Chapter3 Modeling data in the organization

Kun Yue

March, 2009

Outline

Data modeling is the most important part of the system development process.

- **Introduction**
- **The E-R model**
  - E-R diagram
  - Entities
  - Attributes
  - Relationships
  - Modeling time-dependent data

Introduction

- **Business rules —— requirements**

  - Define or constrain some aspects of the business
  - Assert business structure, control or influence the behavior of the business
  - Policies, procedures, events, functions, other business objects and state constraints

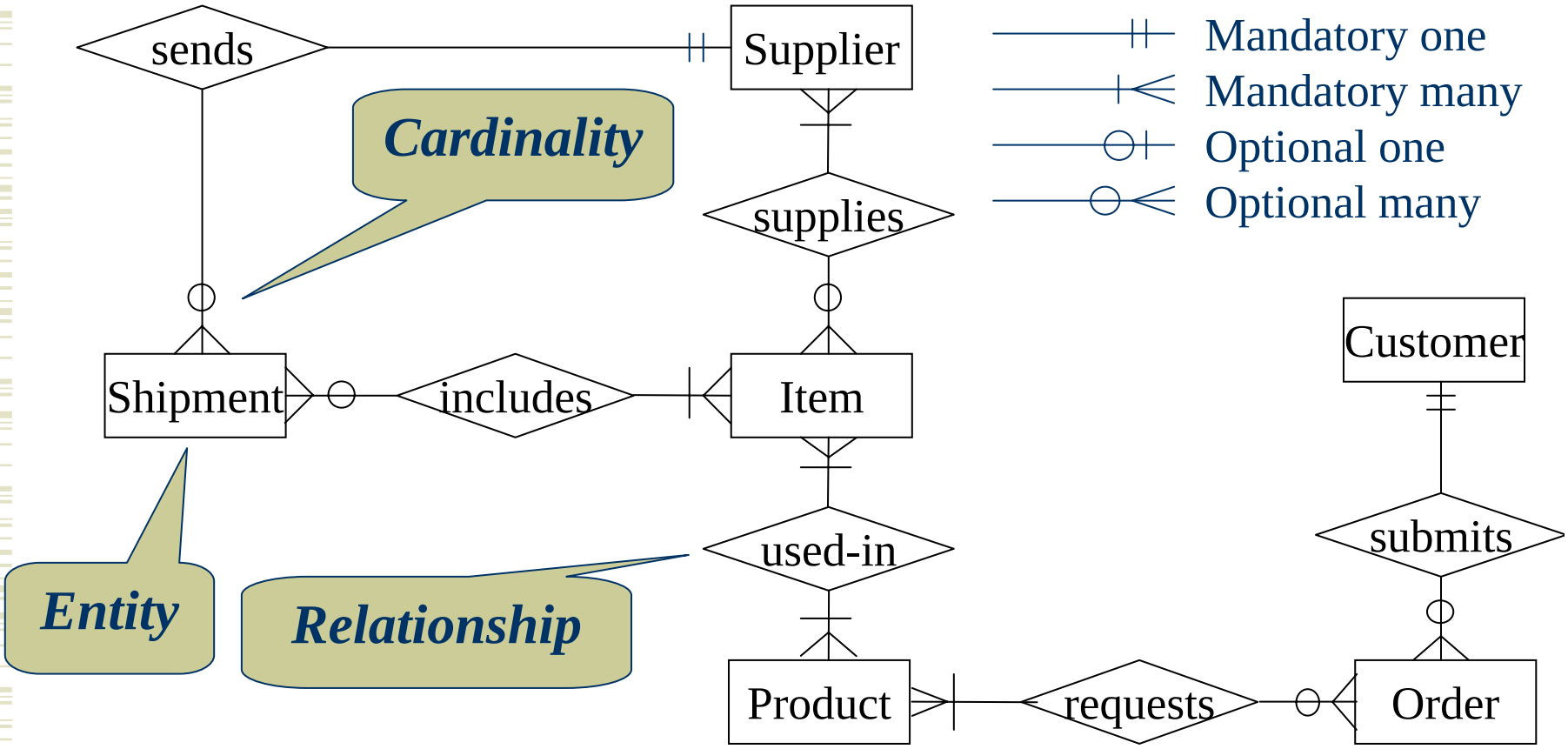
- How to find the business rules? From the iterative inquiry process
- How to map the business rules into the database?

The E-R model

- **Introduction**
- **The E-R model**
  - E-R diagram
  - Entities
  - Attributes
  - Relationships
  - Modeling time-dependent data

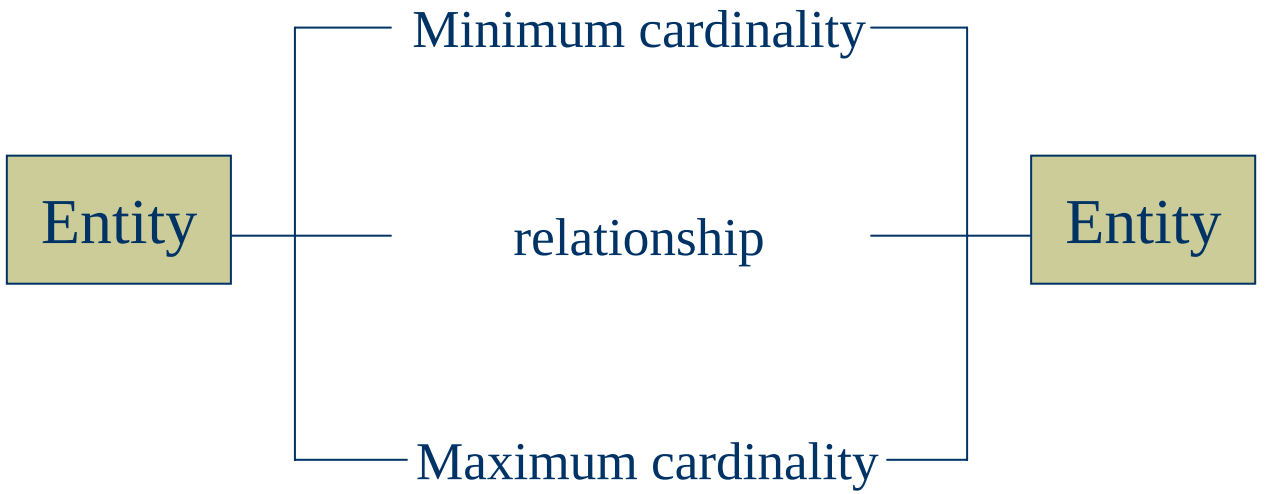
The E-R model

- **Entity-relationship diagram: entities, relationships, attributes**



Revisit the business rule

- **Business rule grammar:**



E.g,

(1) Each customer can submit any number of orders. (zero or more)

