MODELING DATA IN THE ORGANISATION

Modern Database Management

11th Edition

Jeffrey A. Hoffer, V. Ramesh,

Heikki Topi

Learning Outcomes

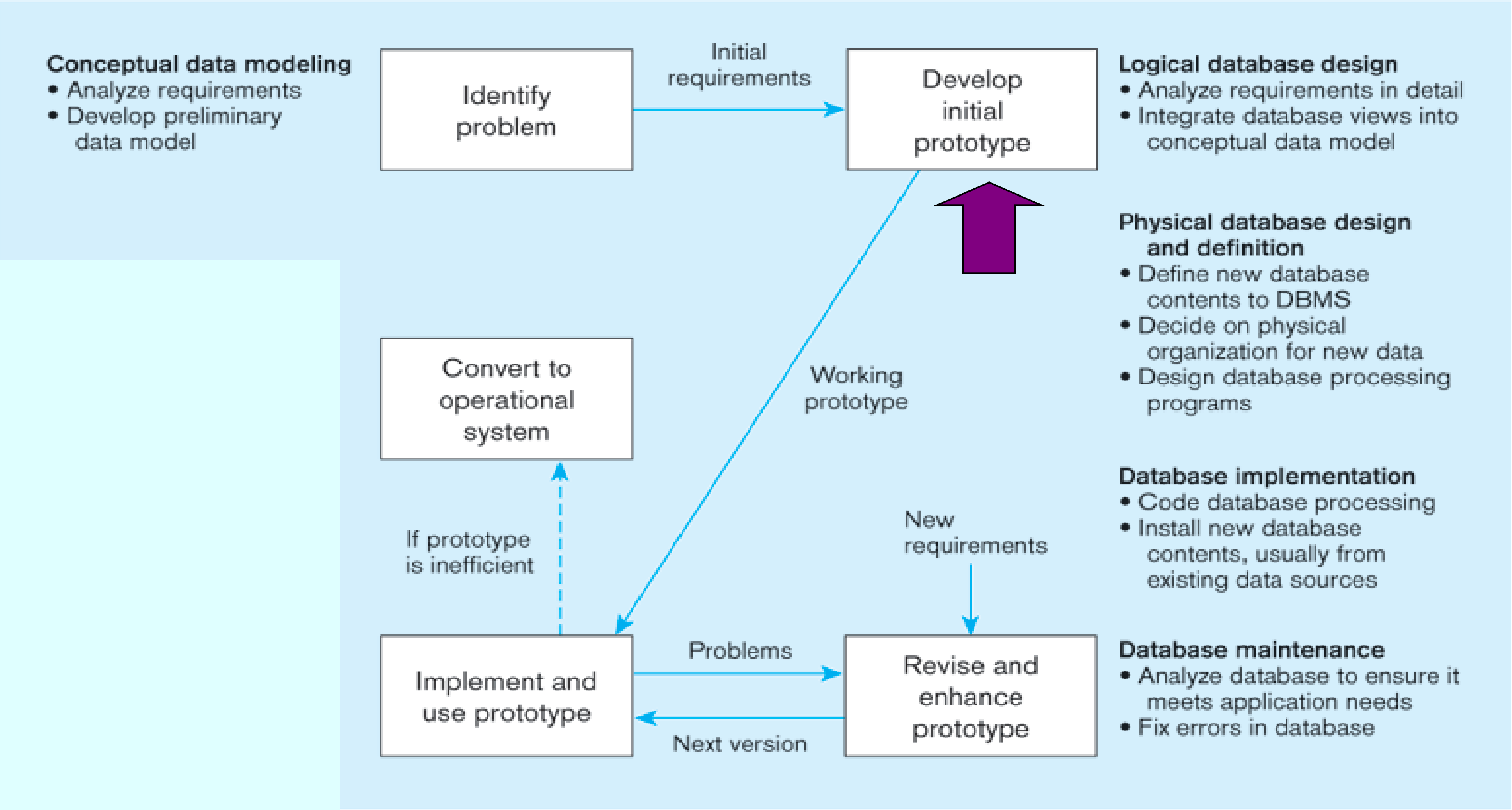
- Prototyping
- Understand importance of data modelling
- Write good names and definitions for entities, relationships, and attributes
- Distinguish unary, binary, and ternary relationships
- Model different types of attributes, entities, relationships, and cardinalities
- Draw E-R diagrams for common business situations

Two approaches to databases and its development

- **SDLC**
  - System Development Life Cycle
  - Detailed, well-planned development process
  - Time-consuming, but comprehensive
  - Long development cycle

- **Prototyping**
  - Rapid application development (RAD)
  - Cursory attempt at conceptual data modeling
  - Define database during development of initial prototype
  - Repeat implementation and maintenance activities with new prototype versions

Prototyping Database Methodology (Figure 1-8) (cont.)



Conceptual data modeling

- Analyze requirements
- Develop preliminary data model

Logical database design

- Analyze requirements in detail
- Integrate database views into conceptual data model

Physical database design and definition

- Define new database contents to DBMS
- Decide on physical organization for new data
- Design database processing programs

Database implementation

- Code database processing
- Install new database contents, usually from existing data sources

Database maintenance

- Analyze database to ensure it meets application needs
- Fix errors in database

Identify problem
Initial requirements
Develop initial prototype
Working prototype
Convert to operational system
If prototype is inefficient
New requirements
Implement and use prototype
Problems
Next version
Revise and enhance prototype

A GOOD DATA NAME IS:

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

DATA DEFINITIONS

- Explanation of a term or fact
  - Term–word or phrase with specific meaning
  - Fact–association between two or more terms
- Guidelines for good data definition
  - A concise description of essential data meaning
  - Gathered in conjunction with systems requirements
  - Accompanied by diagrams
  - Achieved by consensus, and iteratively refined

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

http://agilemodeling.com/

Business rules often focus on access control issues, for example, professors are allowed to input and modify the marks of the students taking the seminars they instruct, but not the marks of students in other seminars.

Name: Tenured professors may administer student grades

Identifier: BR123

Description: Only tenured professors are granted the ability to initially input, modify, and delete grades students receive in the seminars that they and they only instruct. They may do so only during the period a seminar is active.

Example: Dr. Bruce, instructor of "Biology 301 Advanced Uses of Gamma Radiation," may administer the marks of all students enrolled in that seminar, but not those enrolled in "Biology 302 Effects of Radiation on Arachnids," which is taught by Dr. Peters.

Source: University Policies and Procedures
Doc ID: U1701
Publication date: August 14, 2000

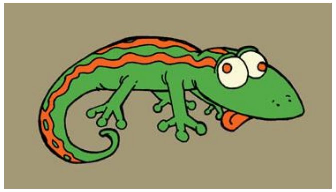
Related rules: BR12 Qualifying For Tenure
BR65 Active Period for Seminars
BR200 Modifying Final Student Grades

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

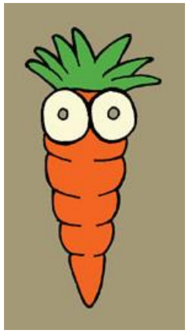
E-R Model Constructs

- 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)
- Relationships:
  - Relationship instance–link between entities (corresponds to primary key-foreign key equivalencies in related tables)
  - Relationship type–category of relationship…link between entity types
- Attributes:
  - Properties or characteristics of an entity or relationship type (often corresponds to a field in a table)

