



## Chapter 4

### Data Modeling for Relational Database Development

การสร้างแบบจำลองข้อมูลเพื่อพัฒนาฐานข้อมูลเชิงสัมพันธ์

ผศ.ดร.เทพฤทธิ์ บัณฑิตวัฒนวงศ์

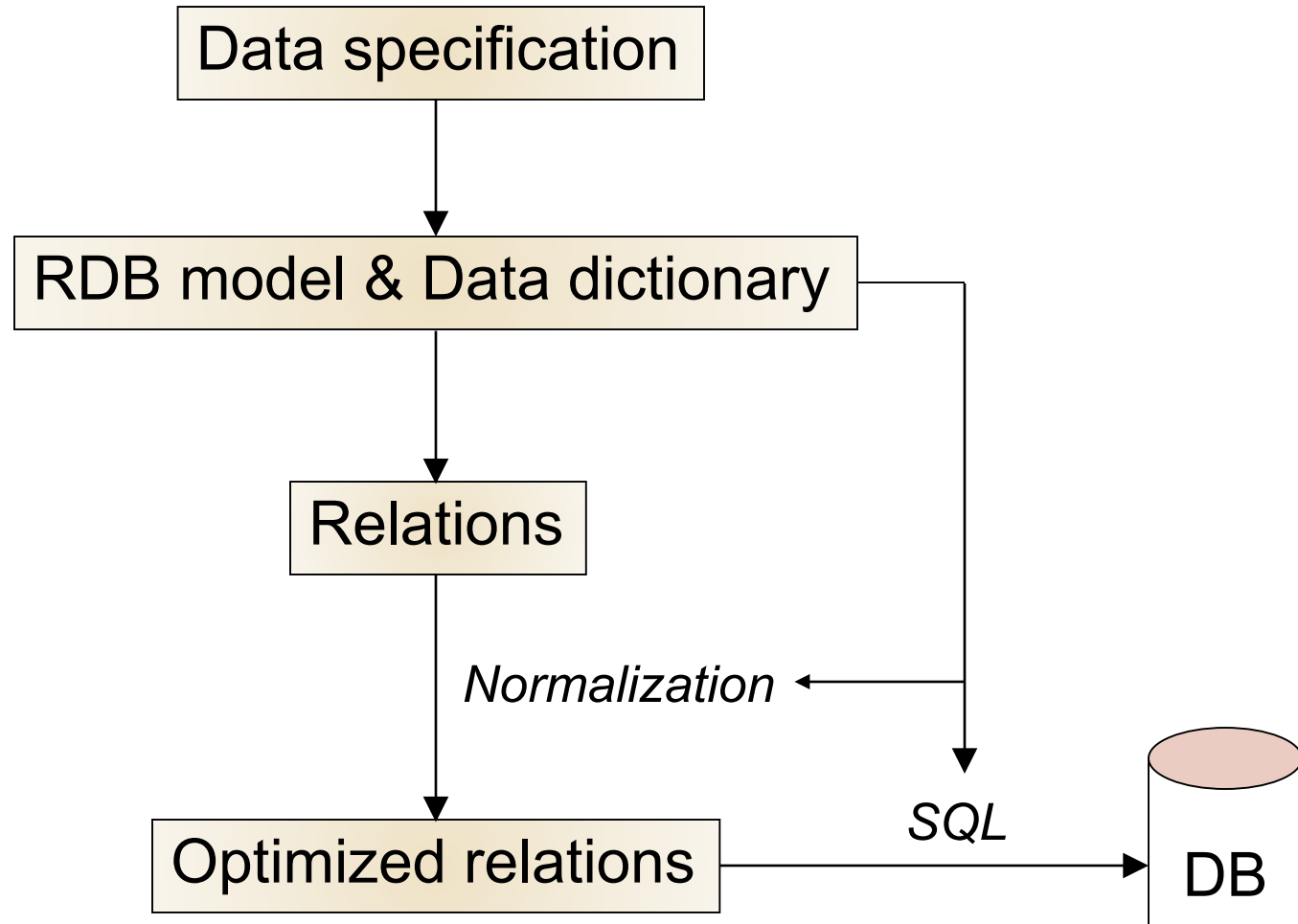
# Content

- Relational DB development process
- Data modeling concepts
- Entity-Relationship model
- Enhanced Entity-Relationship model
- Data dictionary

# Relational Database Development Process

1. Collect data requirement from users
  - Document data requirement as data specification.
2. Analyze data requirement
  - Create an RDB model and data dictionary.
3. Design database
  - Transform the RDB model into conceptual schema and normalize it as well as choose DBMS product.
4. Implement database
  - Transform the conceptual schema together with the RDB model and data dictionary into SQL codes to implement DB on the selected DBMS.