(2) Each product must use at least one item. (one or more)

Entities and their attributes

- Entity type vs. entity instance
- Strong entity types vs. weak entity types

- Simple attributes vs. composite attributes
- Single-valued attributes vs. multivalued attributes
- Stored attributes vs. derived attributes
- Identifier attribute

Entity type vs. entity instance

- **Entity type**

- A collection of entities that share common properties or characteristics

- E-R diagram includes the ***entity types*** that ***can not be computed by others***

- **Entity instance**

- A single occurrence of an entity type

- An entity type is described once in the database

E.g., ***EMPLOYEE*** **(*No, name, address, city, zip, birthday*)**

- Many entity instances of the entity type are stored in the database.

E.g.,

642-17-8630	Michelle	2 North Cuihu Road	Kunming	650091	03-12-1993
534-10-1971	David	220 Handan Road	Shanghai	200433	08-16-1994

Strong entity types vs. weak entity types (1)

- A motivating example:

  Book-Chapter

Book_ID	Book_Name	Chapter_ID	Chapter_Name
B01	Data Structure	C01	Introduction
B01	Data Structure	C02	Graph
B02	Algorithm	C01	Introduction
B02	Algorithm	C02	Graph

  - Does the book exist uniquely and independently?
  - Does the chapter exist uniquely and independently?

Strong entity types vs. weak entity types (2)

- **Strong entity type:**
  - An entity type that exists ***independently*** of other entity types
  - Has a ***unique*** identifier globally
- **Weak entity type:**
  - An entity type whose existence depends on some other entity type
  - Makes sense only in the context of the entity type that it depends on
  - Identifying owner — the entity type on which the weak entity type depends
  - A weak entity type does not have its own globally unique identifier, may have a partial identifier
  - Identifying relationship — the relationship between a weak and its owner

Book

Chapter

Owner of Chapter: Book

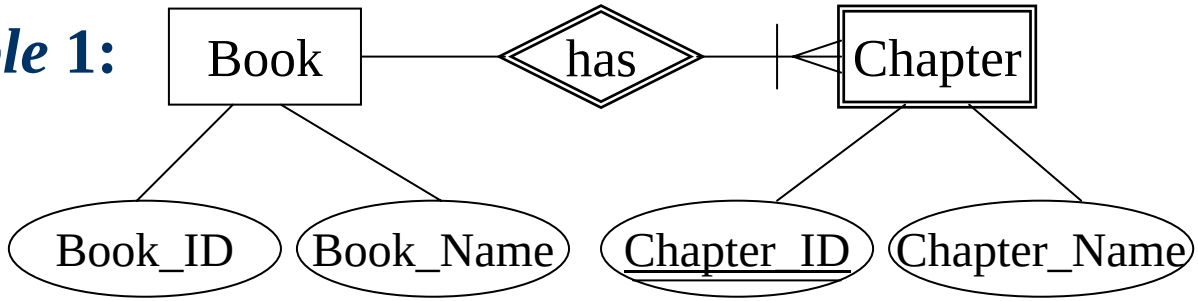
Partial identifier of Chapter: Chapter_ID

How to identify a weak entity?

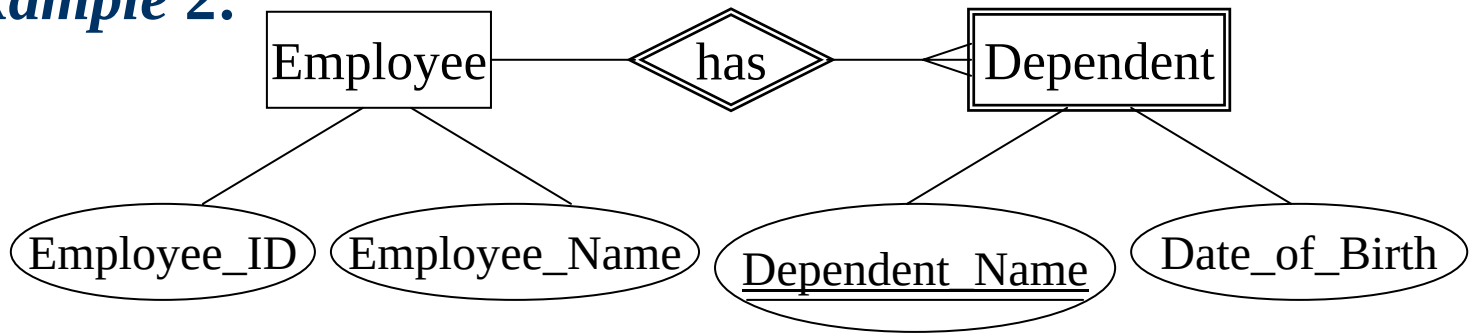
Strong entity types vs. weak entity types (3)

- How to represent the weak entity type and the identifying relationship in E-R diagram?