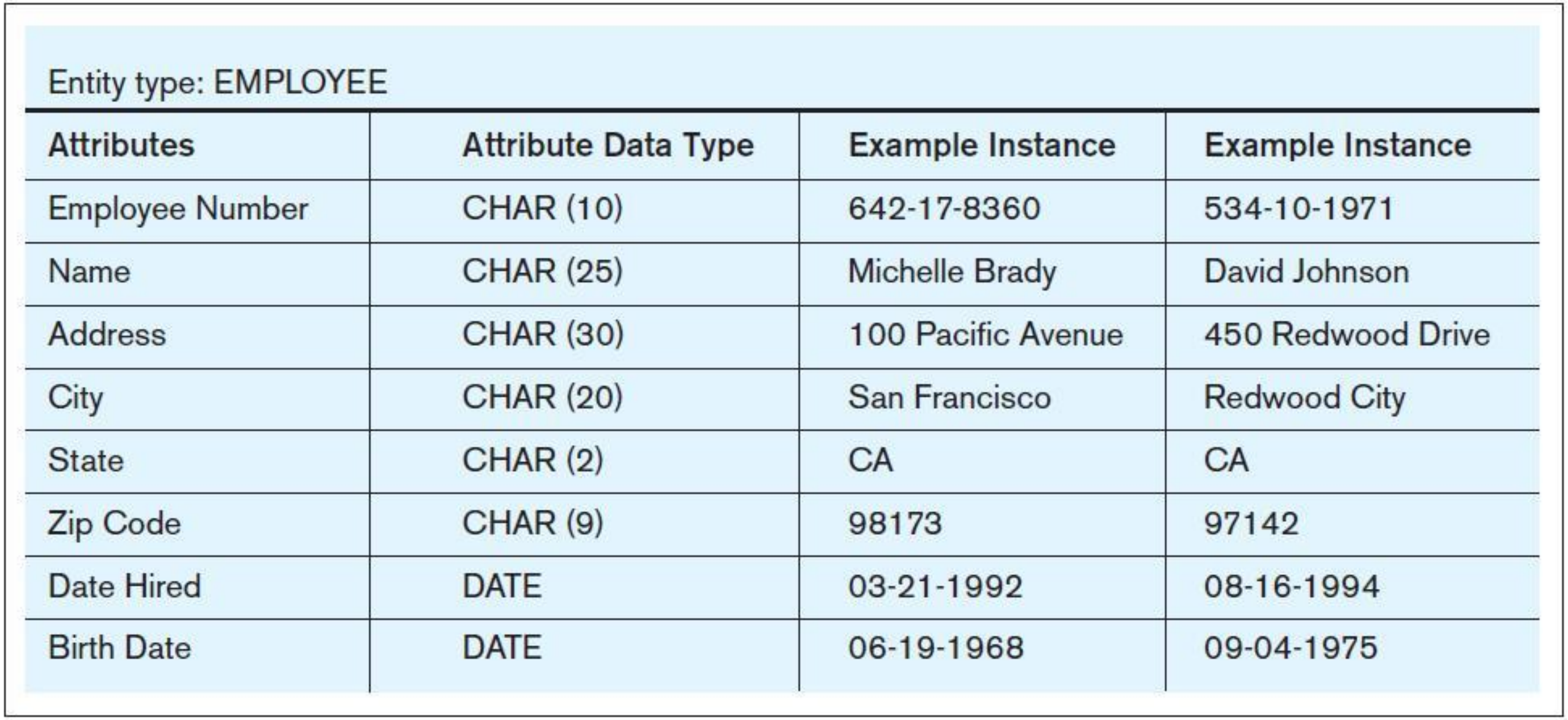


ENTITIES

- **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



Entity type: EMPLOYEE

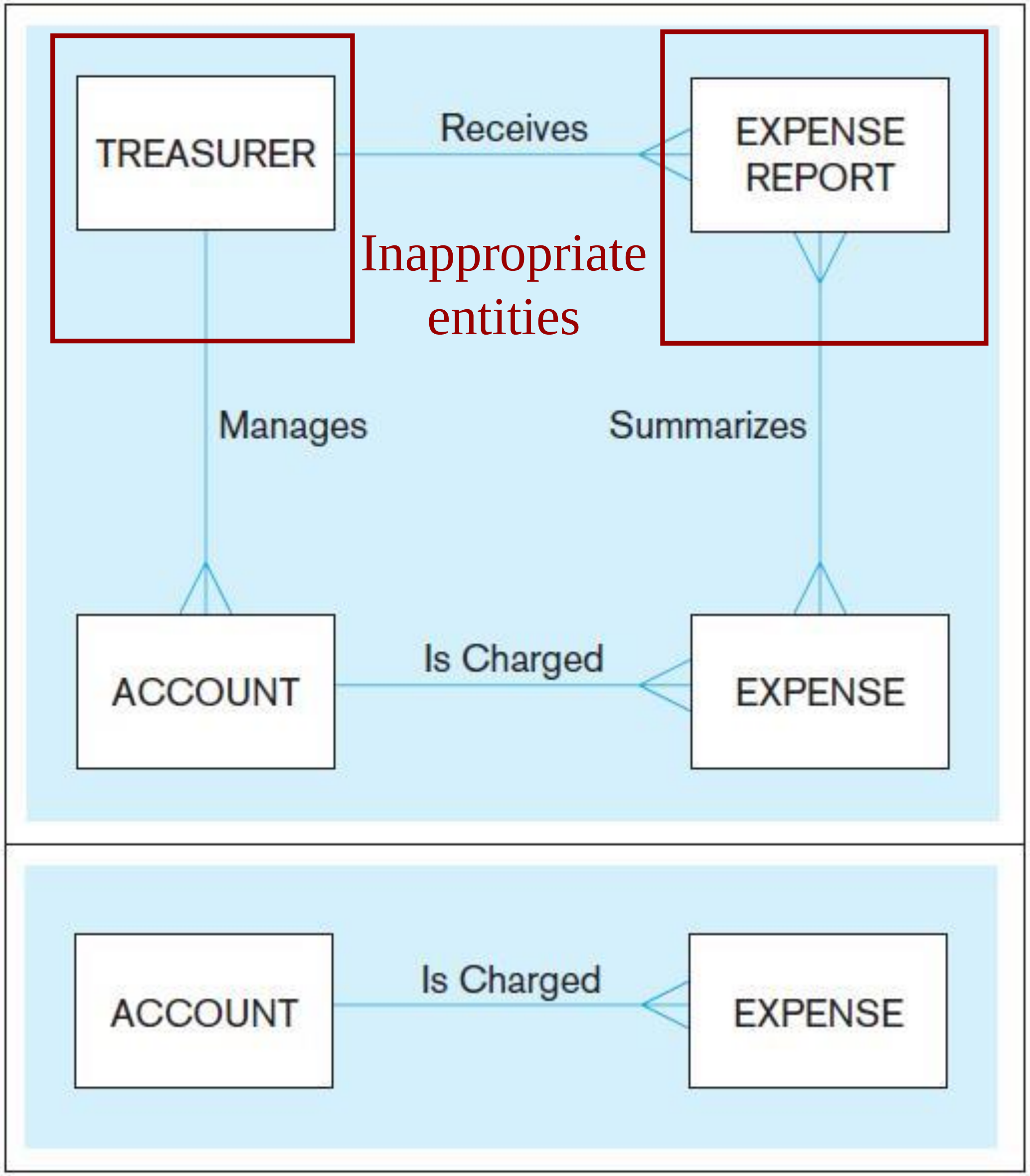
Attributes	Attribute Data Type	Example Instance	Example Instance
Employee Number	CHAR (10)	642-17-8360	534-10-1971
Name	CHAR (25)	Michelle Brady	David Johnson
Address	CHAR (30)	100 Pacific Avenue	450 Redwood Drive
City	CHAR (20)	San Francisco	Redwood City
State	CHAR (2)	CA	CA
Zip Code	CHAR (9)	98173	97142
Date Hired	DATE	03-21-1992	08-16-1994
Birth Date	DATE	06-19-1968	09-04-1975

FIGURE 2-3 Entity type EMPLOYEE with two instances

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



System user

Inappropriate entities

System output

Appropriate entities

ATTRIBUTES

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

NAMING ATTRIBUTES

- Name should be a singular noun or noun phrase
- Name should be unique
- Name should follow a standard format
  - e.g. **[Entity type name { [Qualifier] }] Class**
- Similar attributes of different entity types should use the same qualifiers and classes

DEFINING ATTRIBUTES

- 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
- State source of values
- Specify required vs. optional
- State min and max number of occurrences allowed
- Indicate relationships with other attributes

SIMPLE VS. COMPOSITE ATTRIBUTES

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

The address is broken into component parts

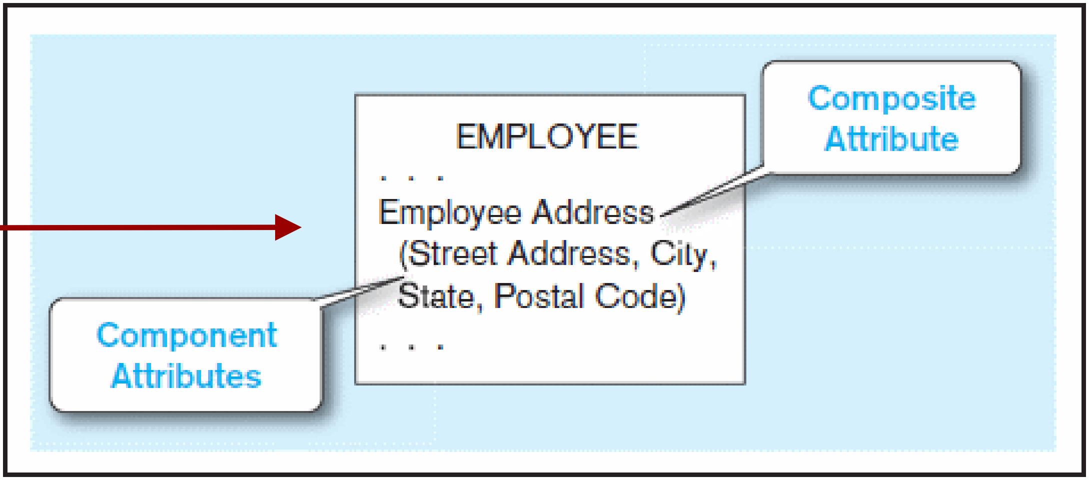
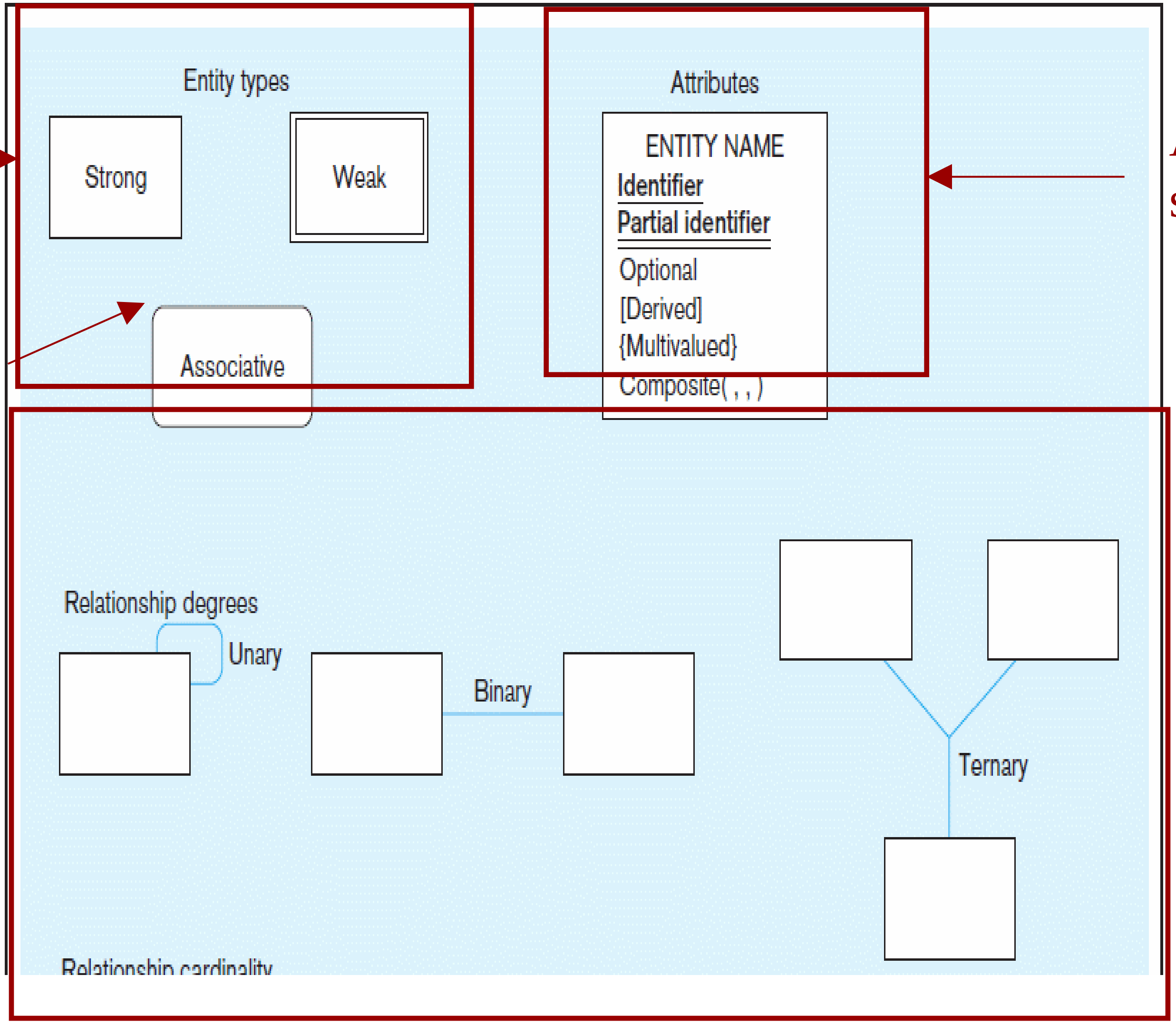


Figure 2-7 A **composite** attribute

Basic E-R notation (Figure 2-2)