## Relational Database Development Process (cont.)



# RDB Model Creation

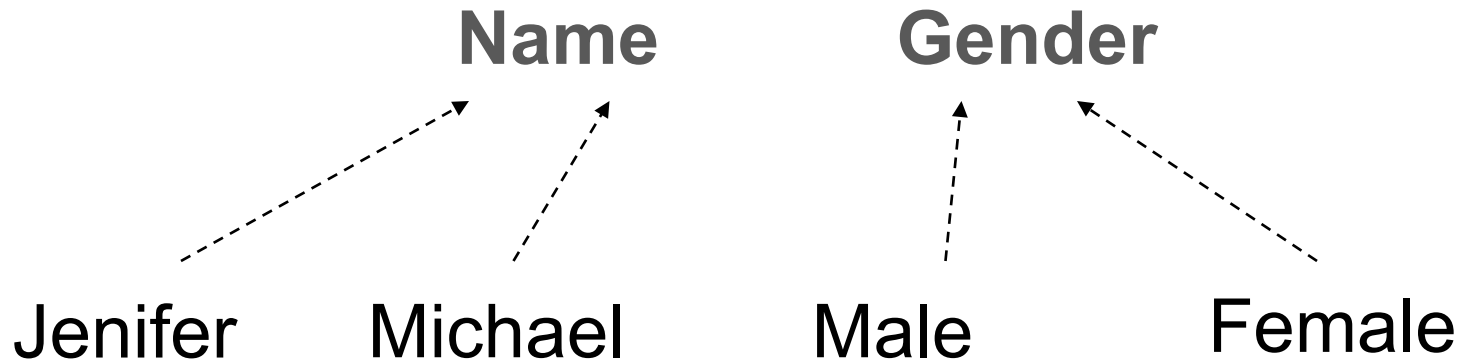
- RDB model is a part of data modeling methodology created by using data modeling language, which is based on data modeling concepts.

# Data Modeling Concepts

- Involve three types of abstraction, which is a mental process of selecting properties of a set of objects.
  1. Classification
  2. Aggregation
  3. Generalization

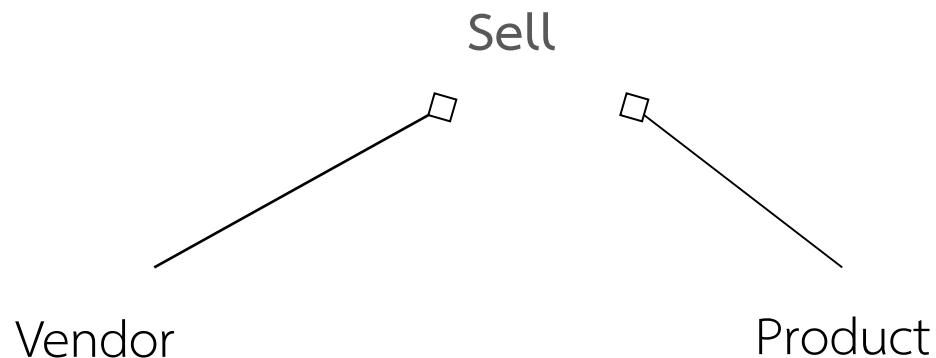
# 1. Classification

- Defines a concept as a class of objects characterized by common properties.
- Represents IS\_MEMBER\_OF relationship.



## 2. Aggregation

- Defines a new class from a set of class(es) composing it.
- Represents IS\_PART\_OF relationship.





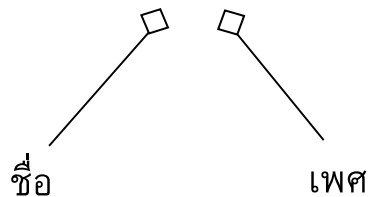
## 2. Aggregation (cont.)