***Example* 1:**



***Example* 2:**



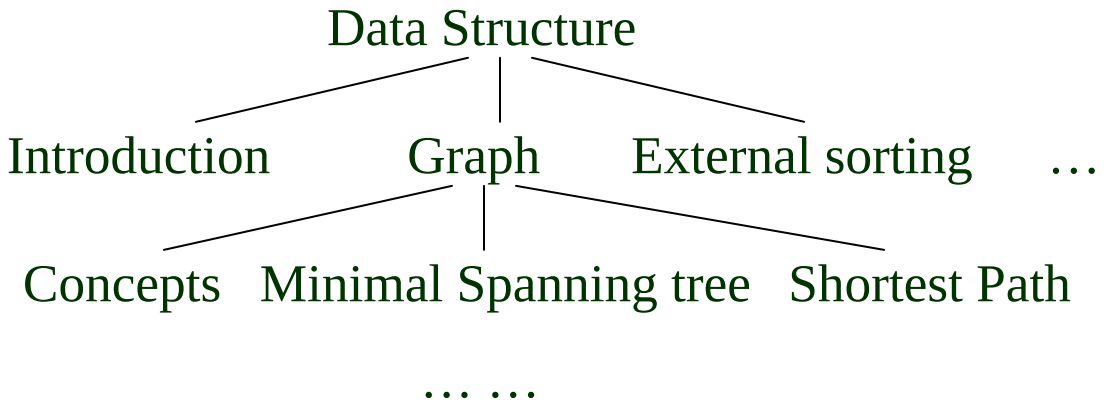
How to identify "Chapter" and "Dependent" globally?

Discussion

- Is the owner of the weak entity sure to be a strong entity?
- Is there any weak entities in the tree-structured data?

***Example*: Book — Chapter — Section**

- Owner is said relatively to the weak entity.
- What are weak entities? What are the owners of them respectively?
- **The owner may be strong or weak entity.**



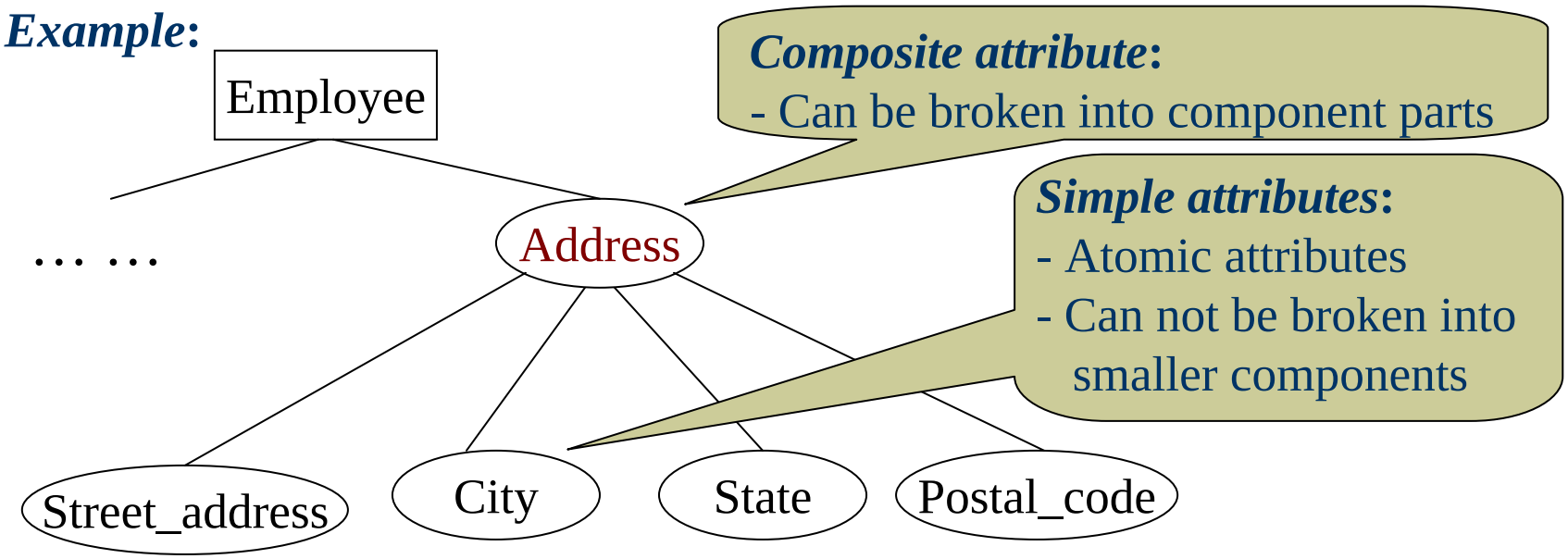
Attributes (1)

- **An attribute**:

- A property or characteristic of an entity type that of interest to the organization
- An entity instance is composed of the attribute values of the entity type

