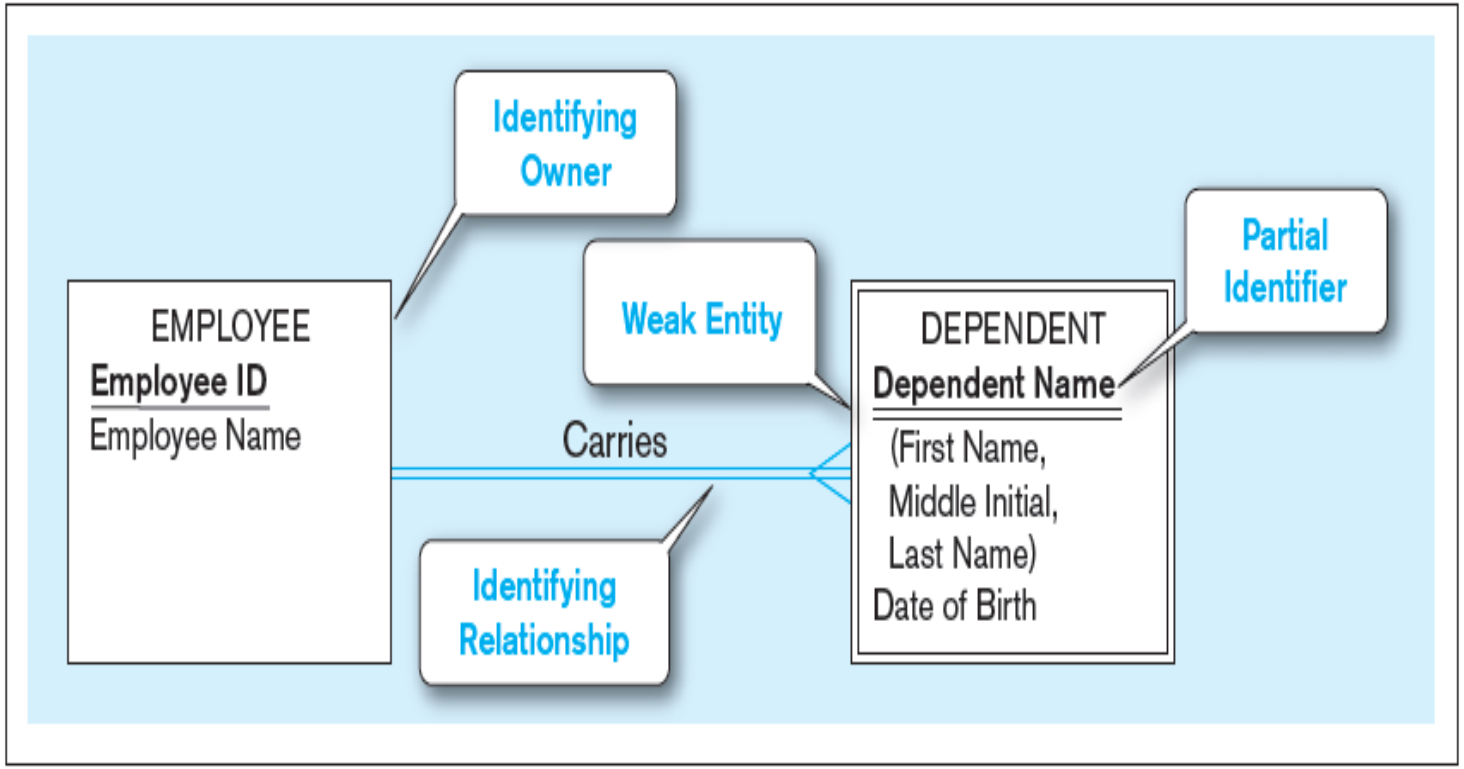
B. A book (see above in (a)) is written by one or multiple authors. Each author is identified by an author number and has a name and date of birth. Each author has either one or multiple books; in addition, occasionally data are needed also regarding prospective authors who have not yet published any books.

Strong vs. Weak Entities, and Identifying Relationships

- **Strong entity**
  - exists independently of other types of entities
  - has its own unique identifier
    - identifier underlined with single line
- **Weak entity**
  - dependent on a strong entity (identifying owner); cannot exist on its own
  - does not have a unique identifier (only a partial identifier)
  - entity box and partial identifier have double lines
- **Identifying relationship**
  - links strong entities to weak entities

Figure 2-5 Example of a Weak Identity and Its Identifying Relationship

Weak entity with identifying relationship.



Guidelines for Naming Entities

- Singular noun
- Specific to organization
- Concise, or abbreviation
- For event entities, the result not the process
- Name consistent for all diagrams

Guidelines for Defining Entities

- “An X is…”
- Describe unique characteristics of each instance
- Explicit about what is and is not the entity
- When an instance is created or destroyed
- Changes to other entity types
- History that should be kept

Attributes

- **Attribute** – property or characteristic of an entity or relationship type
- **Classifications of attributes:**
  - Required versus Optional
  - Simple versus Composite
  - Single-Valued versus Multivalued
  - Stored versus Derived
  - Identifier

Figure 2-6 Required vs. Optional Attributes

Entity type: STUDENT

Attributes	Attribute Data Type	Required or Optional	Example Instance	Example Instance
Student ID	CHAR (10)	Required	28-618411	26-844576
Student Name	CHAR (40)	Required	Michael Grant	Melissa Kraft
Home Address	CHAR (30)	Required	314 Baker St.	1422 Heft Ave
Home City	CHAR (20)	Required	Centerville	Miami
Home State	CHAR (2)	Required	OH	FL
Home Zip Code	CHAR (9)	Required	45459	33321
Major	CHAR (3)	Optional	MIS	

Required – must have a value for every entity (or relationship) instance with which it is associated

Optional – may not have a value for every entity (or relationship) instance with which it is associated

Figure 2-7 A Composite Attribute

- **Composite attribute** – An attribute that has meaningful component parts (sub-attributes)