- Aggregation degrees
  - Unary aggregation
  - Binary aggregation
  - Ternary aggregation
  - N-ary aggregation

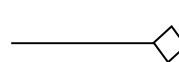
กำกับดูแล



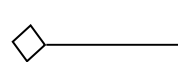
บุคคล



ผู้จำหน่าย



จำหน่าย

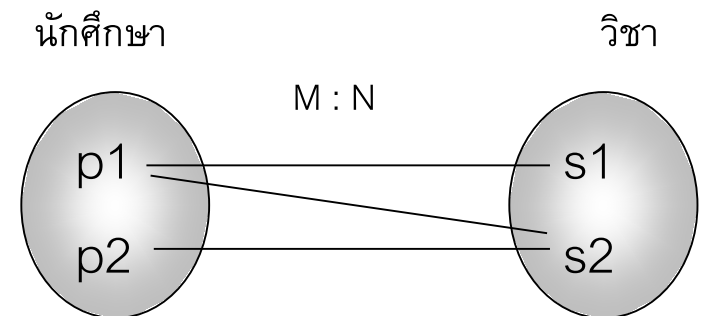
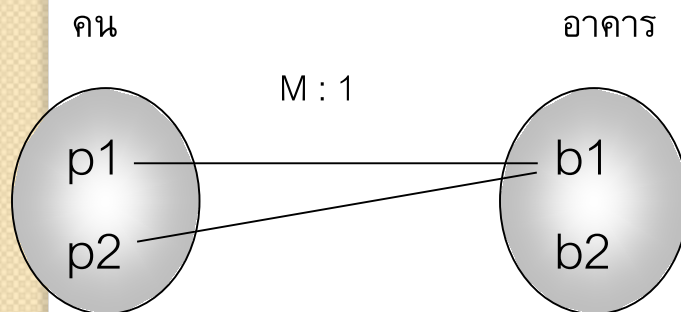
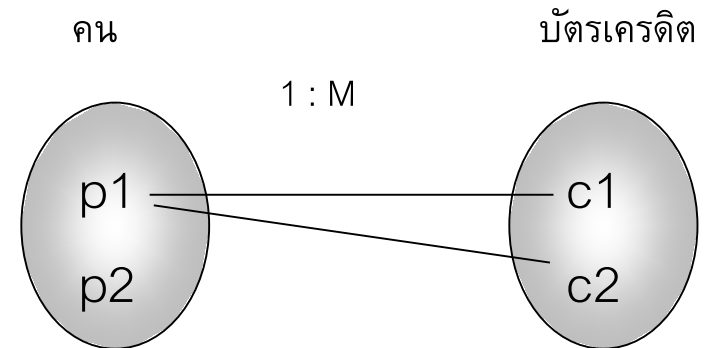
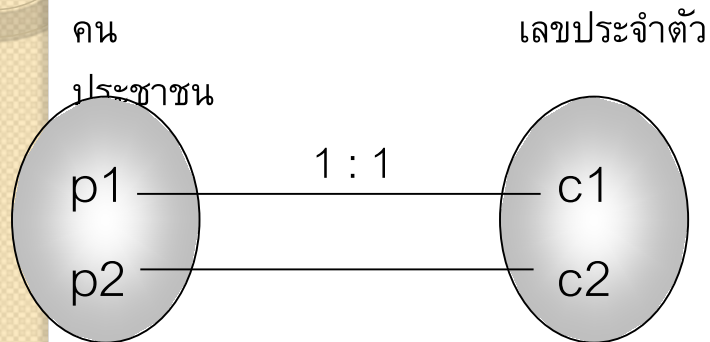