- **Simple attributes** vs. **composite attributes**

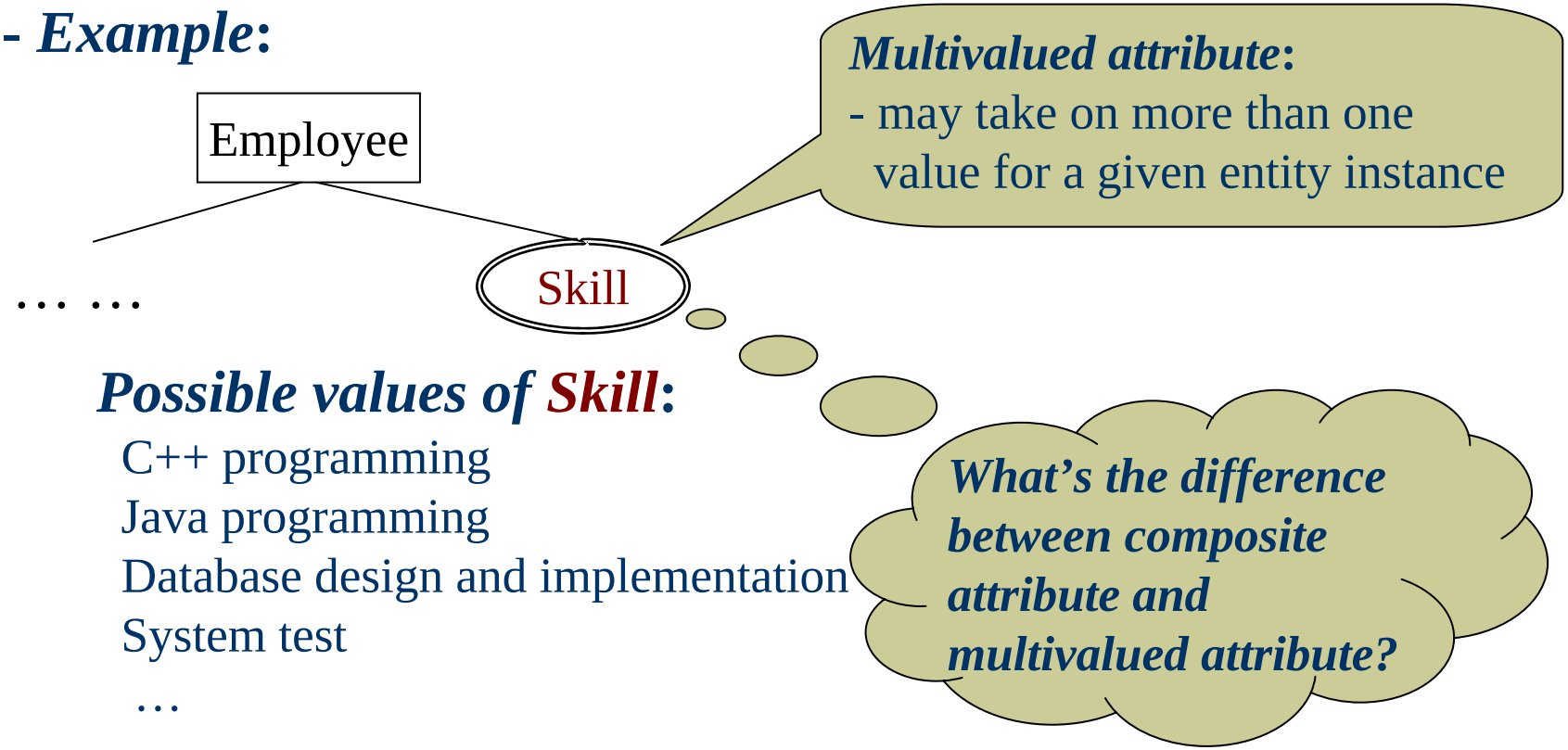
- ***Example*:**



Attributes (2)

- **Single-valued attributes** vs. **multivalued attributes**

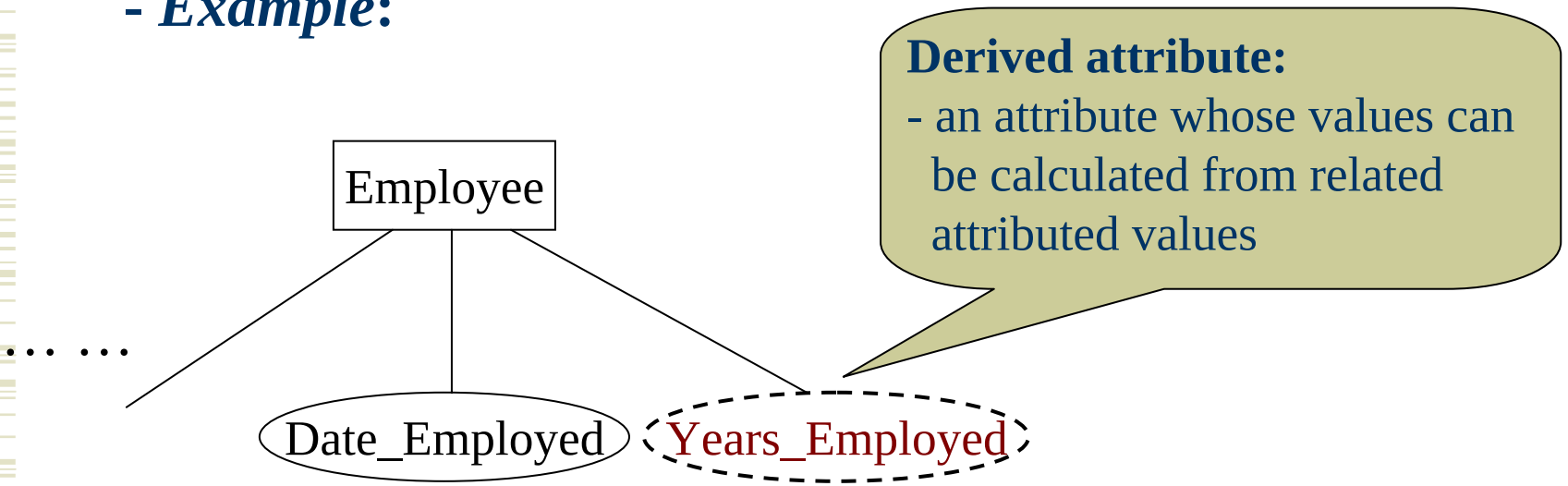
***- Example*:**



Attributes (3)