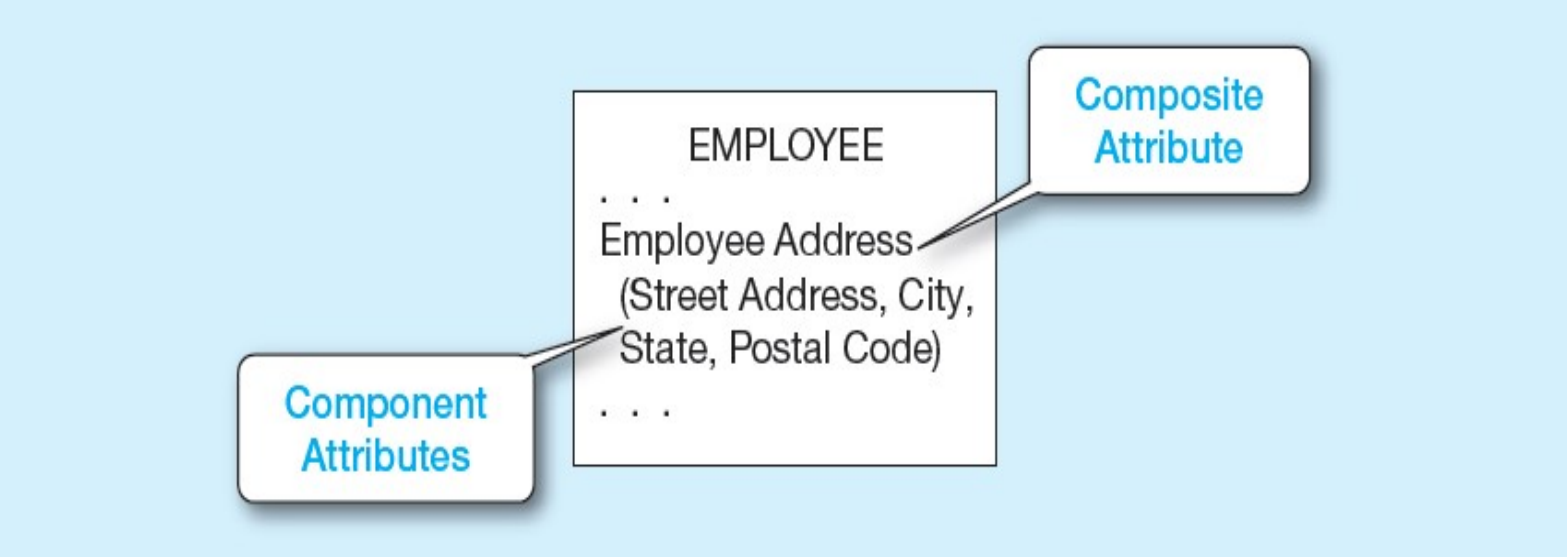
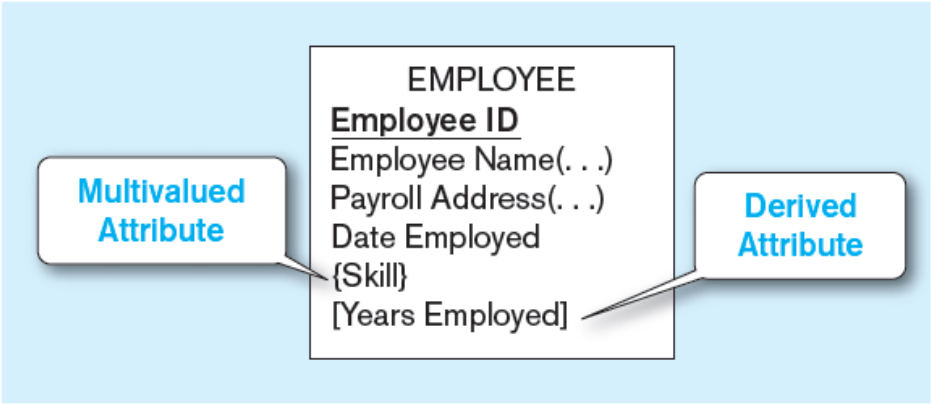


Figure 2-8 Multivalued and Derived Attributes

- **Multivalued**
  - May take on more than one value for a given entity (or relationship) instance
  - An employee can have more than one skill
- **Derived**
  - Values can be calculated from related attribute values (not physically stored in the database)
  - Years employed calculated from date employed and current date



Identifiers (Keys)

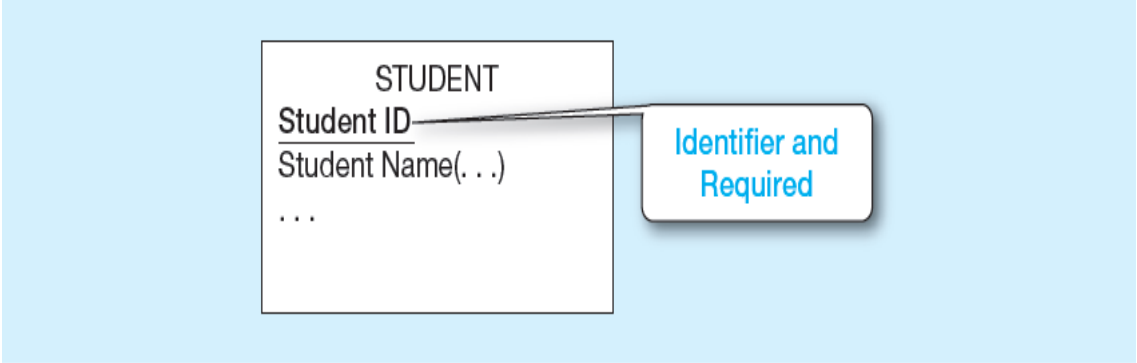
- **Identifier (Key)** – an attribute (or combination of attributes) that uniquely identifies individual instances of an entity type
- **Simple versus Composite Identifier**
- **Candidate Identifier** – an attribute that could be an identifier; it satisfies the requirements for being an identifier

Criteria for Identifiers

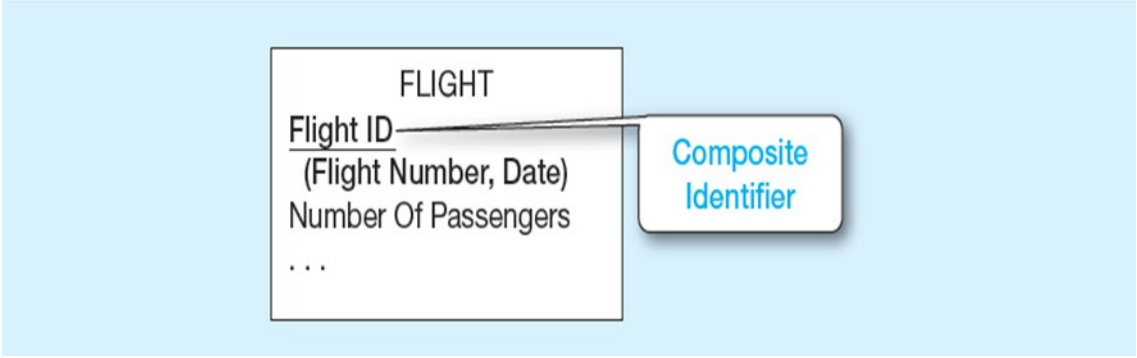
- Choose Identifiers that
  - Will not change in value
  - Will not be null
- Avoid intelligent identifiers (e.g., containing locations or people that might change)
- Substitute new, simple keys for long, composite keys