Entity symbols

A special entity that is also a relationship

Relationship degrees specify number of entity types involved

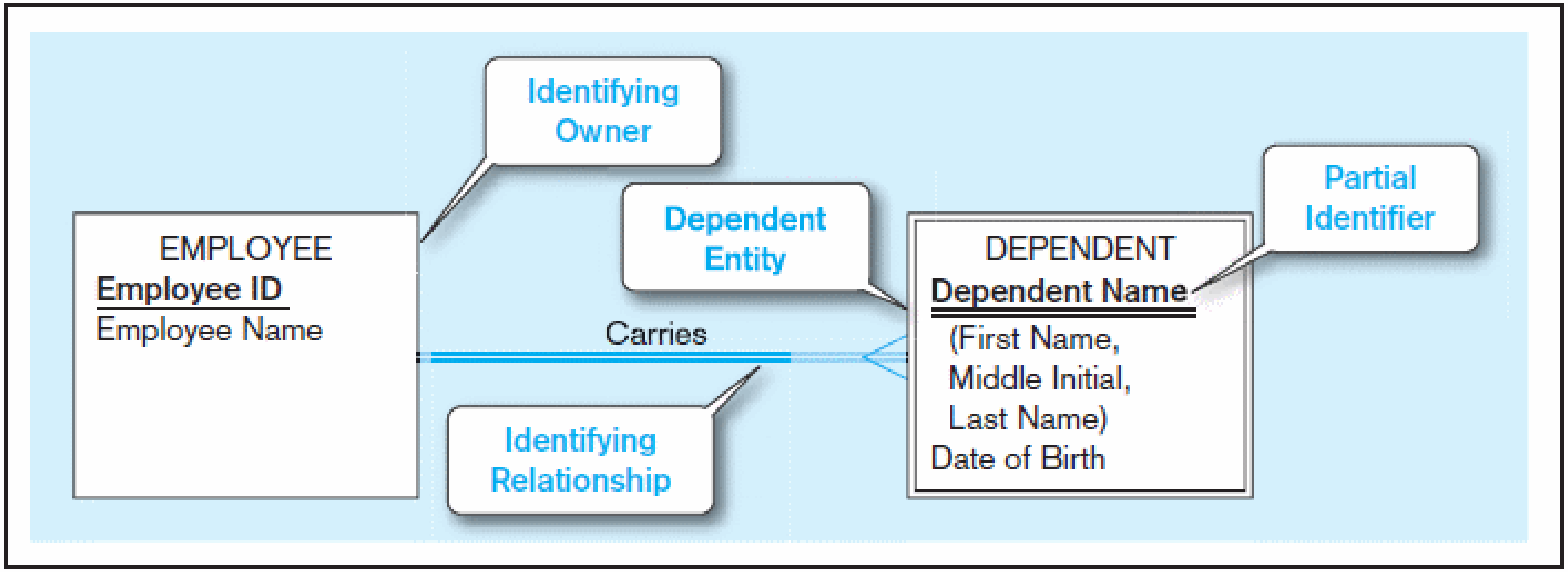
Attribute symbols

Relationship symbols

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

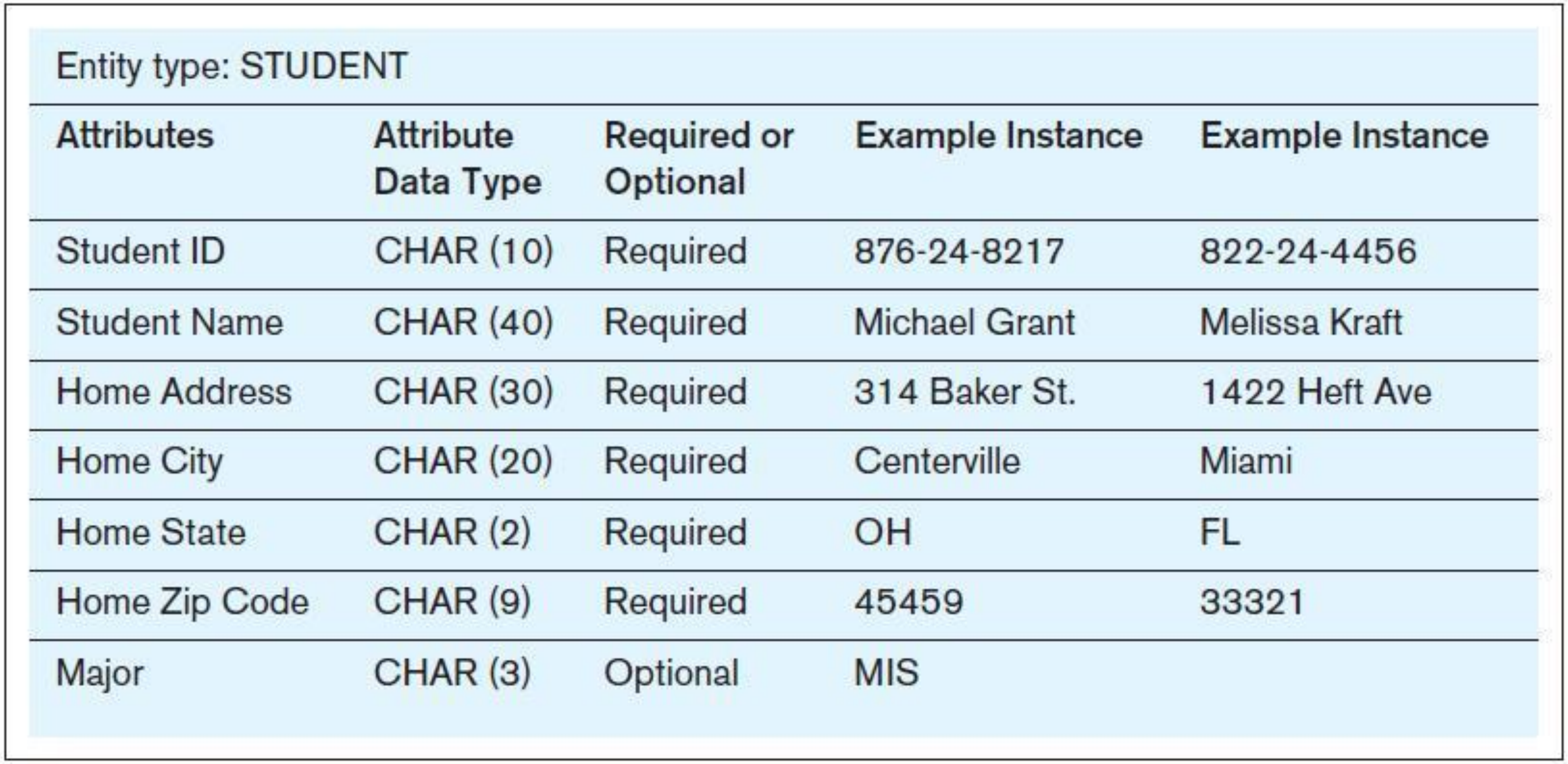
Figure 2-5 Example of a weak identity and its identifying relationship



Strong entity

Weak entity

REQUIRED VS. OPTIONAL ATTRIBUTES



Entity type: STUDENT

Attributes	Attribute Data Type	Required or Optional	Example Instance	Example Instance
Student ID	CHAR (10)	Required	876-24-8217	822-24-4456
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

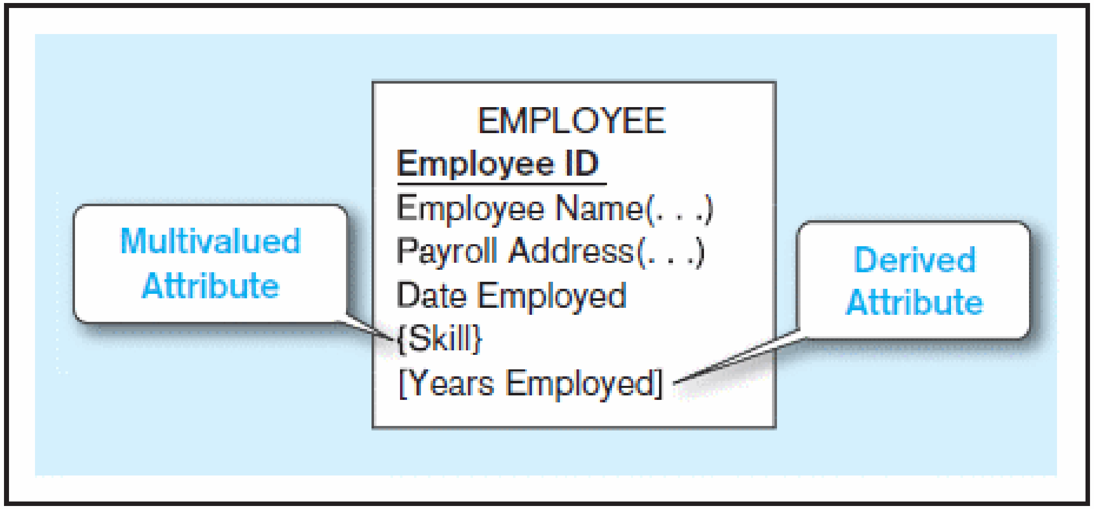
Multi-valued and Derived Attributes

Multivalued – may take on more than one value for a given entity (or relationship) instance

Derived – values can be calculated from related attribute values (not physically stored in the database)

Figure 2-8 Entity with **multivalued** attribute (Skill) and **derived** attribute (Years Employed)

Multivalued
an employee can have more than one skill

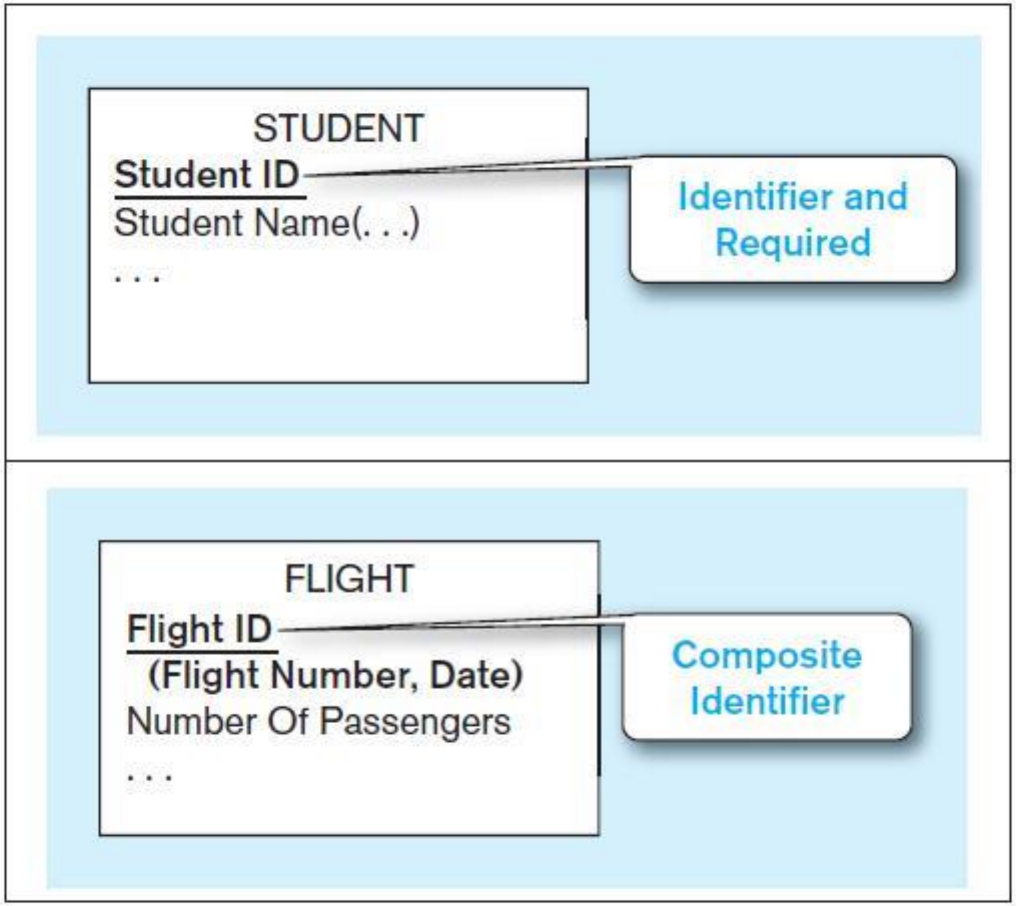


Derived
Calculated from date employed and current date

Figure 2-9 Simple and composite identifier attributes

(a) Simple identifier attribute

(b) Composite identifier attribute



The identifier is boldfaced and underlined

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 a key…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