- **Stored attributes** vs. **derived attributes**

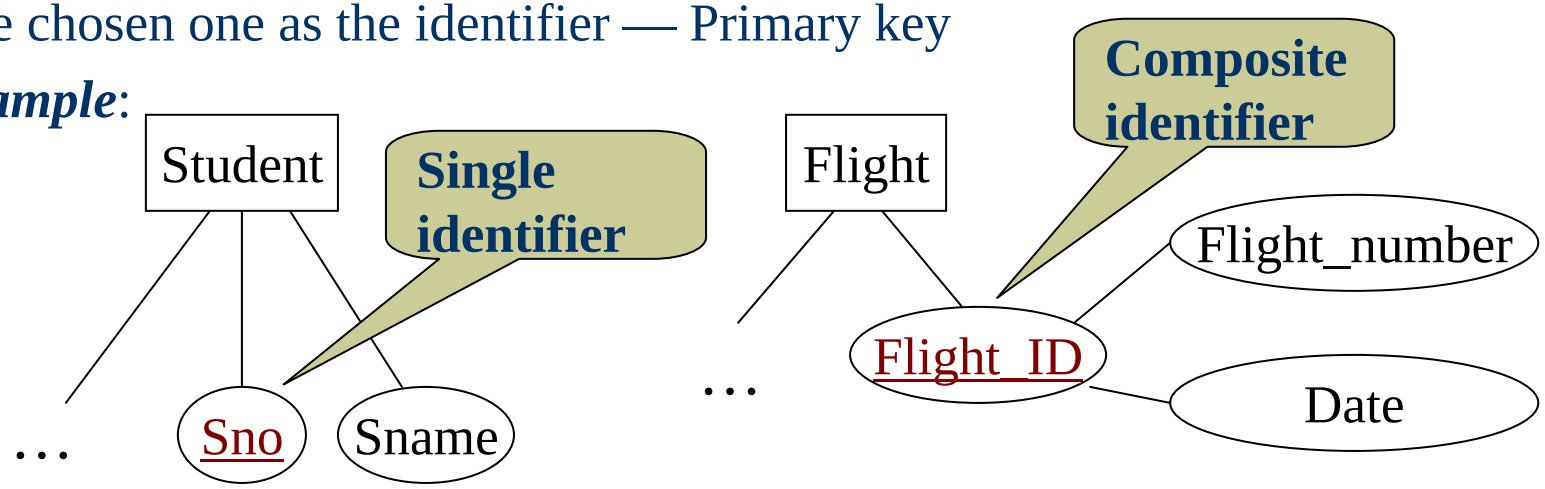
  - ***Example*:**



Attributes (4)

- **Identifier attribute**

- An attribute or combination of attributes that ***uniquely*** identifies individual instances of an entity type
- **Single identifier** vs. **composite identifier —**
  Single attribute vs. Composite attribute
- More than one candidate identifiers — Keys
  The chosen one as the identifier — Primary key
- ***Example***:



Relationships

- Basic concepts and definitions
- Degree of a relationship
- Attributes or relationships?
- Cardinality constraints

Basic concepts and definitions of relationships (1)

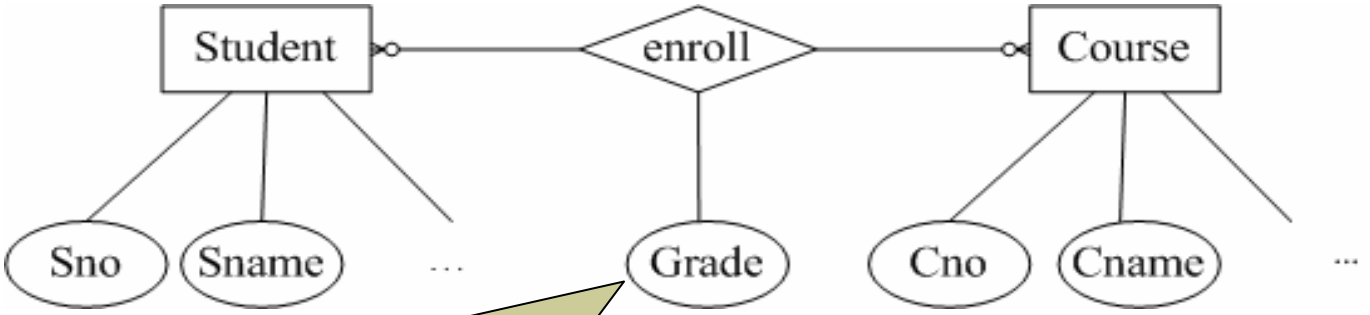
- **Relationship type**

A meaningful association between (among) entity types

- **Relationship instance**

An association between (among) entity instances where each relationship instance includes ***exactly one*** entity instance from each participating entity type.

- **Attributes on relationships**

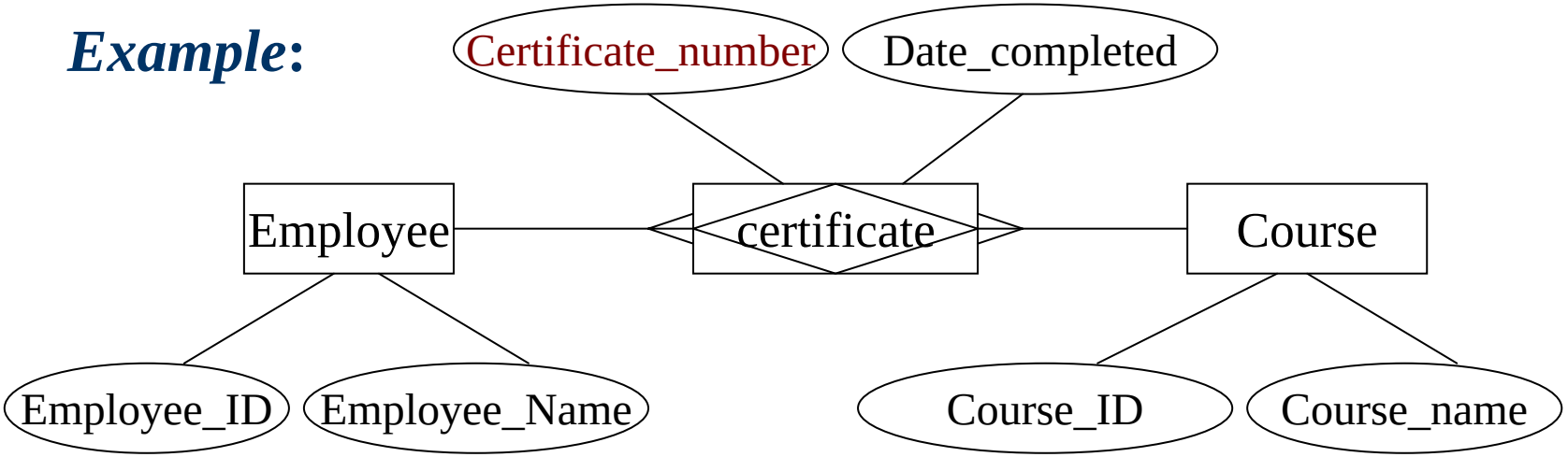


enroll: Many-to-Many relationship
"**Grade**" :the attribute of "enroll"