Figure 2-9 Simple and Composite Identifier Attributes

a) Simple identifier attribute



b) Composite identifier attribute



Naming Attributes

- Name should be a singular noun or noun phrase
- Name should be unique
- Name should follow a standard format
- Similar attributes of different entity types should use the same qualifiers and classes

Defining Attributes (1 of 2)

- State what the attribute is and possibly why it is important
- Make it clear what is and is not included in the attribute's value
- Include aliases in documentation

Defining Attributes (2 of 2)

- State source of values
- State whether attribute value can change once set
- Specify whether required or optional
- State min and max number of occurrences allowed
- Indicate relationships with other attributes

Problem 1 (Draw Entities):

A. A book is identified by its I S B N, and it has a title, a price, and a date of publication. It is published by a publisher, each of which has its own I D number and a name. Each book has exactly one publisher, but one publisher typically publishes multiple books over time.

Note: This solution assumes that we have a reason to track a Publisher even if it does not yet have a Book published.

Problem 2 (Draw Entities):

B. A book (see above in (a)) is written by one or multiple authors. Each author is identified by an author number and has a name and date of birth. Each author has either one or multiple books; in addition, occasionally data are needed also regarding prospective authors who have not yet published any books.

Modeling Relationships

1. **Relationship Types vs. Relationship Instances**
   - The relationship type is modeled as lines between entity types. The relationship instance is between specific entity instances

2. **Relationships can have attributes**
   - These describe features pertaining to the association between the entities in the relationship

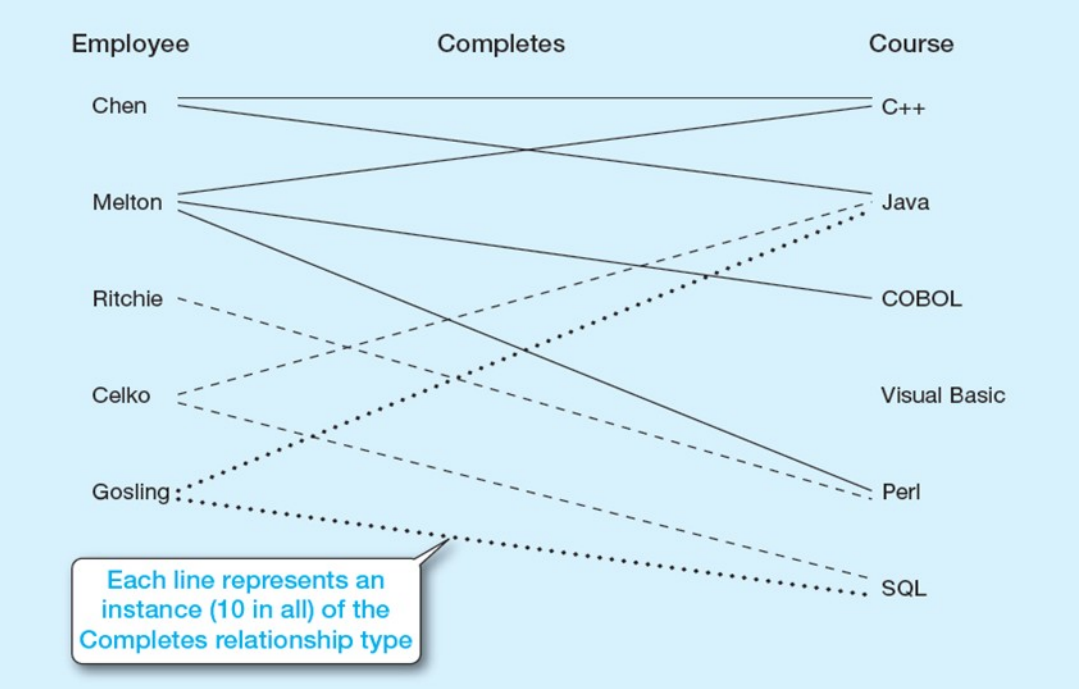
3. **Two entities can have more than one type of relationship between them** (multiple relationships)

4. **Associative Entity** – combination of relationship and entity

1. Relationship Type and Instances

a) Relational type (Completes)