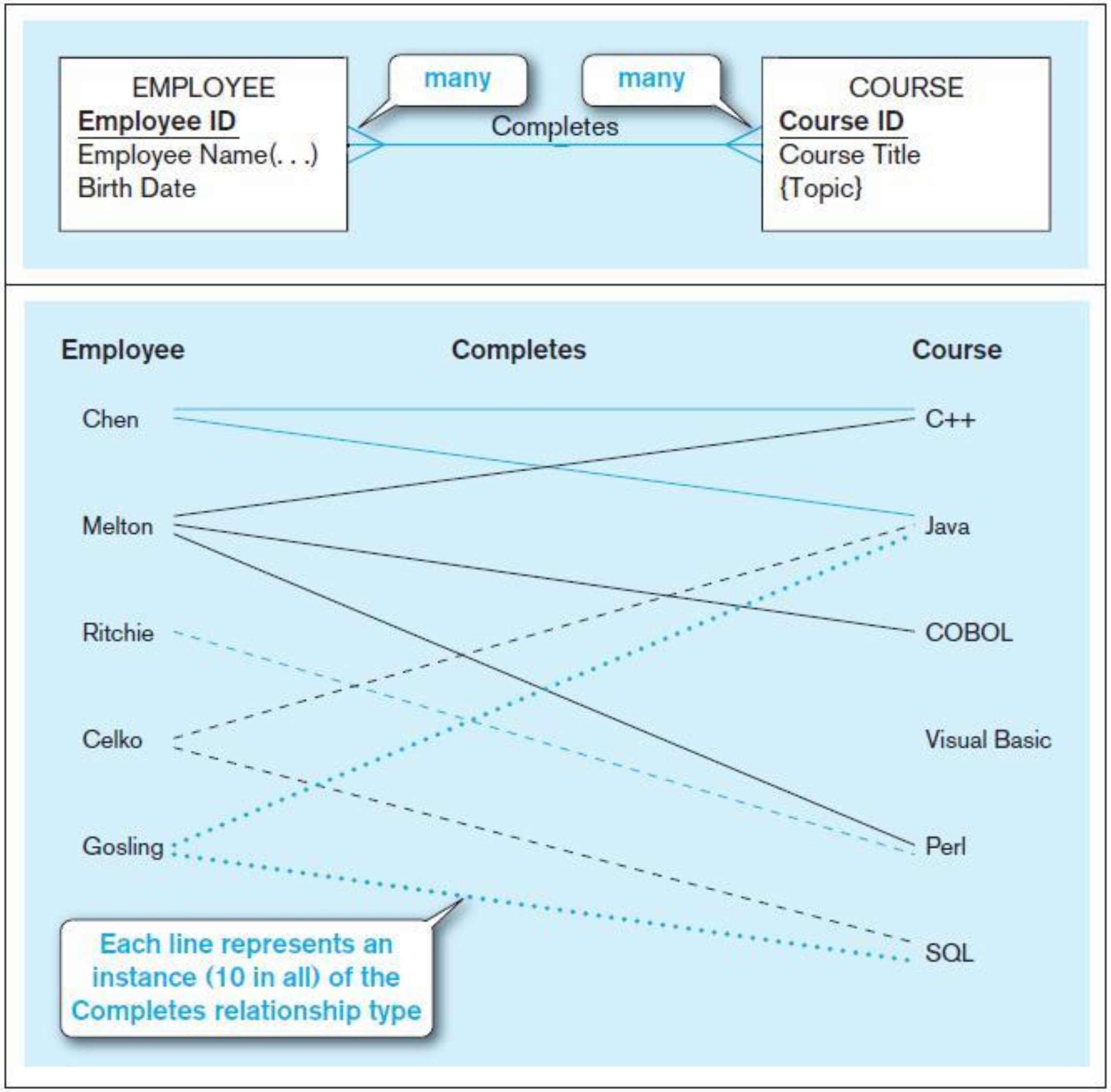
MODELING RELATIONSHIPS

- Relationship Types vs. Relationship Instances
  - The relationship type is modeled as lines between entity types…the instance is between specific entity instances
- Relationships can have attributes
  - These describe features pertaining to the association between the entities in the relationship
- Two entities can have more than one type of relationship between them (multiple relationships)
- Associative Entity–combination of relationship and entity

Figure 2-10 Relationship types and instances

a) Relationship type (Completes)

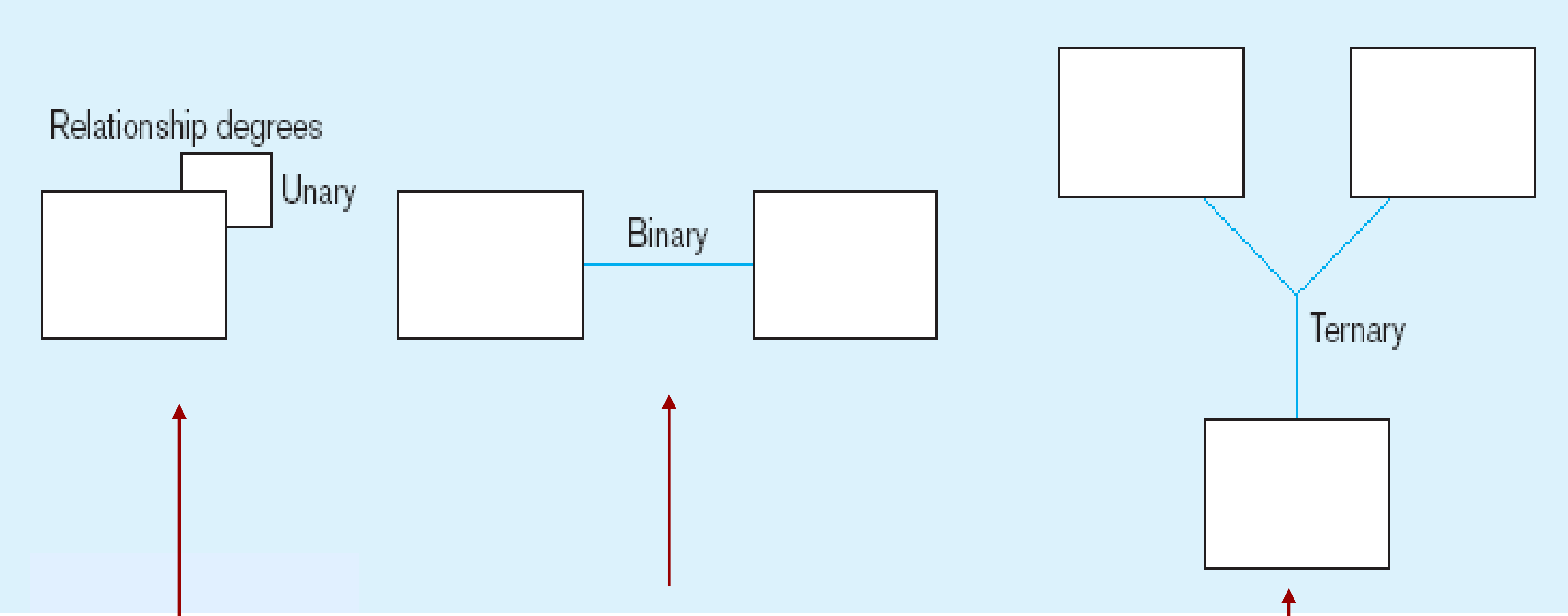
b) Relationship instances (tuples or rows)



DEGREES OF RELATIONSHIPS

- Degree of a relationship is the number of entity types that participate in it
  - –Unary Relationship
  - –Binary Relationship
  - –Ternary Relationship

Degree of relationships – from Figure 2-2



One entity related to another of the same entity type

Entities of two different types related to each other

Entities of three different types related to each other

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

a) Unary relationships

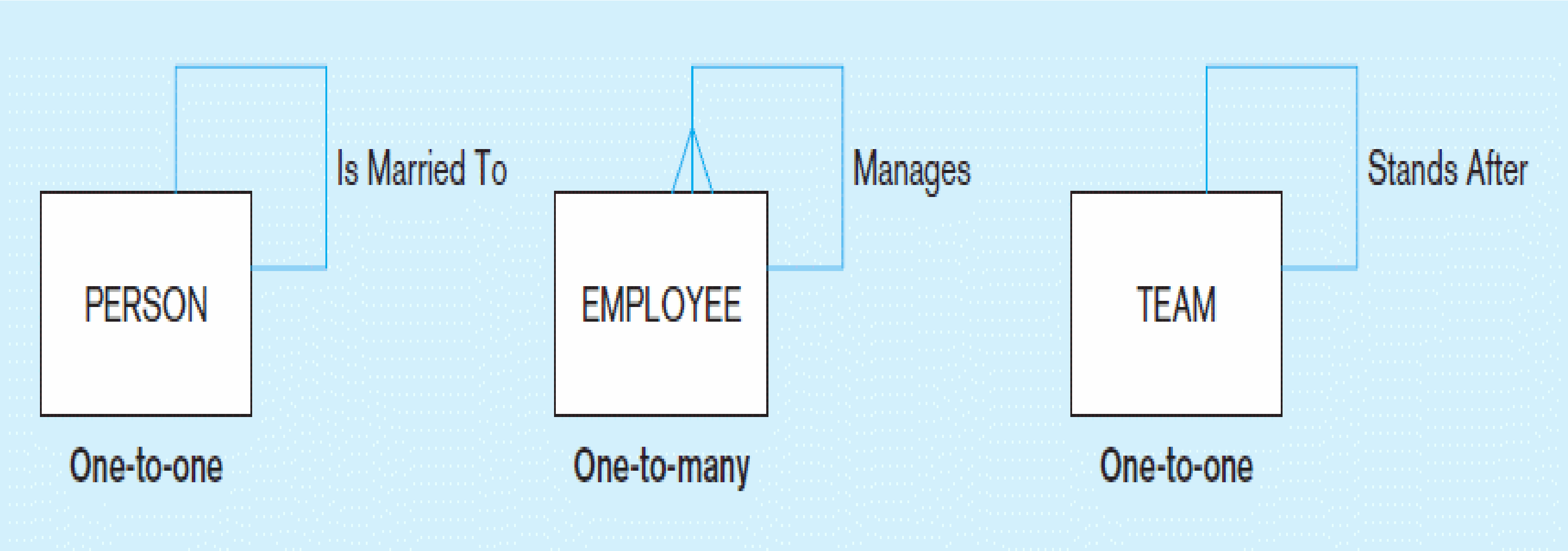


Figure 2-12 Examples of relationships of different degrees (cont.)