Basic concepts and definitions of relationships (2)

- **Associative entities** (different from the attributes of the relationship)

***Example*:**

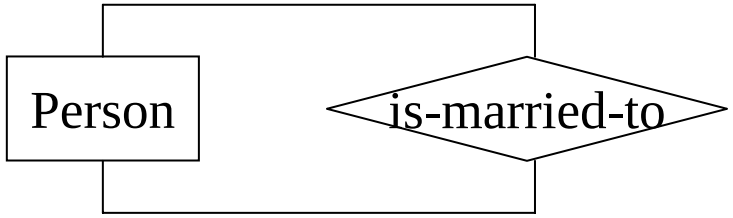


- An entity type that associates the instances of one or more entity types and **contains attributes that are peculiar to the relationship** between those entity instances.
- A single-attribute identifier
- Independent on the entities in the associated relationship

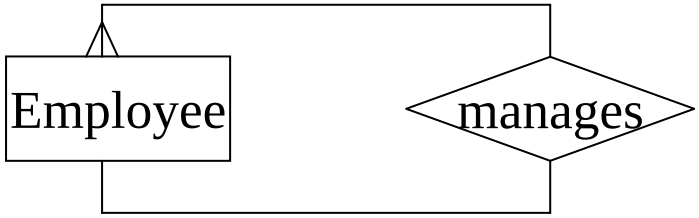
Degree of a relationship (1)

- **Degree**
  - The number of entity types that participate in a relationship
  - The 3 common relationship degrees in E-R models:
    unary (degree 1), binary (degree 2), ternary (degree 3)
- **Unary relationship**

(1) one-to-one



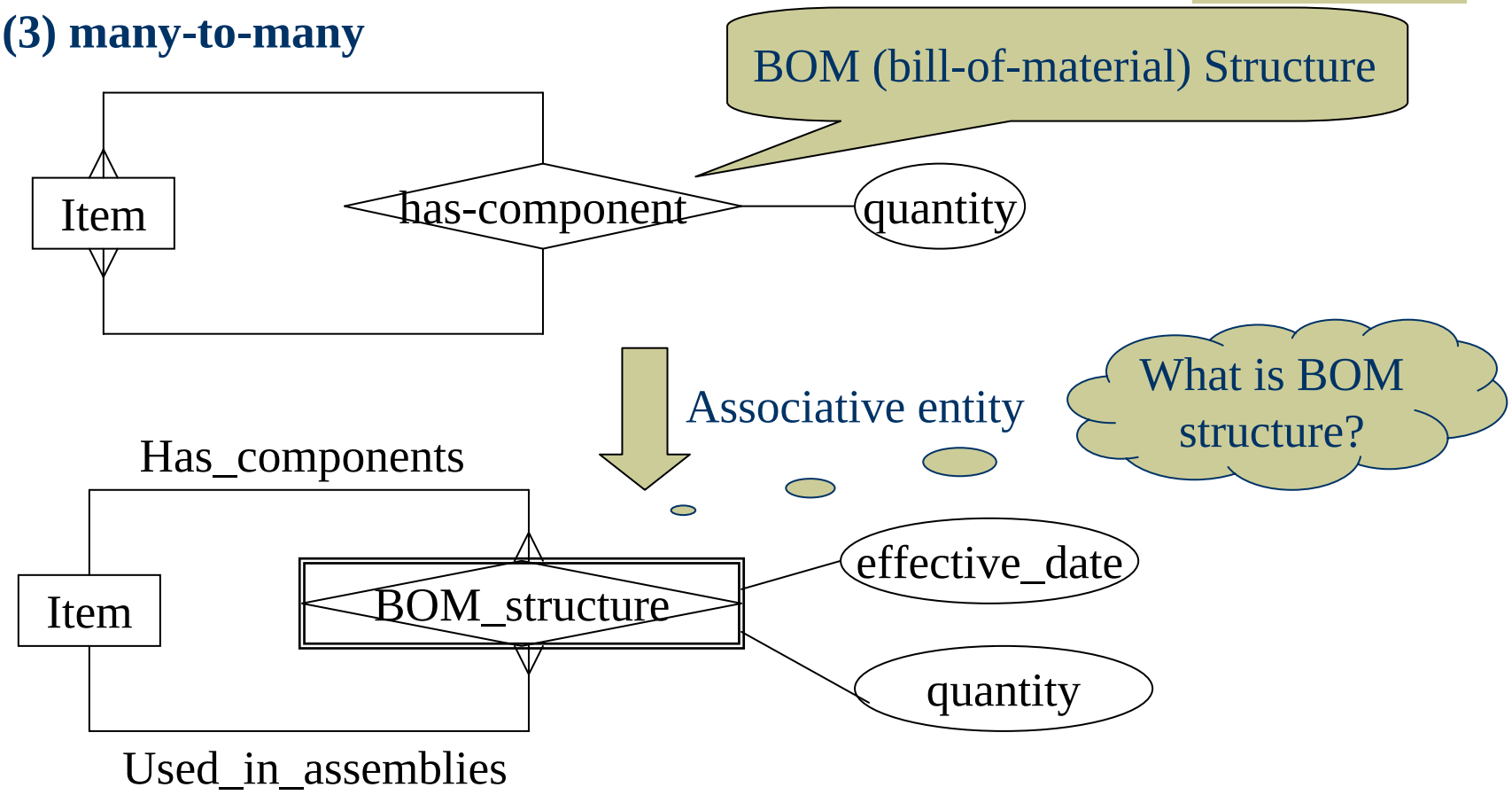
(2) One-to-many



Unary relationships are recursive relationship,
between the instances of a single entity type