โครงการ



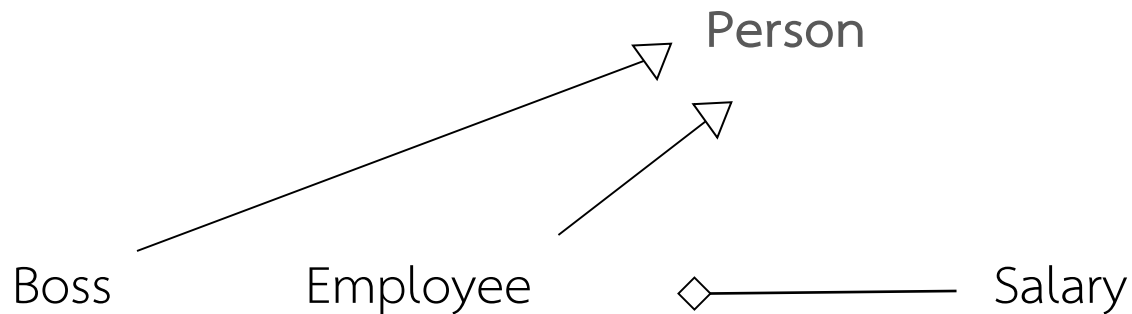
ชิ้นส่วน

# Cardinality Ratios of Binary Aggregation



### 3.Generalization

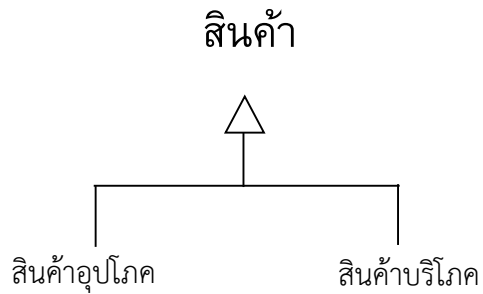
- Defines a class to be inherited by one or more specific classes (i.e., subclasses).
- Represents IS\_A relationship.



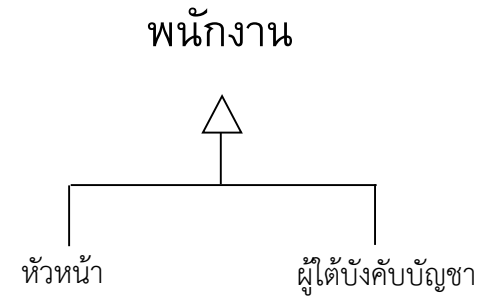
## 3.Generalization (cont.)

- Generalization coverages
  - **Total:** all objects are mapped to at least one of sub-classes.
  - **Partial:** there exists some object that is not mapped to any of sub-classes.
  - **Exclusive:** every object is mapped to at most one sub-class.
  - **Overlapping:** there exists some object that is mapped to more than one sub-class.