b) Binary relationships

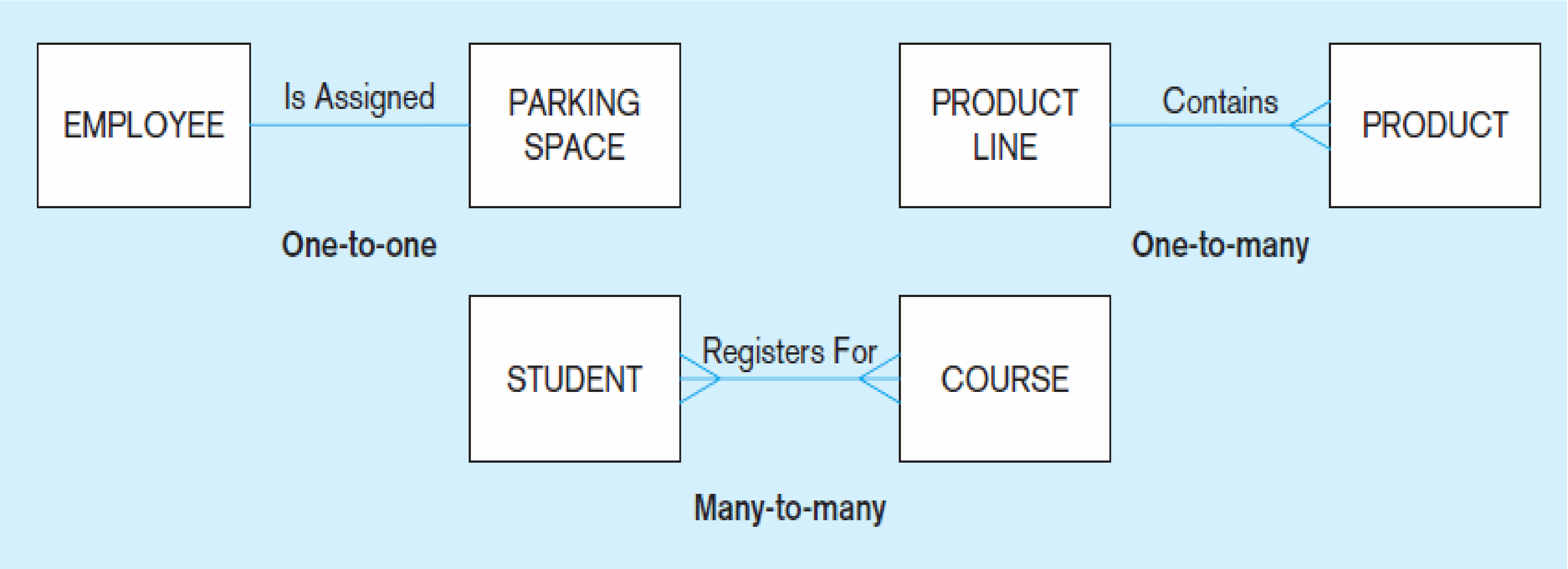
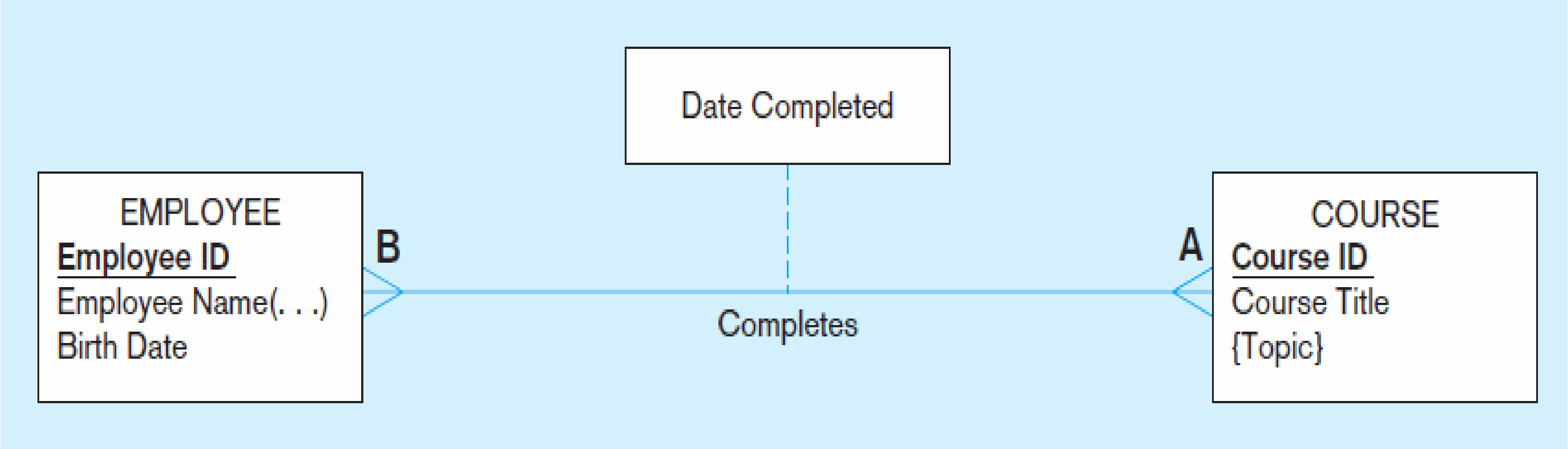


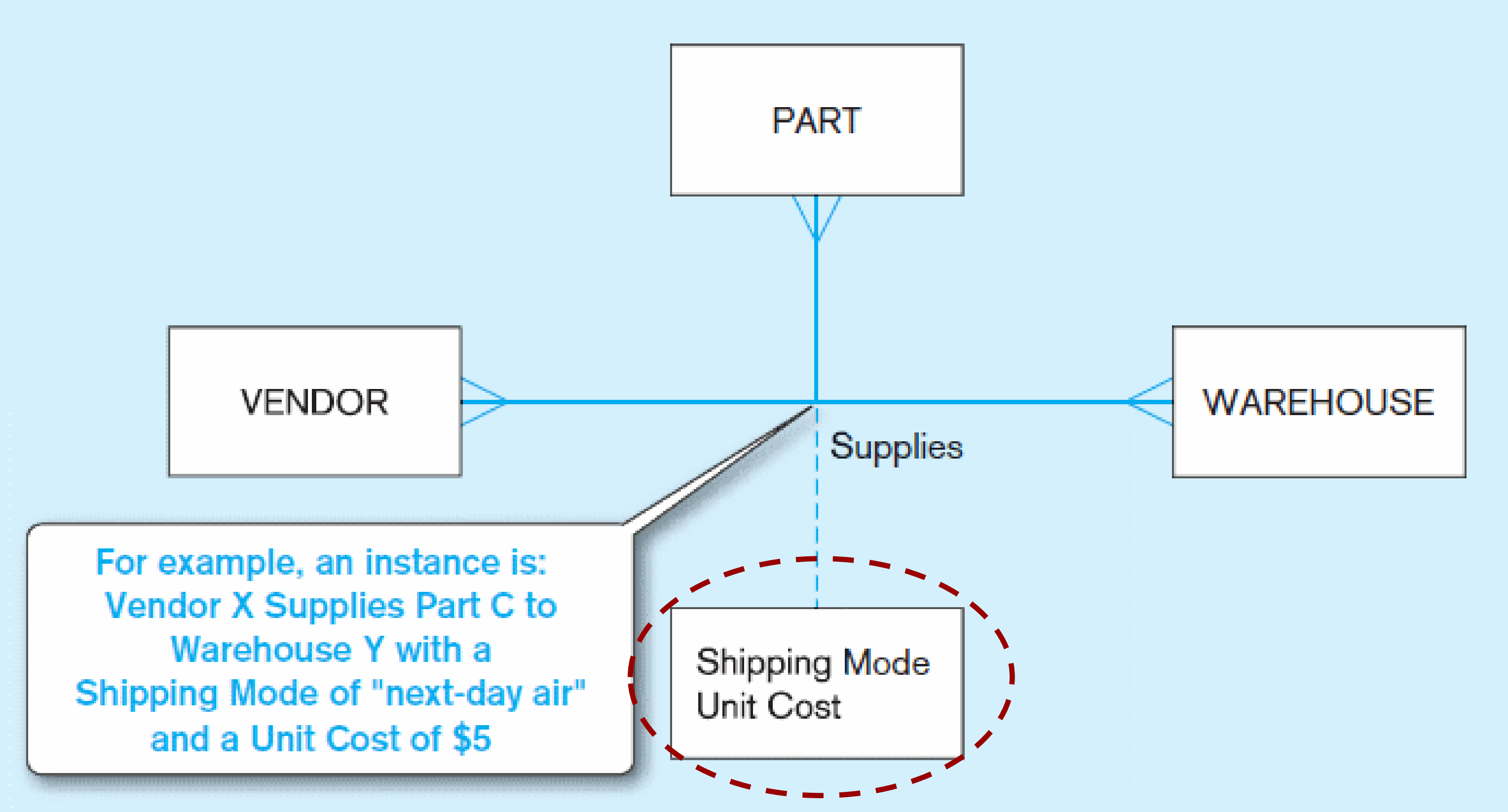
Figure 2-11a A binary relationship with an attribute



Here, the date completed attribute pertains specifically to the employee's completion of a course…it is an attribute of the *relationship*.

Figure 2-12 Examples of relationships of different degrees (cont.)

c) Ternary relationship



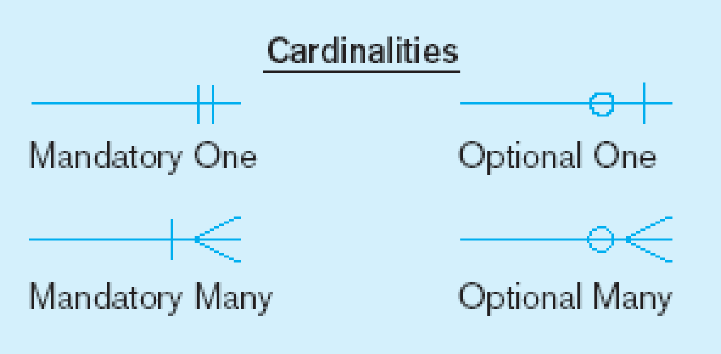
For example, an instance is: Vendor X Supplies Part C to Warehouse Y with a Shipping Mode of "next-day air" and a Unit Cost of $5