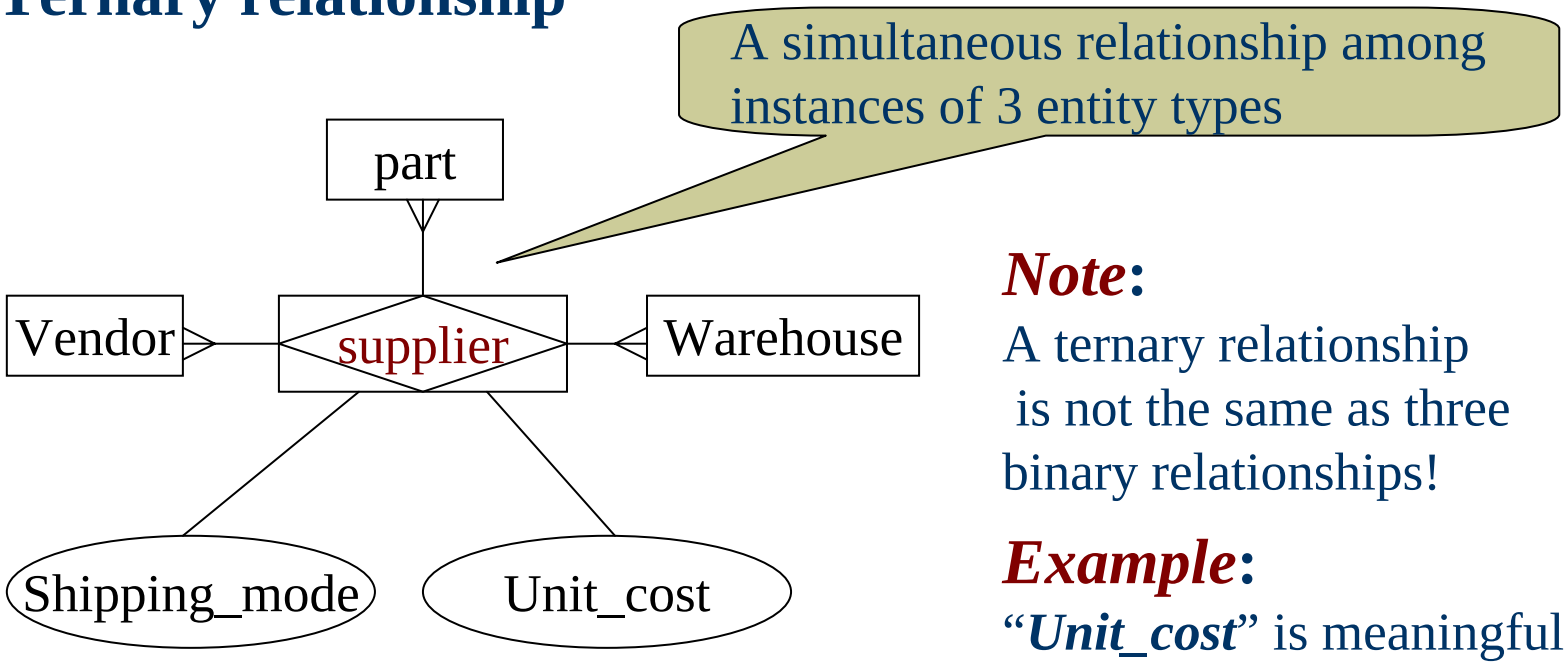
Degree of a relationship (2)

(3) many-to-many



Degree of a relationship (3)

- **Ternary relationship**



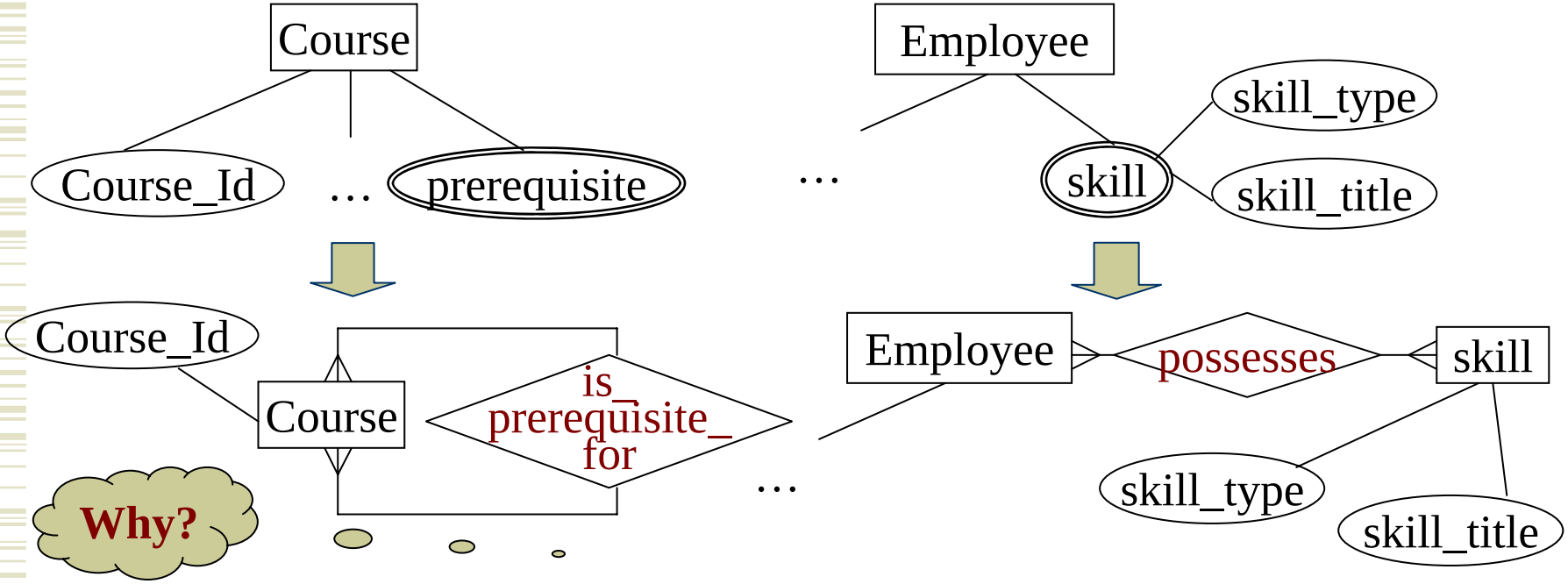
***Note*:**
A ternary relationship is not the same as three binary relationships!