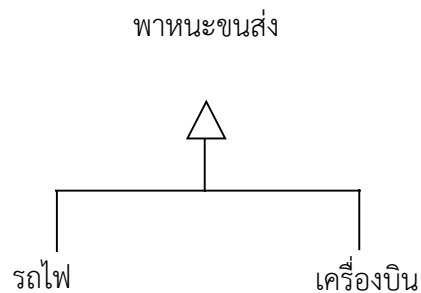
# Generalization Coverage Pairing



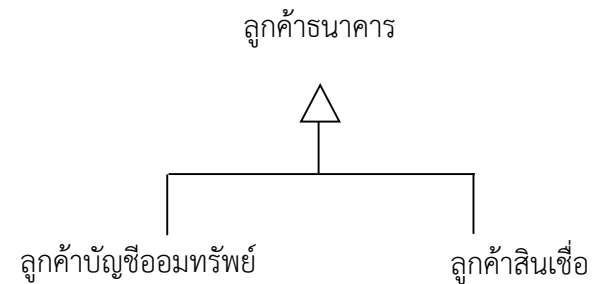
Total, Exclusive



Total, Overlapping

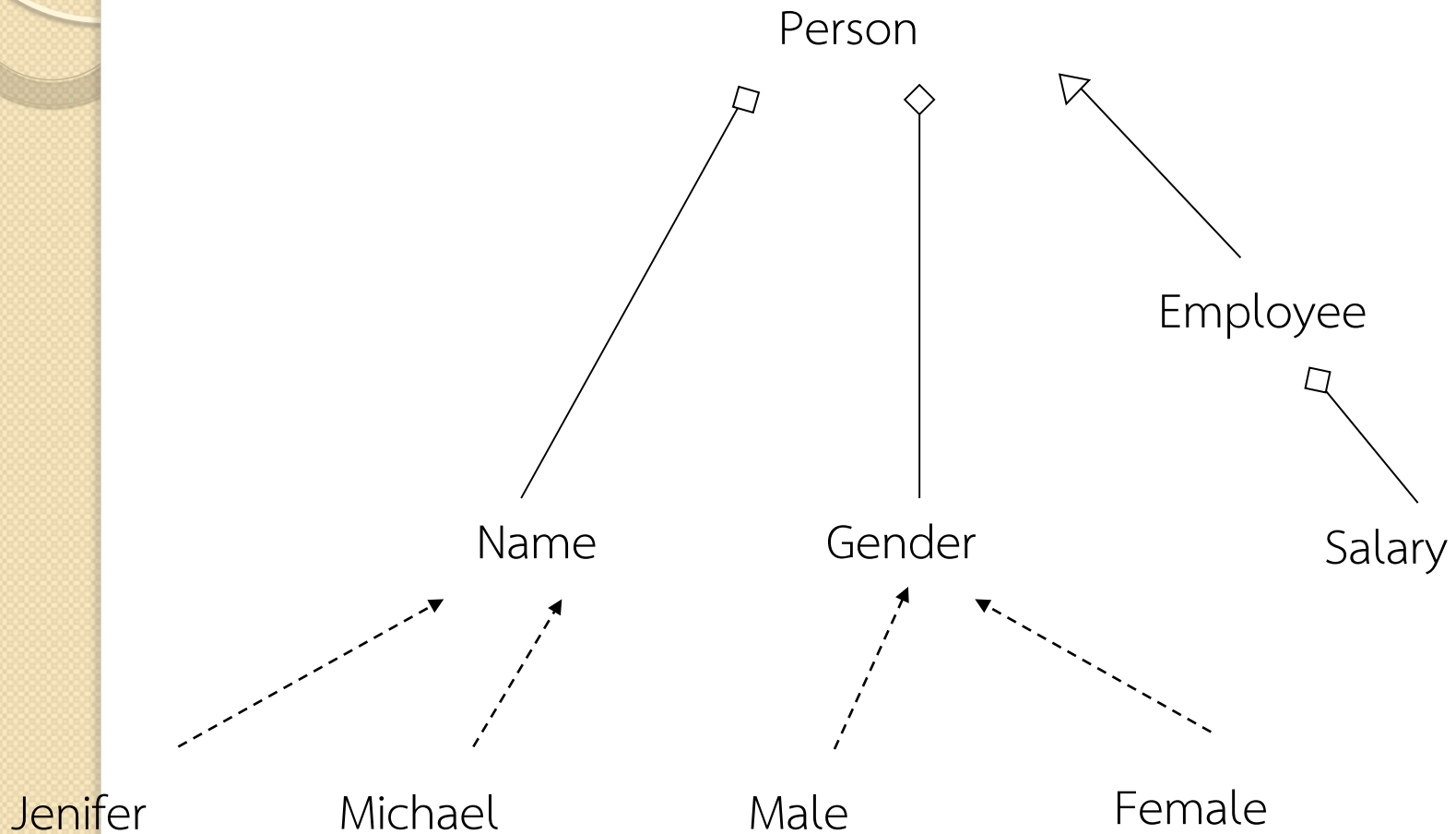


Partial, Exclusive



Partial, Overlapping

# Classification, Aggregation, Generalization



# Data Modeling Language

- A language used for creating DB models to describe user data specifications based on the data modeling concepts.
- Analogous to drawing language used in house blueprints.

# Entity-Relationship (ER) Model

- A de-facto standard data modeling language for creating RDB models.
- Most widely used for conceptual database design (i.e., for developing conceptual schemas).
- Originally invented by Prof. Peter Pin-Shan Chen in 1976.
- Comprises primitive constructs and syntax.
- There are currently several ER model variants. We will be learning the original one.



# ER Model (cont.)

- Primitive constructs
  1. Entity
  2. Relationship
  3. Attribute

# 1.Entity

- A tangible or intangible thing whose attribute(s) are to be stored in a database.
- Created by means of classification abstraction.
- Represented graphically by rectangle with capital name.