Note: a relationship can have attributes of its own

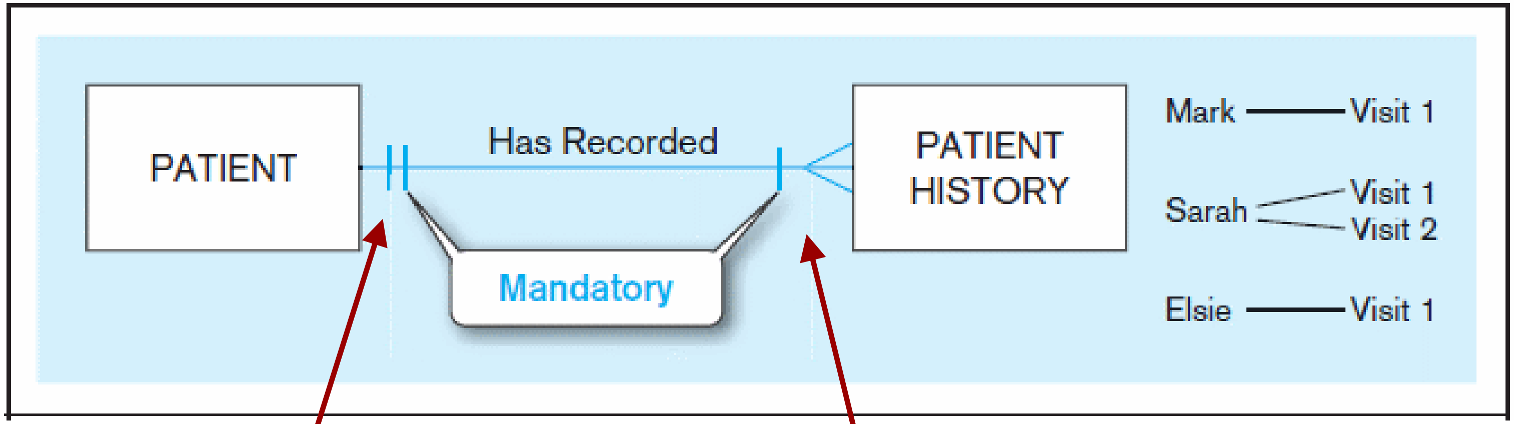
CARDINALITY CONSTRAINTS

- 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



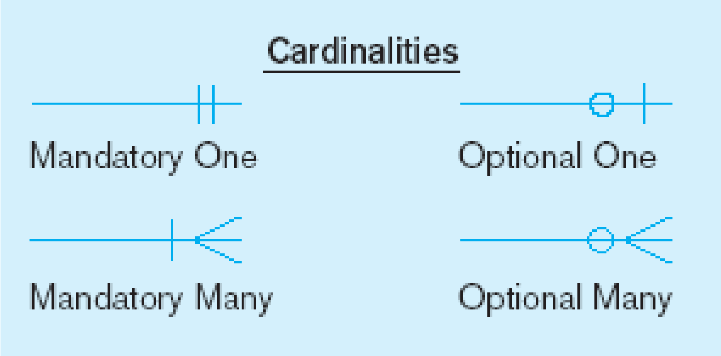
a) Mandatory cardinalities



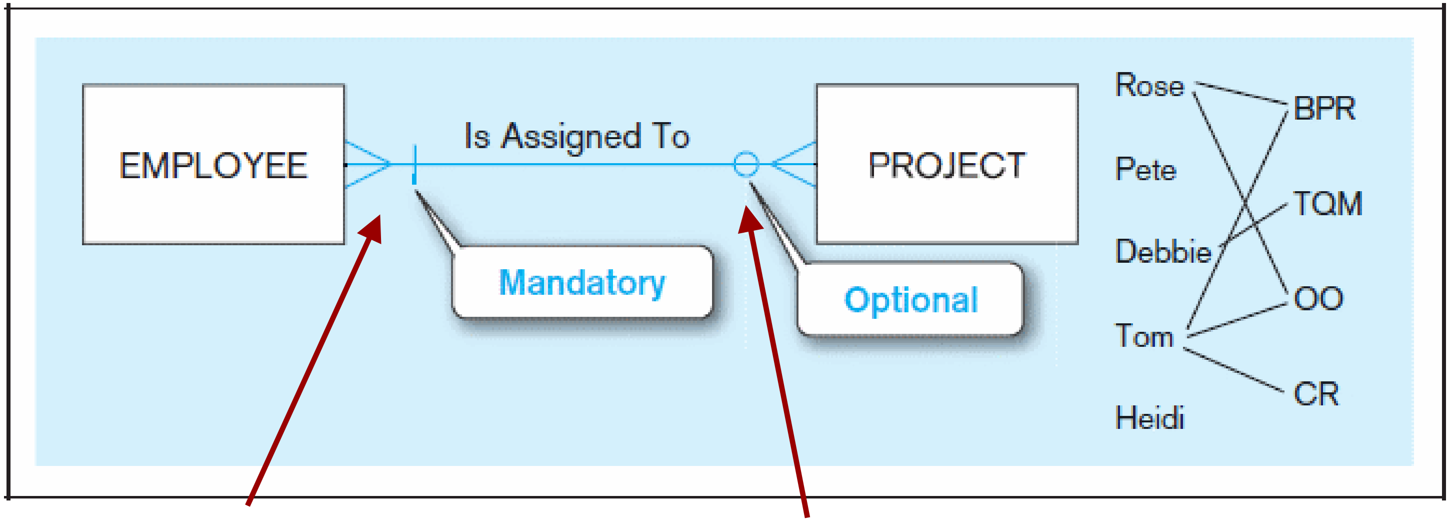
A patient history is recorded for one and only one patient

A patient must have recorded at least one history, and can have many

Figure 2-17 Examples of cardinality constraints (cont.)



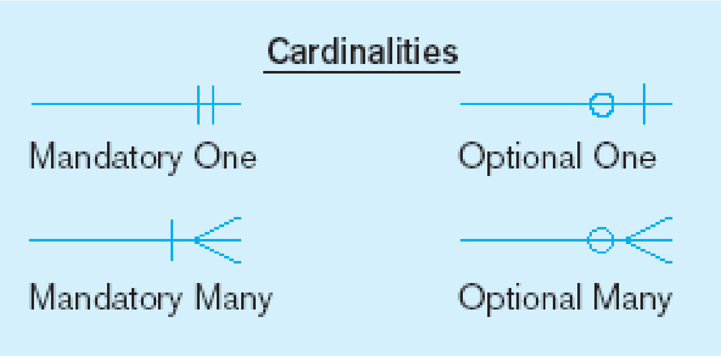
b) One optional, one mandatory



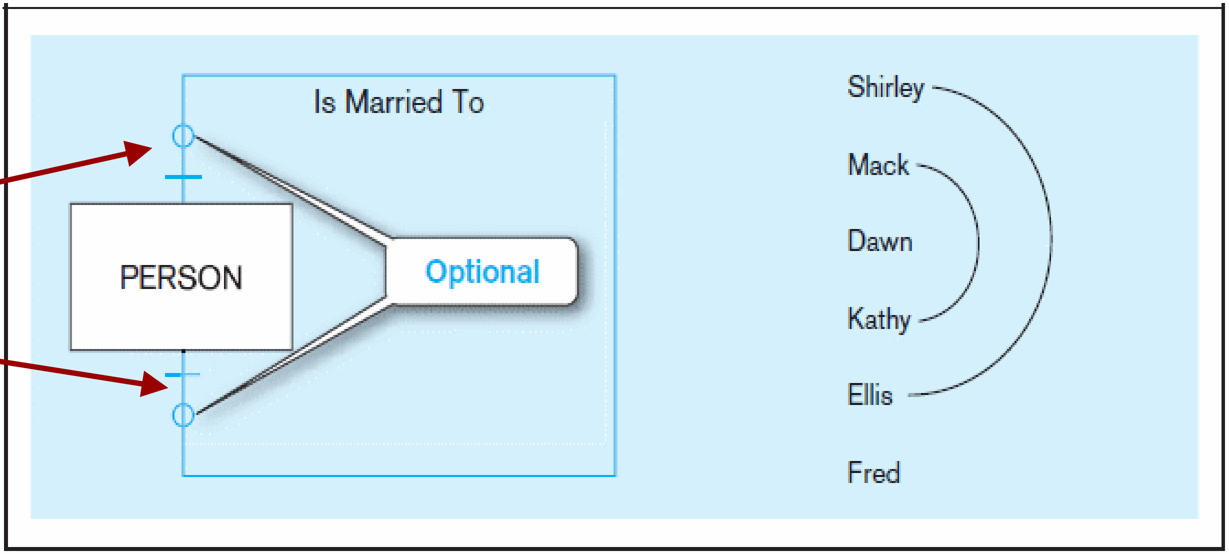
A project must be assigned to at least one employee, and may be assigned to many

An employee can be assigned to any number of projects, or may not be assigned to any at all

Figure 2-17 Examples of cardinality constraints (cont.)



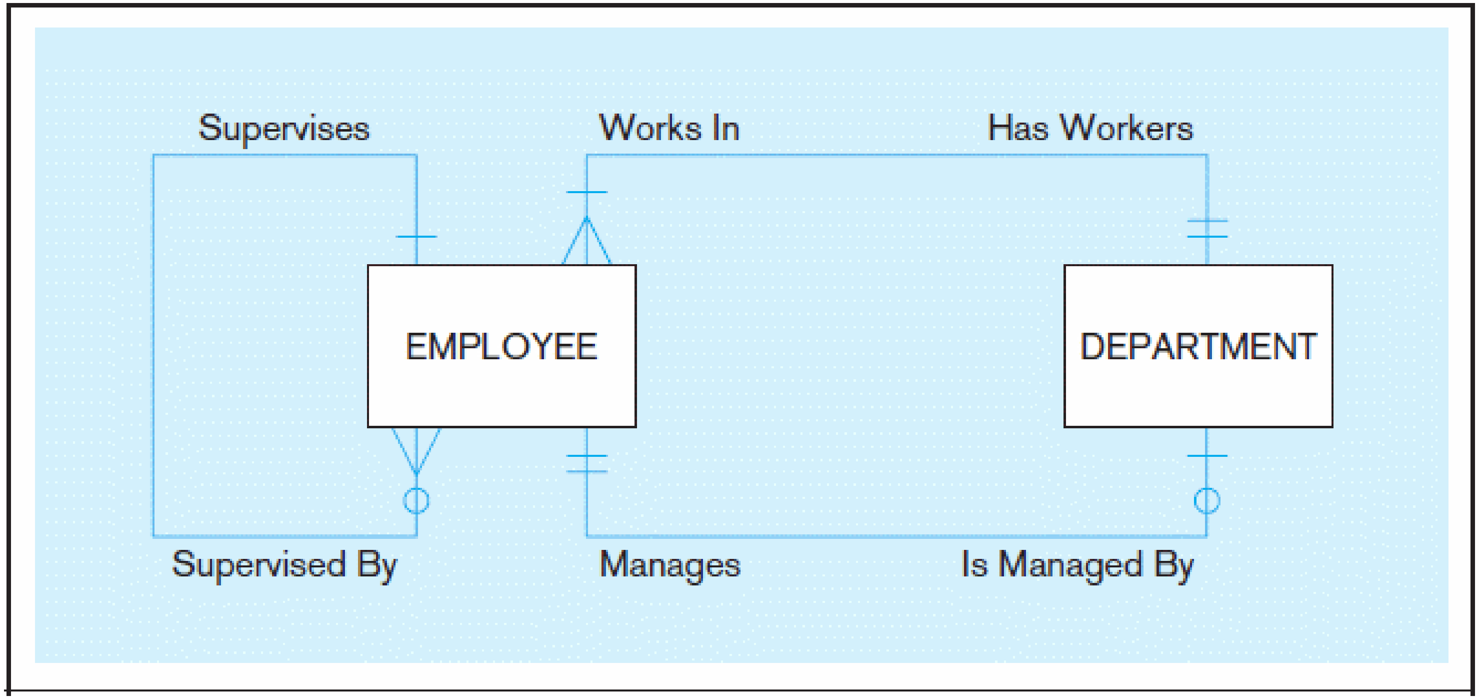
c) Optional cardinalities



A person is married to at most one other person, or may not be married at all

Figure 2-21 Examples of multiple relationships

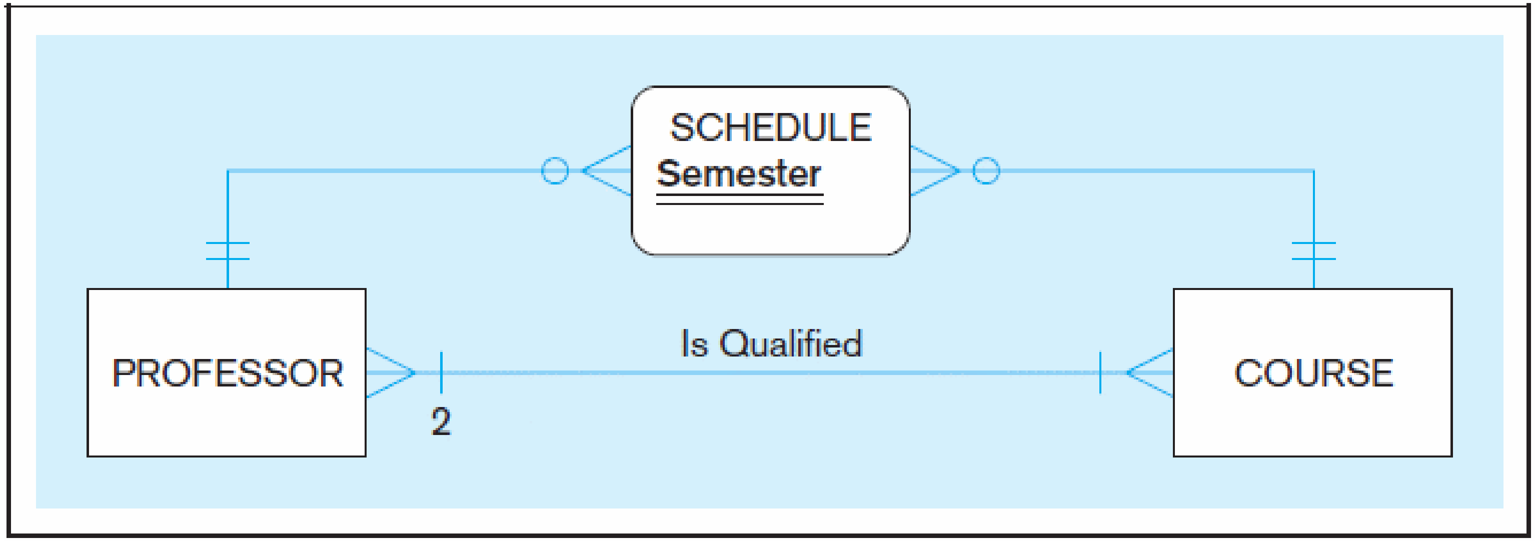
a) Employees and departments



Entities can be related to one another in more than one way

Figure 2-21 Examples of multiple relationships (cont.)