b) Relationship instances



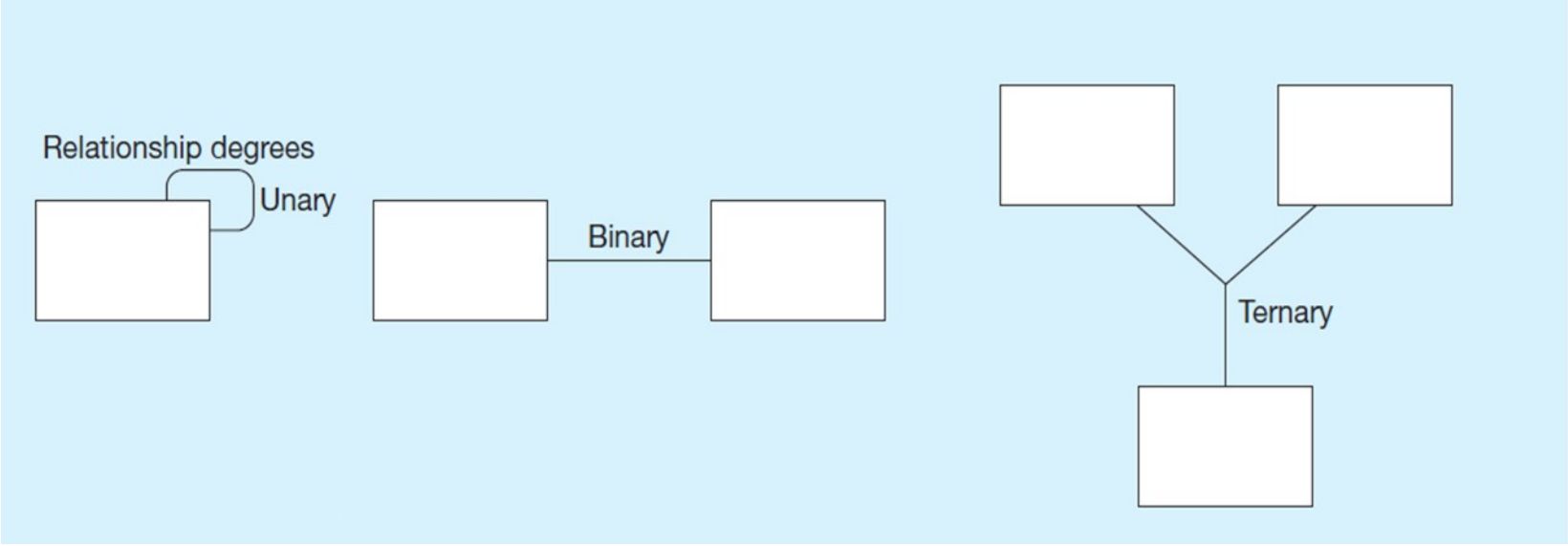
Degrees of Relationships

- Degree of a relationship is the number of entity types that participate in it

Unary – entities of the same entity type related to each other

Binary – entities of one type related to entities of another

Ternary – entities of three different types involved in the same relationship



Instances & Cardinality of Relationships

- **One-to-One**
  - Each entity in the relationship will have exactly one related entity
- **One-to-Many**
  - An entity on one side of the relationship can have many related entities, but an entity on the other side will have a maximum of one related entity
- **Many-to-Many**
  - Entities on both sides of the relationship can have many related entities on the other side

Figure 2-12 Examples of Relationships of Different Degrees (1 of 3)

a) Unary relationships

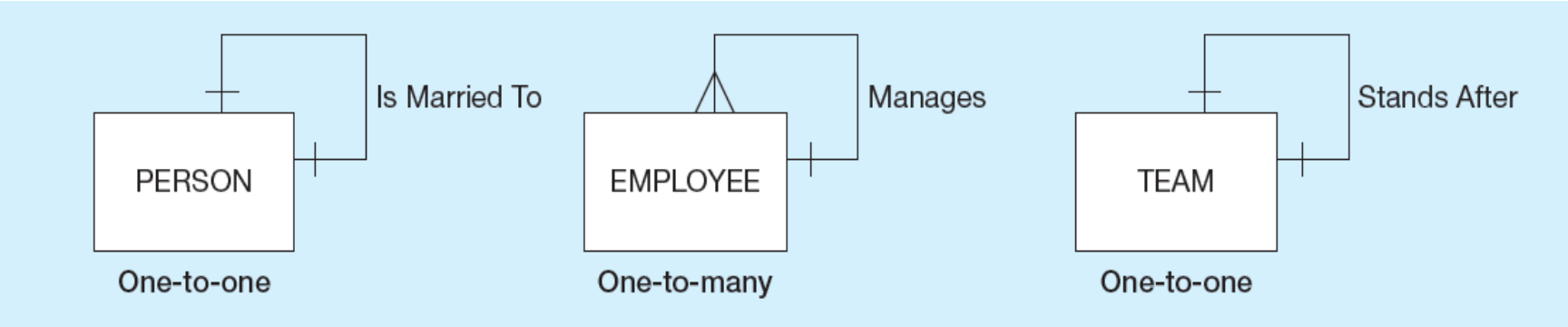


Figure 2-12 Examples of Relationships of Different Degrees (2 of 3)

b) Binary relationships

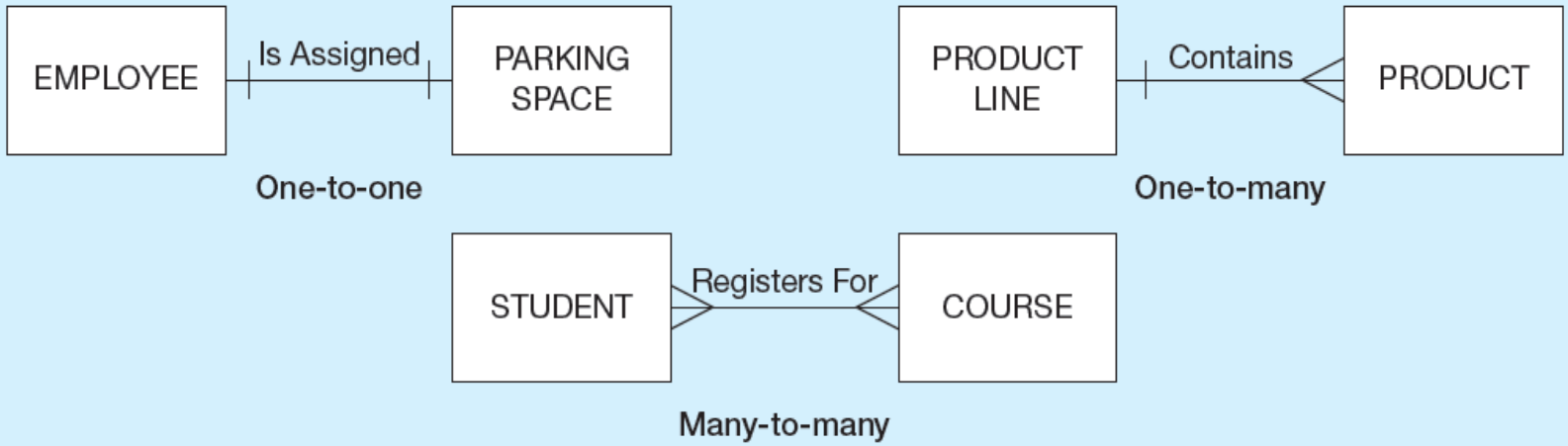
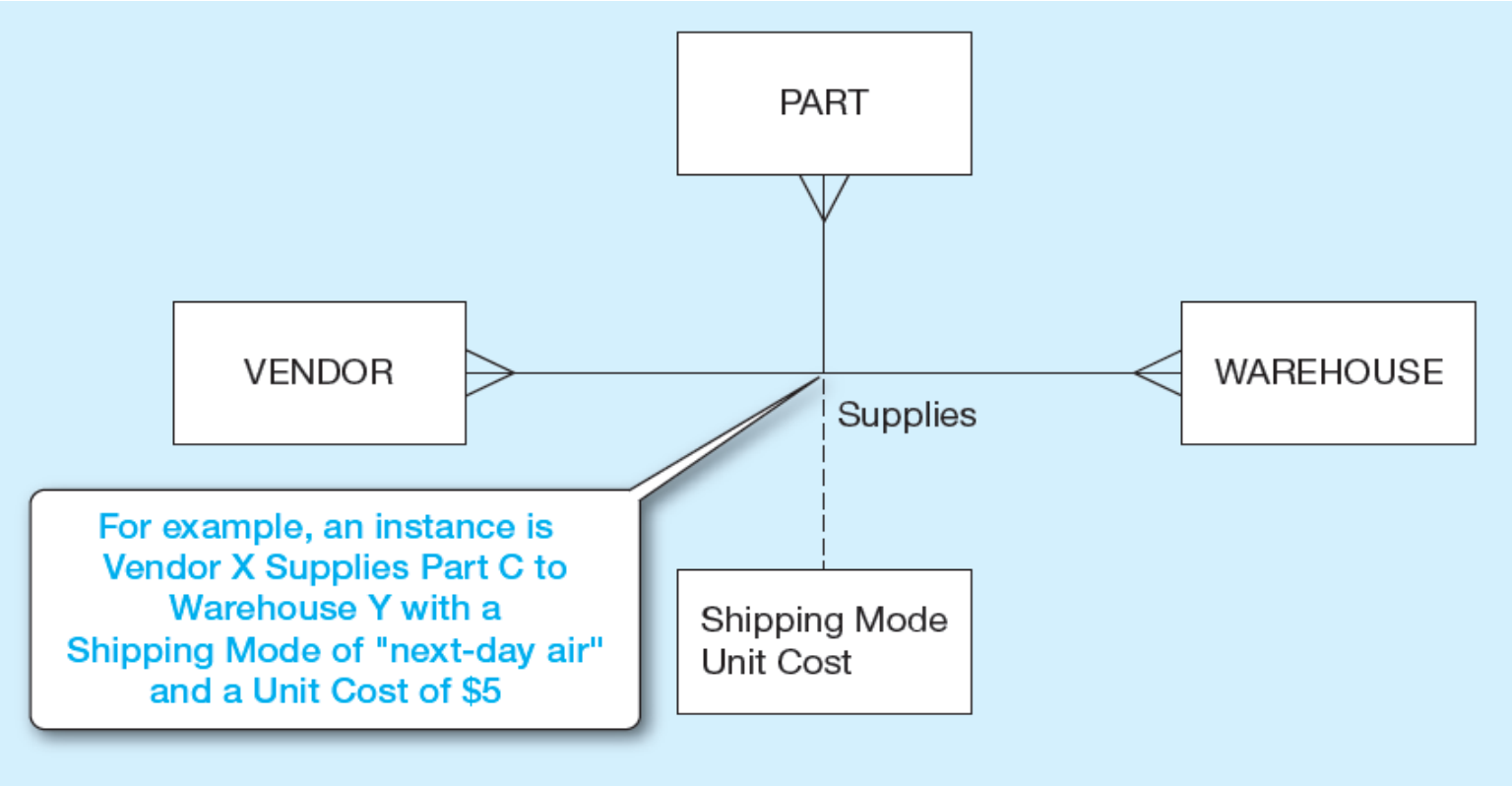