PUBLISHER

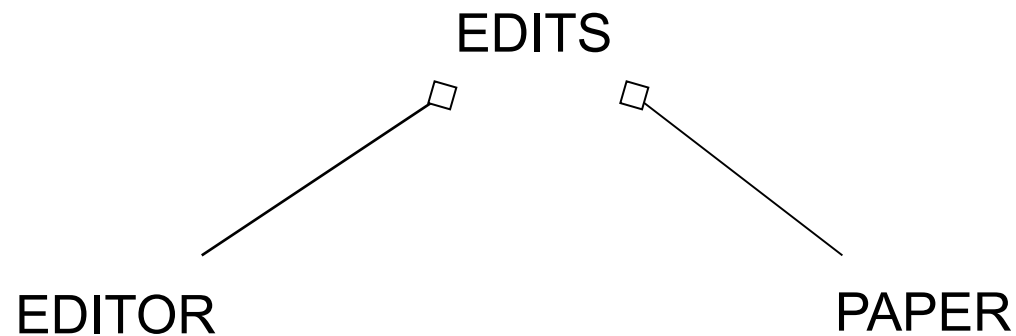
The diagram shows two rectangular boxes representing entities. The first box on the left contains the word 'PUBLISHER' in all capital letters. The second box on the right contains the word 'STUDENT' in all capital letters. Both boxes are simple rectangles with a thin black border.

STUDENT

- In each ER model, there can be no duplicated entity name.
- **Entity instance:** an occurrence of an entity.

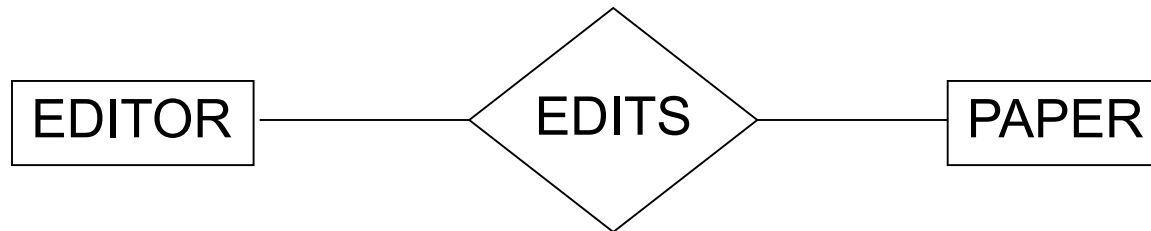
## 2.Relationship

- An aggregation abstraction of one or more entities.



## 2.Relationship (cont.)

- Represented graphically by a named diamond-shape symbol.



## 2.Relationship (cont.)

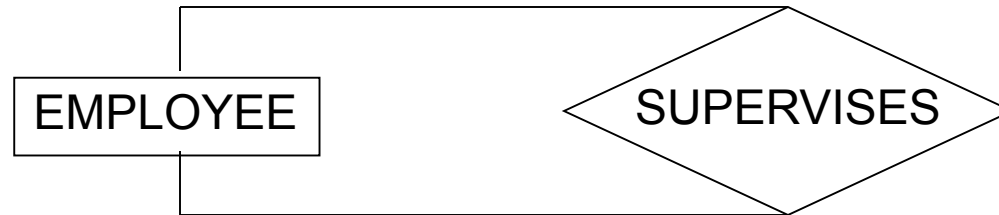
- Its name represents a role played by an entity aggregated in the relationship.
  - Use capital letters.
  - Use underscore to denote white space.
  - Usually a verb word.
  - Conventionally arranged for left-to-right or top-to-bottom readings.

## 2.Relationship (cont.)

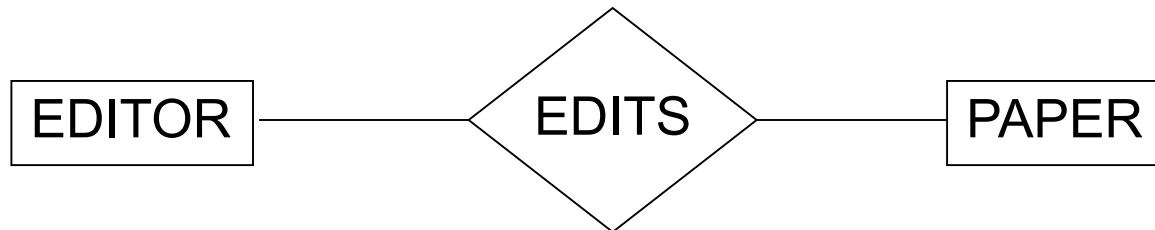
- Described in terms of:
  - **Degree:** a number of entities participating in an aggregation.
  - **Connectivity:** a cardinality ratio of entities in an aggregation.
  - **Existence:** a presence of entities in aggregation.
  - **Structural constraint:** minimum and maximum instance numbers of each entity in aggregation e.g. (1, 1), (1, N), (2, 5)