b) Professors and courses (fixed lower limit constraint)

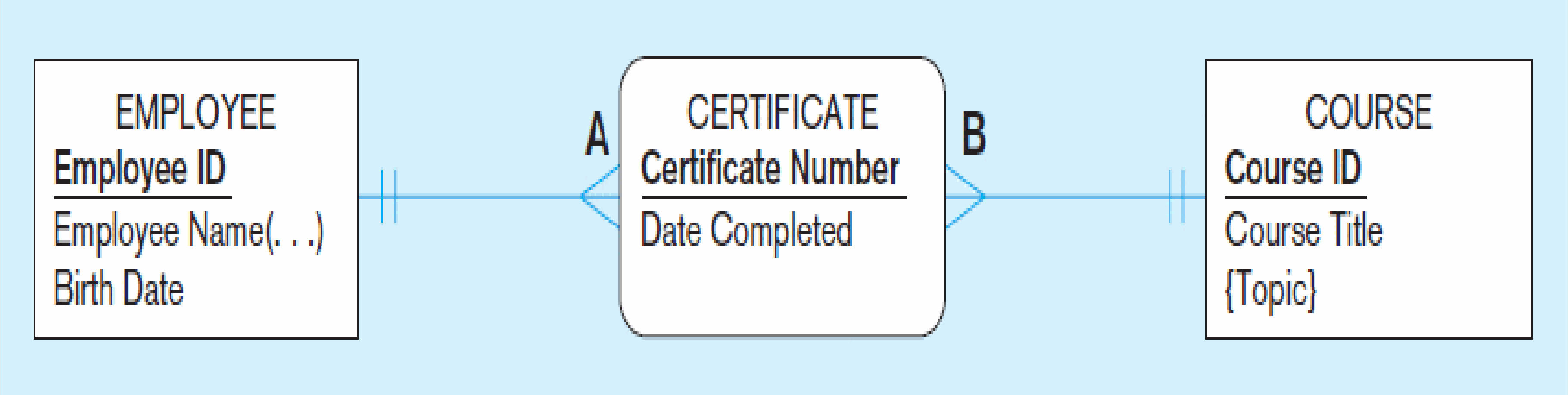


Here, min cardinality constraint is 2. At least two professors must be qualified to teach each course. Each professor must be qualified to teach at least one course.

ASSOCIATIVE ENTITIES

- An entity–has attributes
- A relationship–links entities together
- 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
  - Ternary relationships should be converted to associative entities

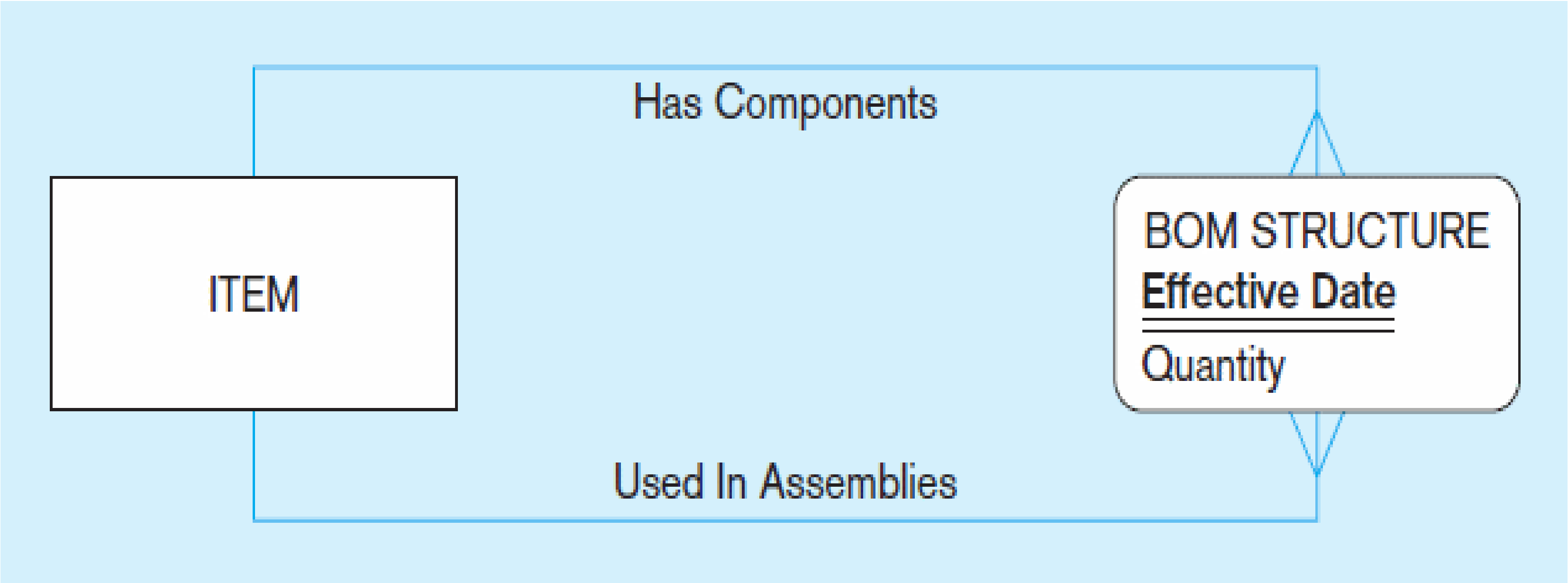
Figure 2-11b An associative entity (CERTIFICATE)



Associative entity is like a relationship with an attribute, but it is also considered to be an entity in its own right.

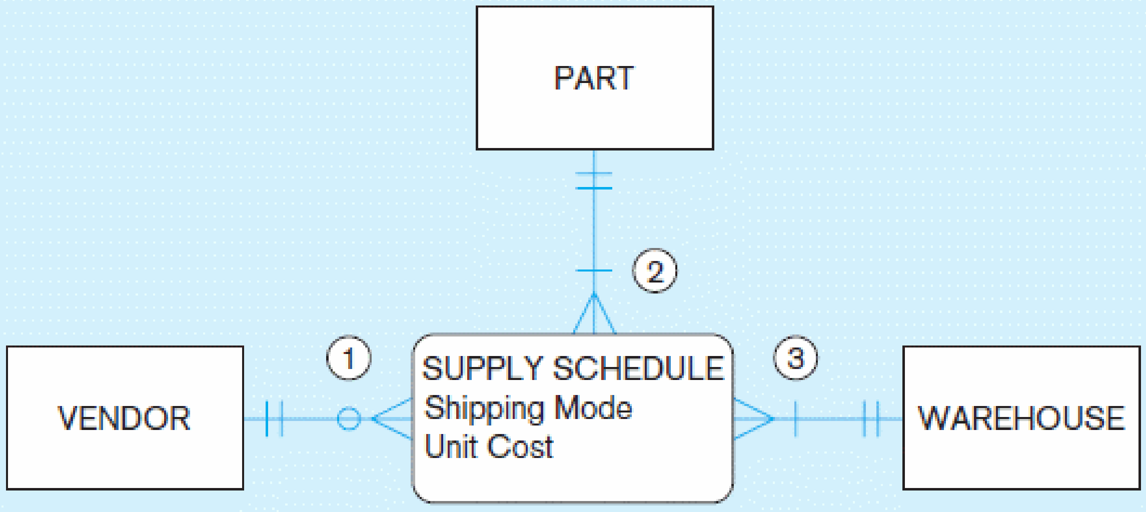
Note that the many-to-many cardinality between entities in Figure 2-11a has been replaced by two one-to-many relationships with the associative entity.

Figure 2-13c An associative entity – bill of materials structure



This could just be a relationship with attributes…it' s a judgment call.

Figure 2-18 Cardinality constraints in a ternary relationship



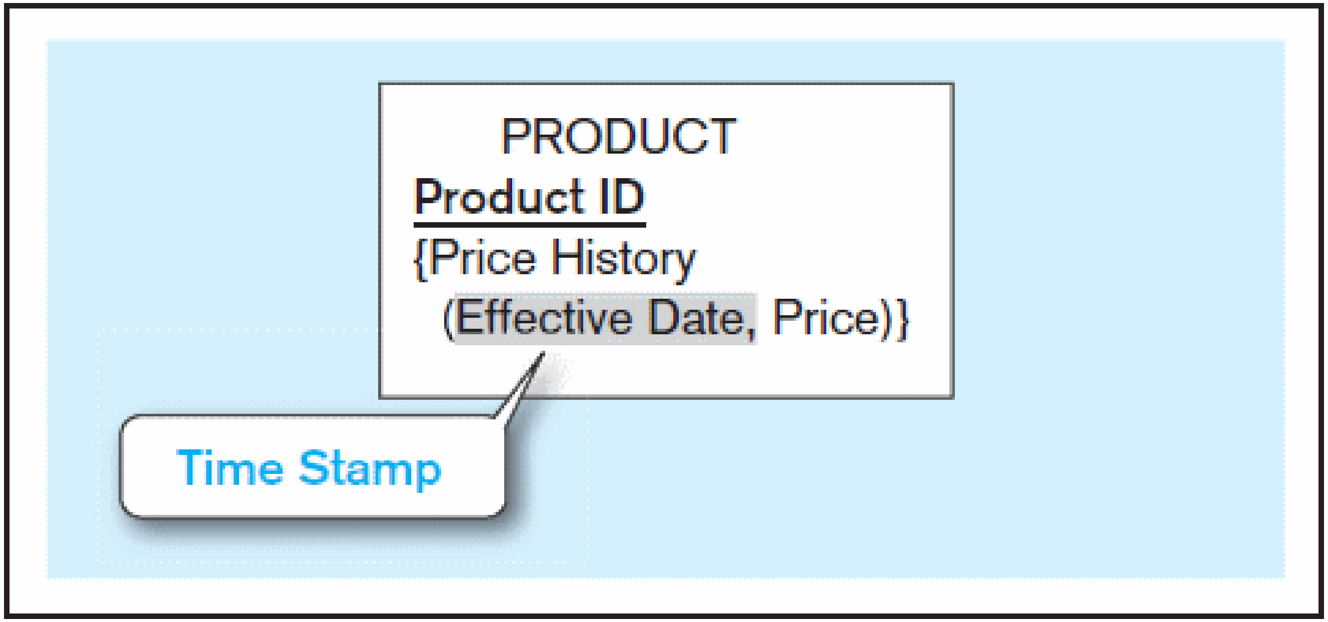
Business Rules

1. Each vendor can supply many parts to any number of warehouses but need not supply any parts.

2. Each part can be supplied by any number of vendors to more than one warehouse, but each part must be supplied by at least one vendor to a warehouse.

3. Each warehouse can be supplied with any number of parts from more than one vendor, but each warehouse must be supplied with at least one part.