Figure 2-12 Examples of Relationships of Different Degrees (3 of 3)

c) Ternary relationships



Cardinality Constraints

- Cardinality Constraints — the number of instances of one entity that can or must be associated with each instance of another entity
- Minimum Cardinality
  - If zero, then optional
  - If one or more, then mandatory
- Maximum Cardinality
  - The maximum number

Figure 2-17 Examples of Cardinality Constraints (1 of 3)

a) Mandatory cardinalities

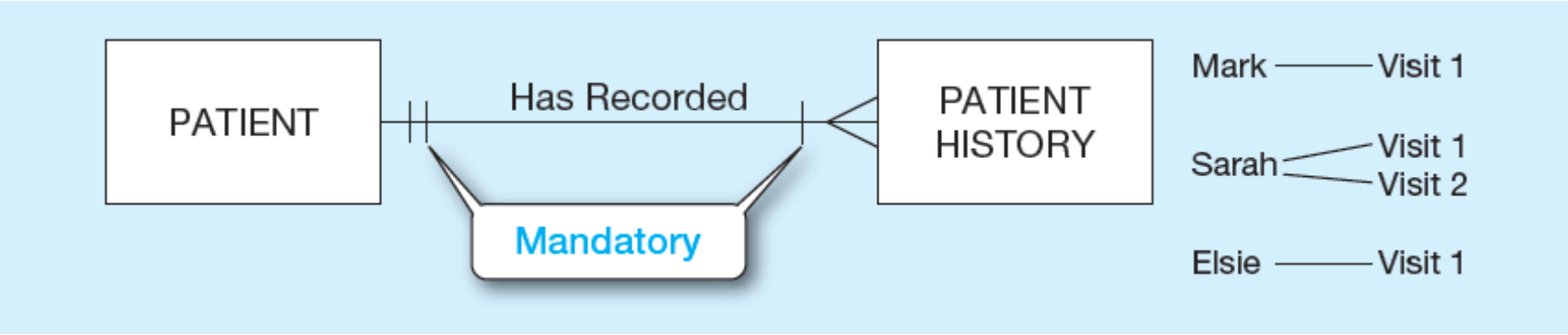


Figure 2-17 Examples of Cardinality Constraints (2 of 3)