***Example*:**
"***Unit_cost***" is meaningful only when it is associated with "***supplier***"

Attributes or relationships? (1)

- **When should an attribute be linked to an entity type via a relationship?**

  - If the ***multivalued*** or ***composite attributes*** are directly linked to the entity type, then **add a new relationship** linked to the entity type



Attributes or relationships? (2)

Why should the multivalued or composite attributes be linked to an entity type via a relationship?

- Decrease the redundancy of data storage
- Unify the data query processing method
- Simplify and standardize data update methods

Why can we do the conversion like this?

- No semantics loss after the conversion
- The inherent relationships are preserved
- … …

Should all the multivalued or composite attribute cases be processed like this?

Cardinality constraints

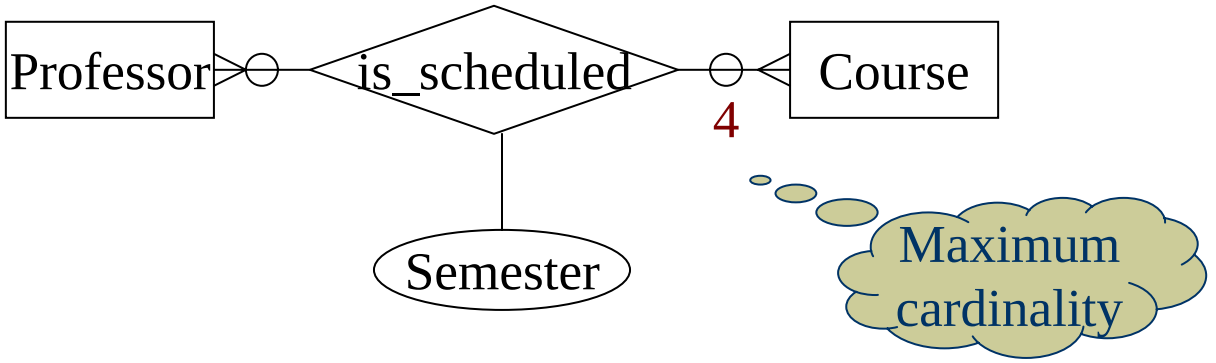
- **Cardinality constraint**

  The ***number*** of instances of entity B that can (or must) be associated with each instance of entity A

- **The range of cardinalities for a relationship**

  - Minimum cardinality (participation constraint)
  - Maximum cardinality

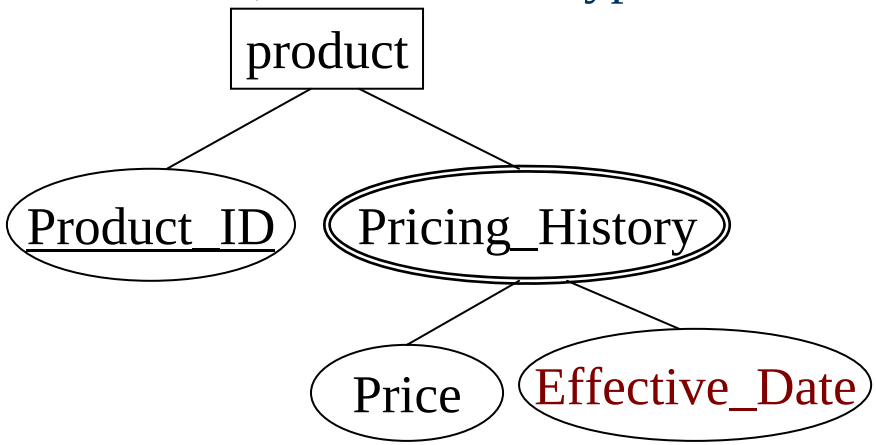
- **Cardinalities in E-R diagrams**



Modeling time-dependent data

- **Time stamp**

- A time value that is associated with a data value
- A time stamp may be associated with any data value that changes over time when we need to maintain a history of those data value
- An attribute about the time, with the data type of "Time_Stamp"
- Example:



- Frequently, time-dependent data are processed in data warehousing applications

Naming and defining in E-R diagrams

- **Entity type**
  - A singular noun
  - Specific to the organization
  - concise
  - …
- **Attribute**
  - A noun
  - Unique
  - Similar attributes should use the same qualifiers and classes (structures)
  - …
- **Relationship**
  - Verb phrase (action)
  - Avoid vague names
  - …

Summary

- Business rules as requirements
- E-R model

- Type and instance (entity type and instance, relationship type and instance)
- E-R diagram (entities, attributes, relationships; cardinality)
- Entities (strong, weak)
- Attributes (entity's, relationship's; single and composite, multivalued, derived)
- Relationships (degree, cardinality, associative entity, attributes or relationships?)
- Cardinality constraints
- Modeling time-dependent data (time stamp)
- Naming and defining in E-R diagrams

Assignments

- Page 119: 3(e), (f)

  3. Contrast the following terms:

     (e) strong entity type; relationship type

     (f) degree; cardinality

- Page 120: 3(e)

BTW*: *Some concepts should also be distinguished and clarified other than above exercises! ☺

The end

Thanks !