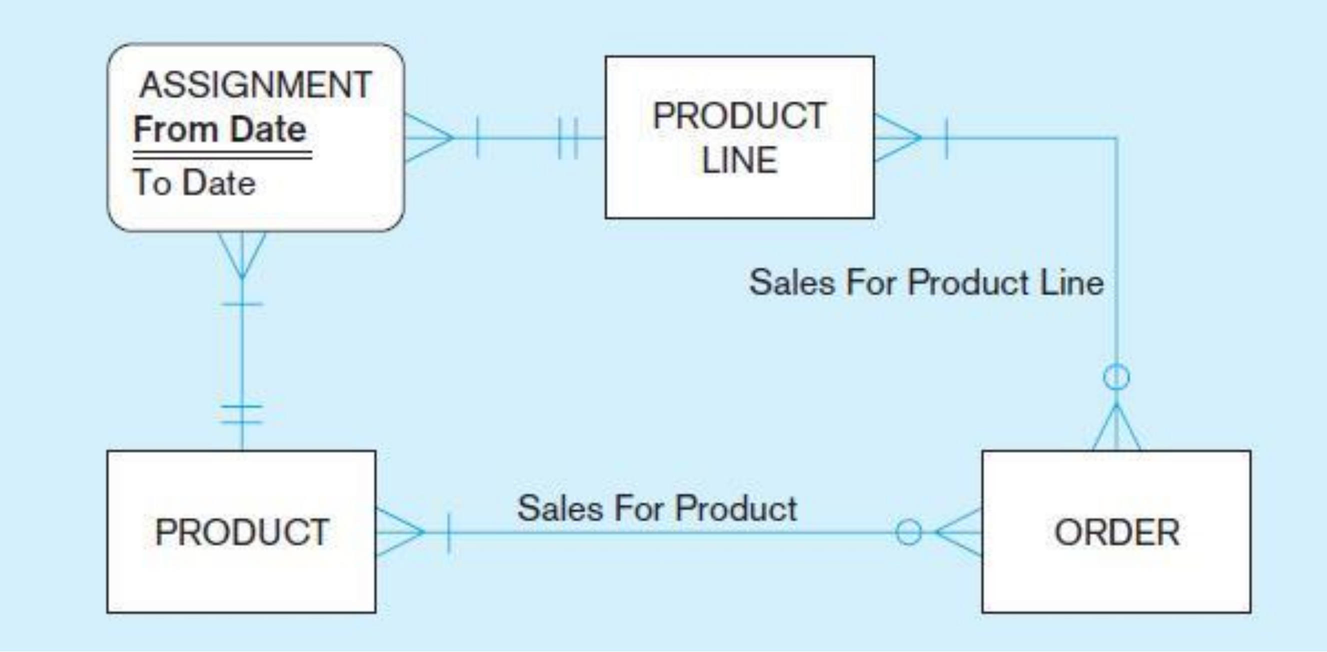
Figure 2-19 Simple example of time-stamping



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.

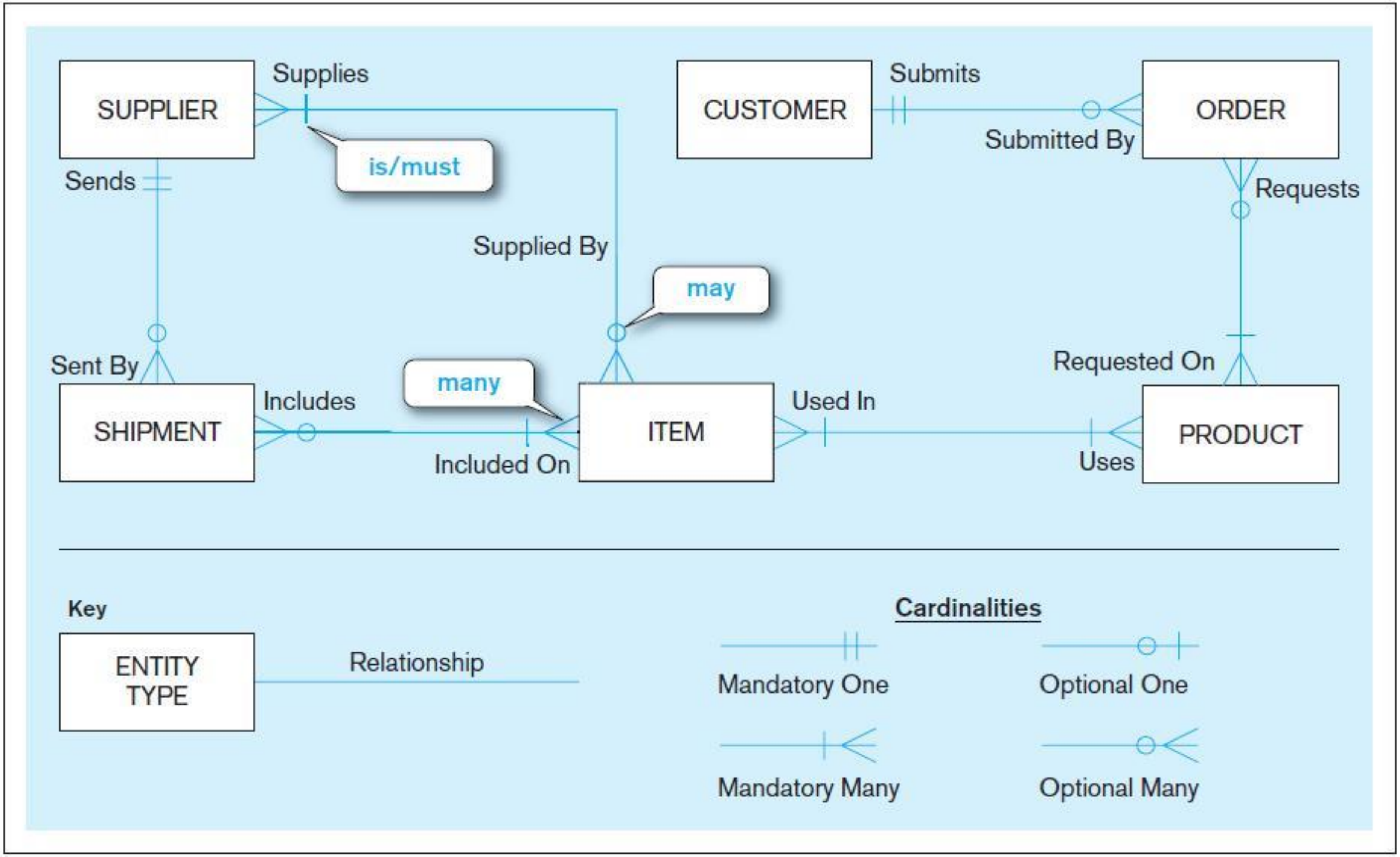
Figure 2-20c E-R diagram with associative entity for product assignment to product line over time

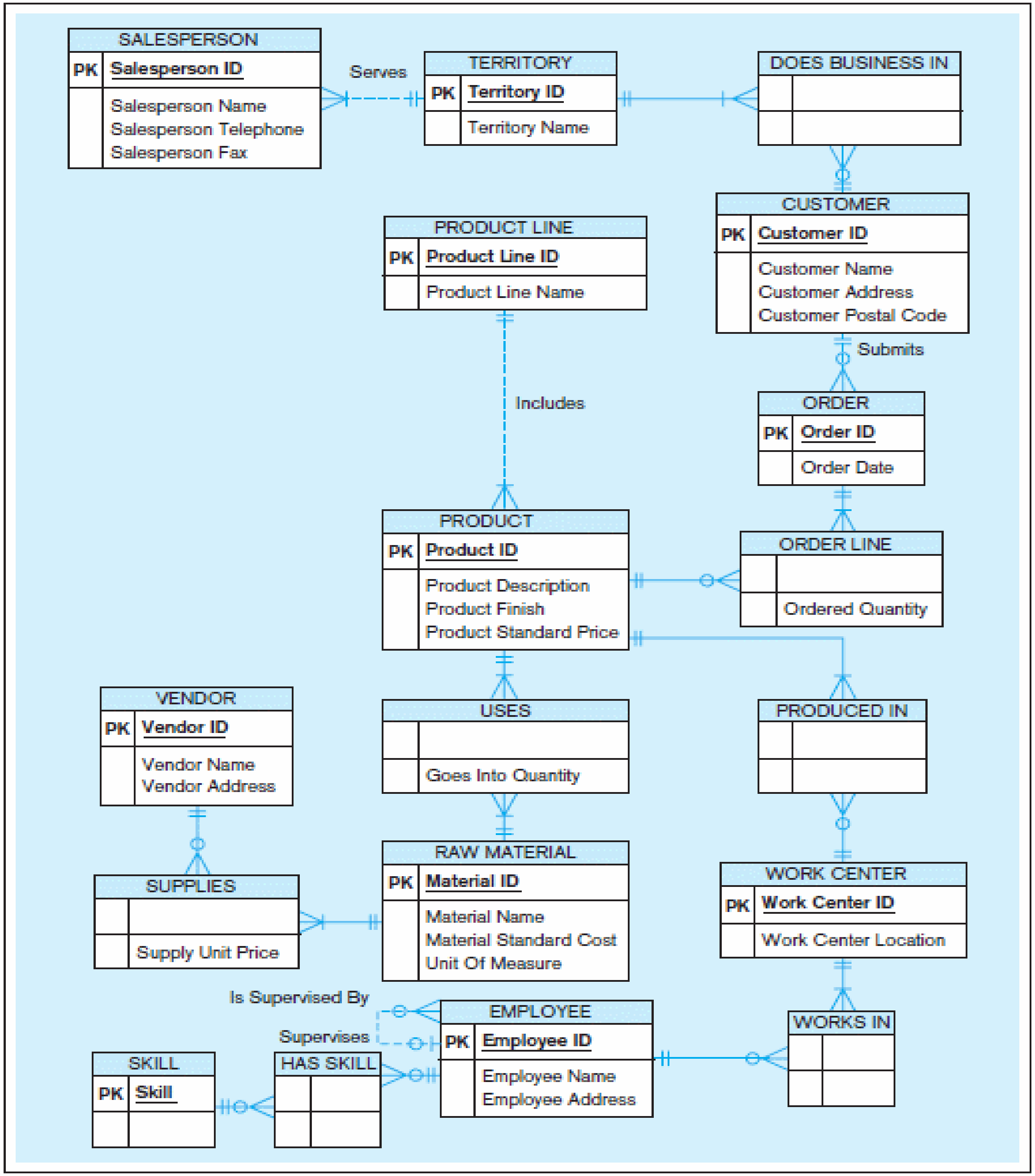


Modeling time-dependent data has become more important due to regulations such as HIPAA and Sarbanes-Oxley.

The Assignment associative entity shows the date range of a product’ s assignment to a particular product line.

Sample E-R Diagram (Figure 2-1)





SALESPERSON
PK Salesperson ID
Salesperson Name
Salesperson Telephone
Salesperson Fax

Serves

TERRITORY
PK Territory ID
Territory Name

DOES BUSINESS IN

CUSTOMER
PK Customer ID
Customer Name
Customer Address
Customer Postal Code

Submits

PRODUCT LINE
PK Product Line ID
Product Line Name

Includes

ORDER
PK Order ID
Order Date

PRODUCT
PK Product ID
Product Description
Product Finish
Product Standard Price

ORDER LINE
Ordered Quantity

VENDOR
PK Vendor ID
Vendor Name
Vendor Address

USES
Goes Into Quantity

PRODUCED IN

SUPPLIES
Supply Unit Price

RAW MATERIAL
PK Material ID
Material Name
Material Standard Cost
Unit Of Measure

WORK CENTER
PK Work Center ID
Work Center Location

Is Supervised By

Supervises

EMPLOYEE
PK Employee ID
Employee Name
Employee Address

WORKS IN

SKILL
PK Skill

HAS SKILL

Figure 2-22 Data model for Pine Valley Furniture Company in Microsoft Visio notation

Different modeling software tools may have different notation for the same constructs.