b) One mandatory, one optional cardinality

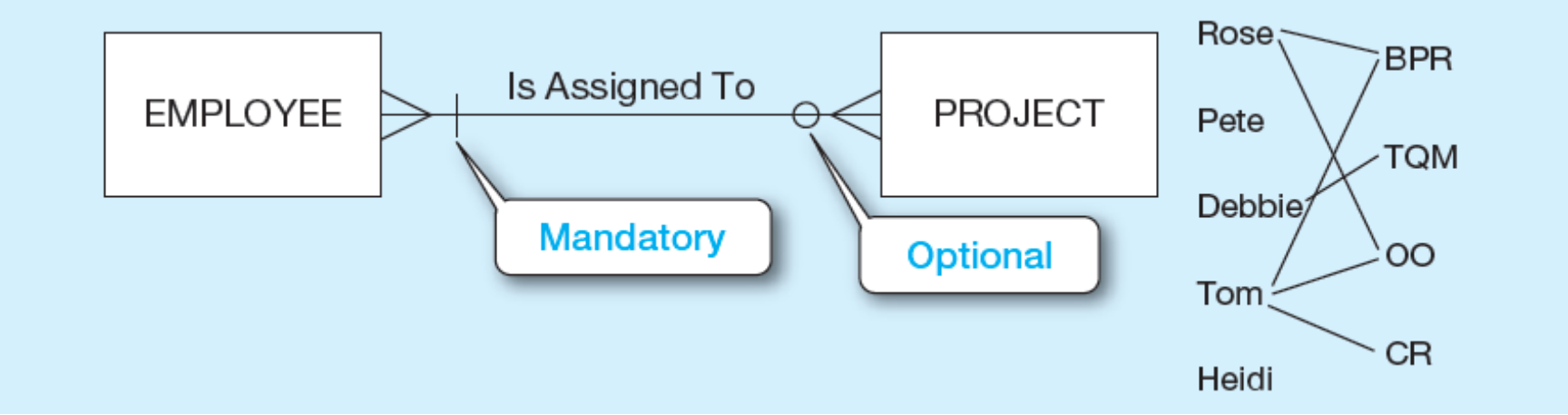
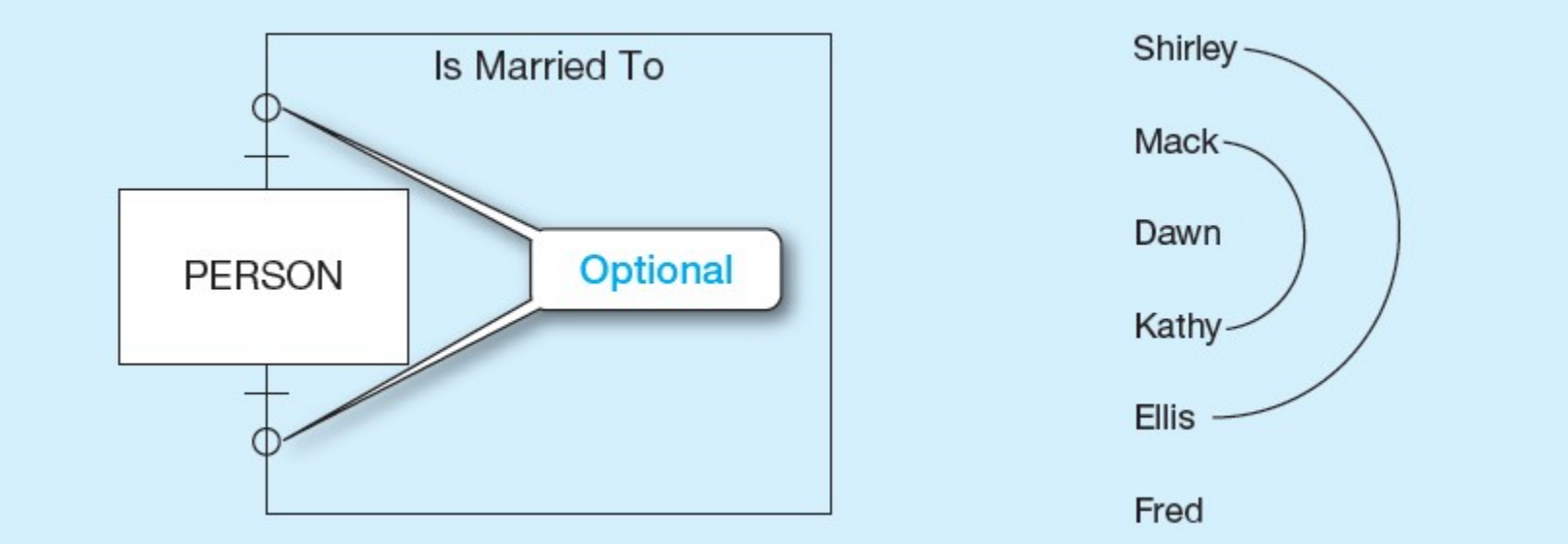


Figure 2-17 Examples of Cardinality Constraints (3 of 3)

c) Optional cardinalities



3. Example of Multiple Relationships (1 of 2)

a) Employees and departments

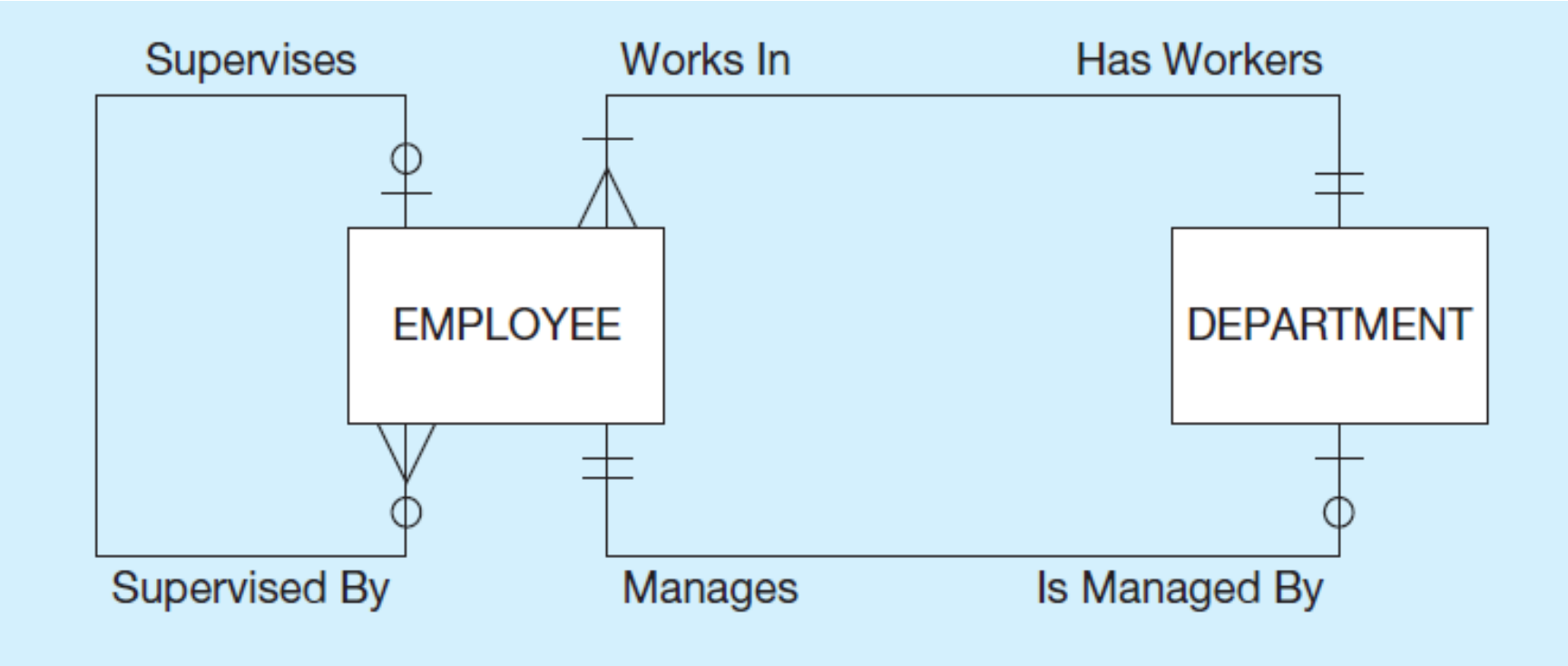
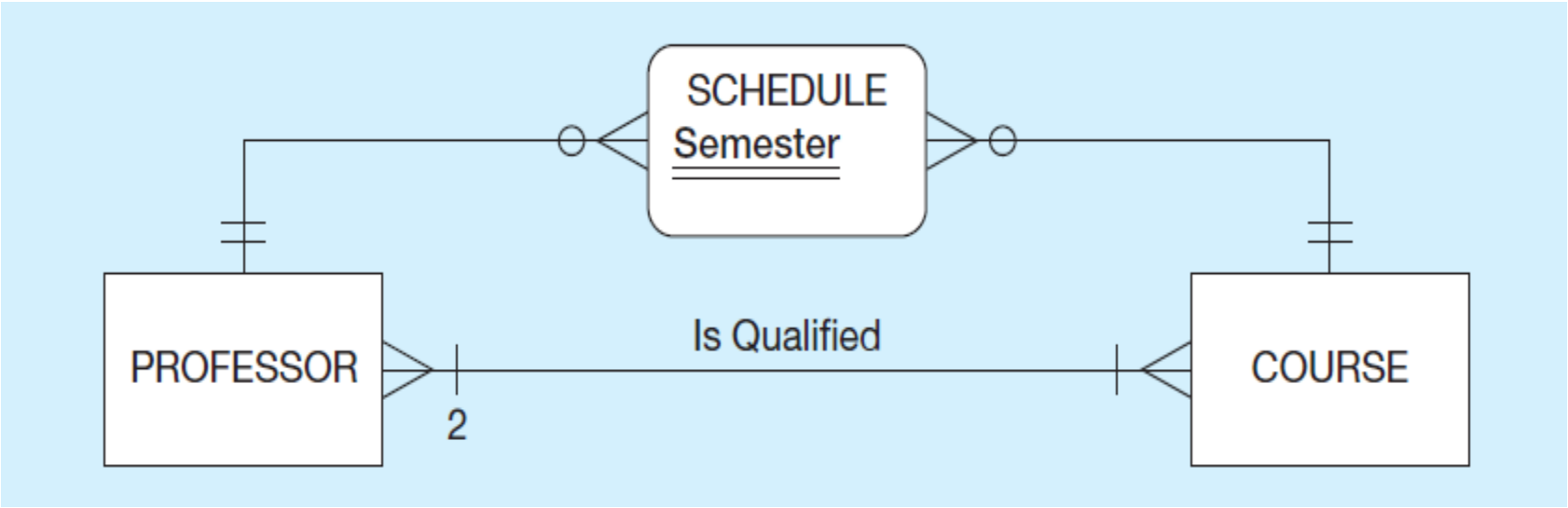


Figure 2-21 Example of Multiple Relationships (2 of 2)

b) Professors and courses (fixed lower limit constraint)



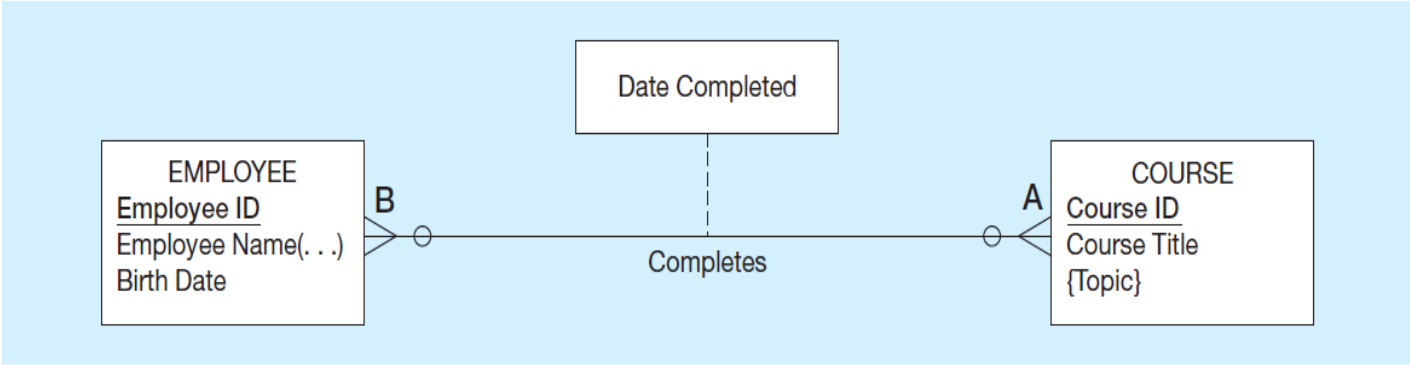
4. Associative Entities

- When should a **relationship with attributes** instead be an **associative entity**?
  - All relationships for the associative entity should be many
  - The associative entity could have meaning independent of the other entities
  - The associative entity preferably has a unique identifier, and should also have other attributes
  - The associative entity may participate in other relationships other than the entities of the associated relationship
  - Convert ternary relationships to associative entities

Figure 2-11 Associative Entities

An associative entity is an entity. It has attributes. It is also a relationship. It serves to link other entities together in a many-to-many relationship.

(a) Attribute on a relationship



(b) An associative entity (CERTIFICATE)

Figure 2-13 Representing a Bill-of-Materials Structure (1 of 2)

Two ways to represent a bill-of-materials structure

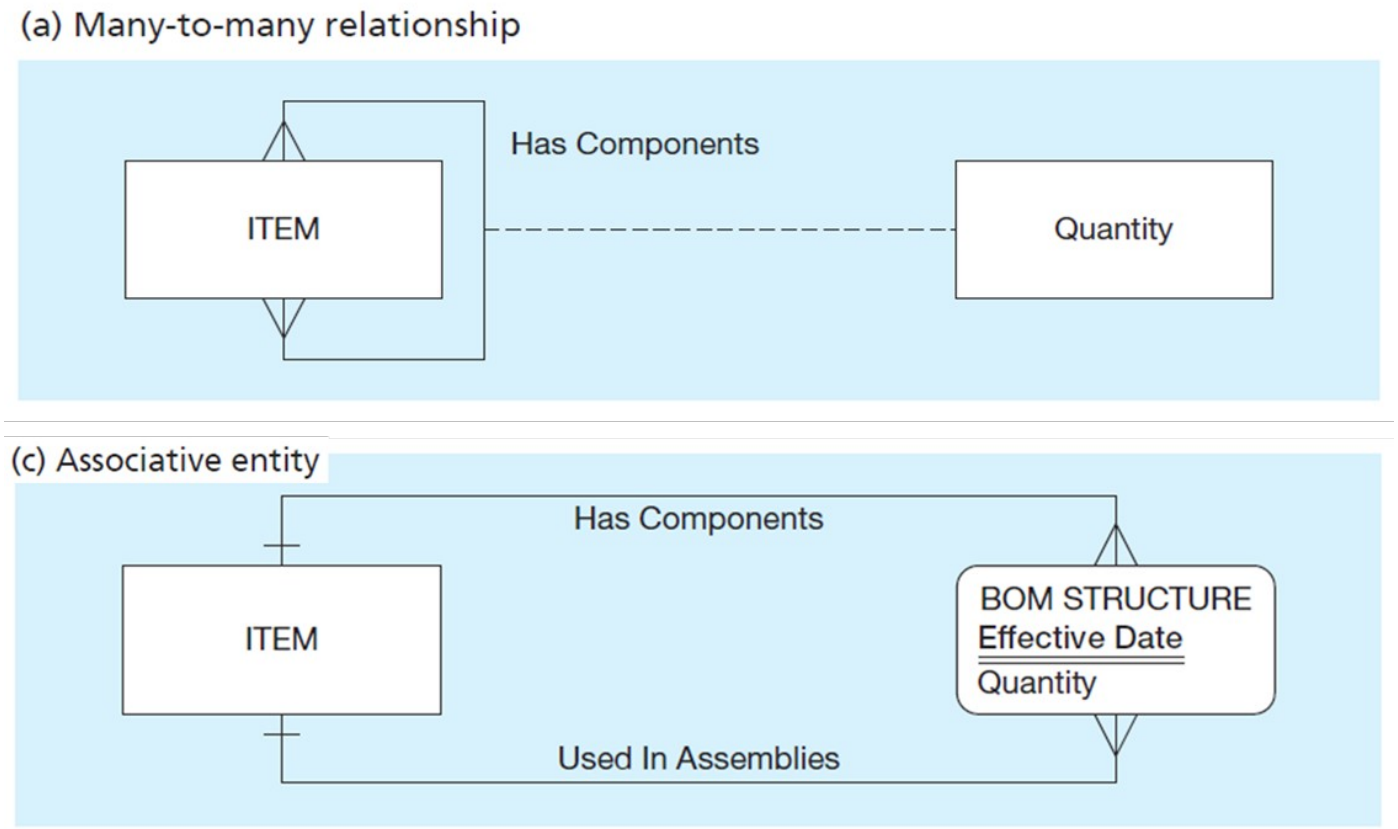


Figure 2-18 Cardinality Constraints in a Ternary Relationship

The three relationships are marked by their mandatory or optional constraints. #1 states that a vendor may or may not supply parts to warehouses. #2 states that each part must come from a vendor and go to a warehouse. #3 says a warehouse must have at least one part from a vendor.

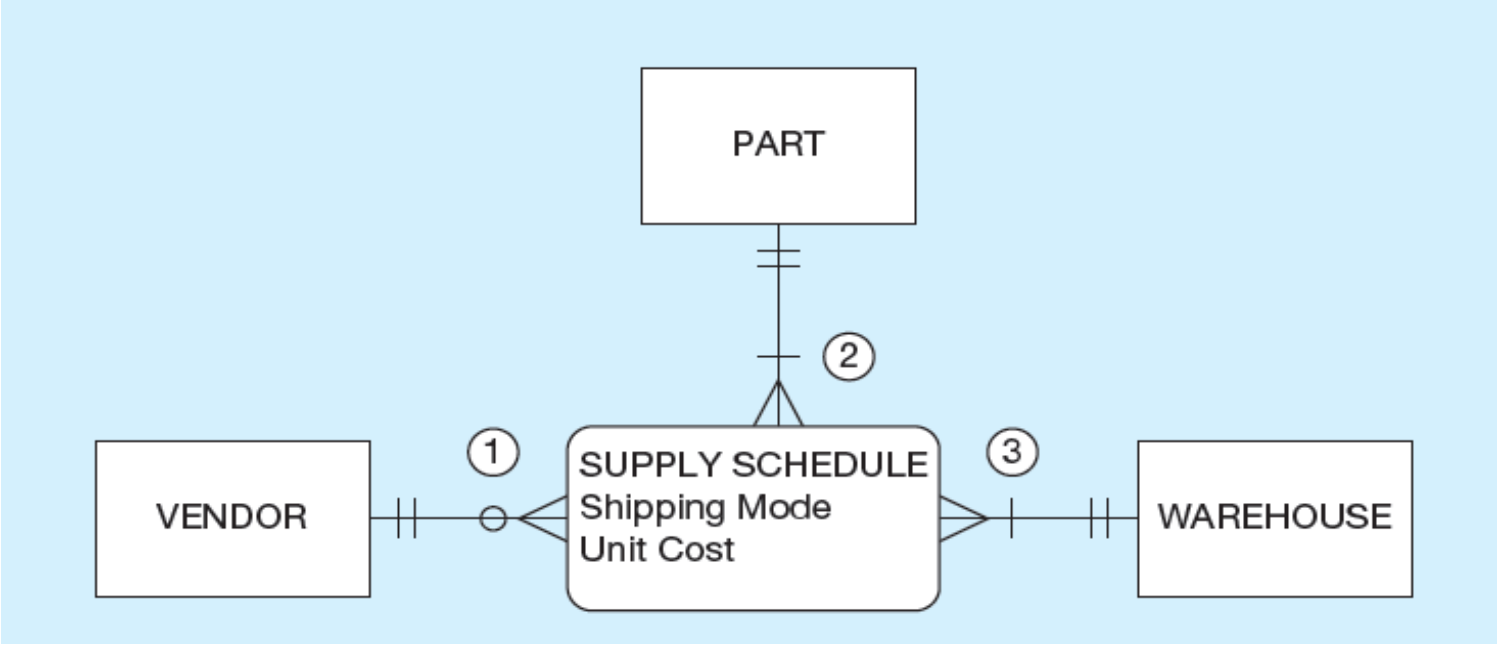


Figure 2-19 Simple Example of Time-Stamping

Time stamp – a time value that is associated with a data value, often indicating when some event occurred that affected the data value.

The Price History attribute is both multivalued and composite.