# Degree

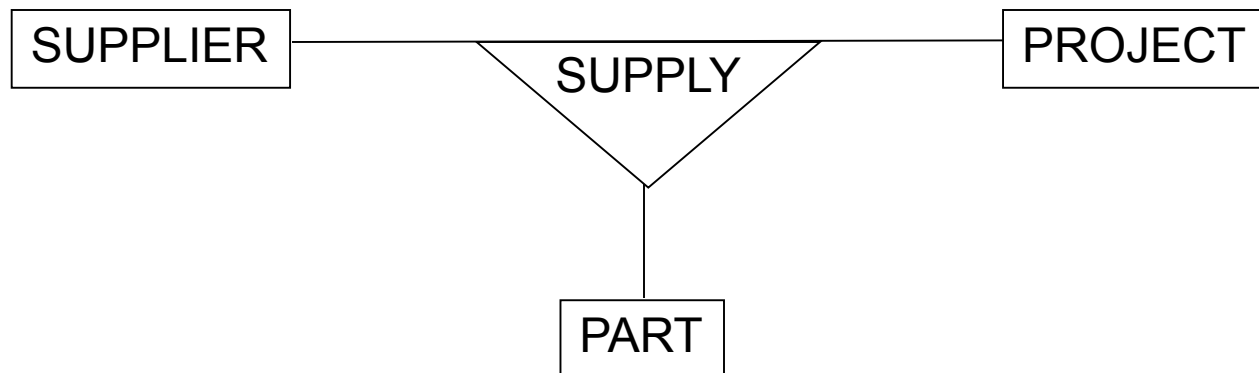
- Unary



- Binary



- Ternary

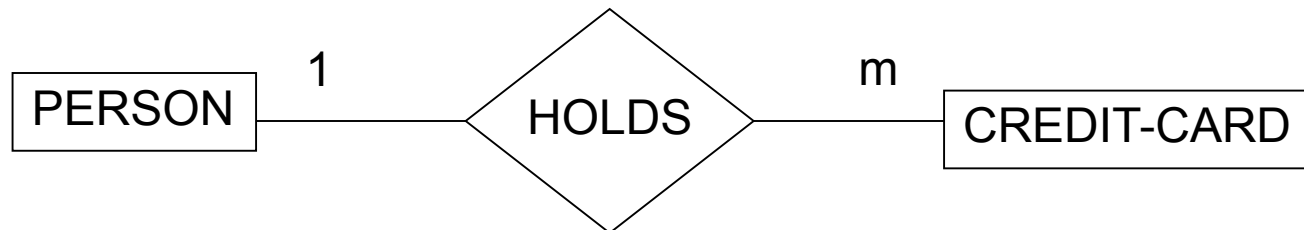


# Connectivity

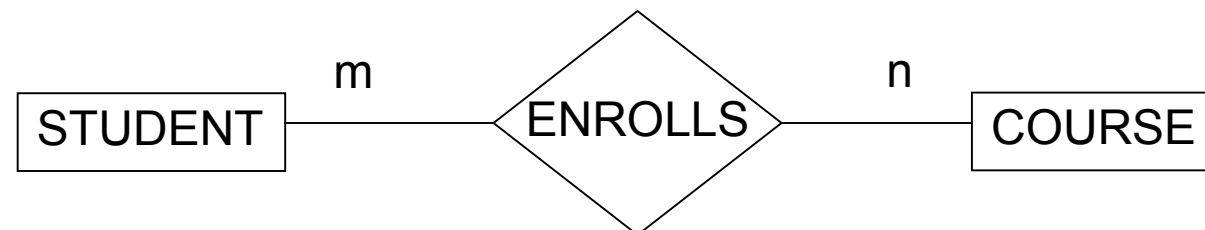
- One-to-one



- One-to-many



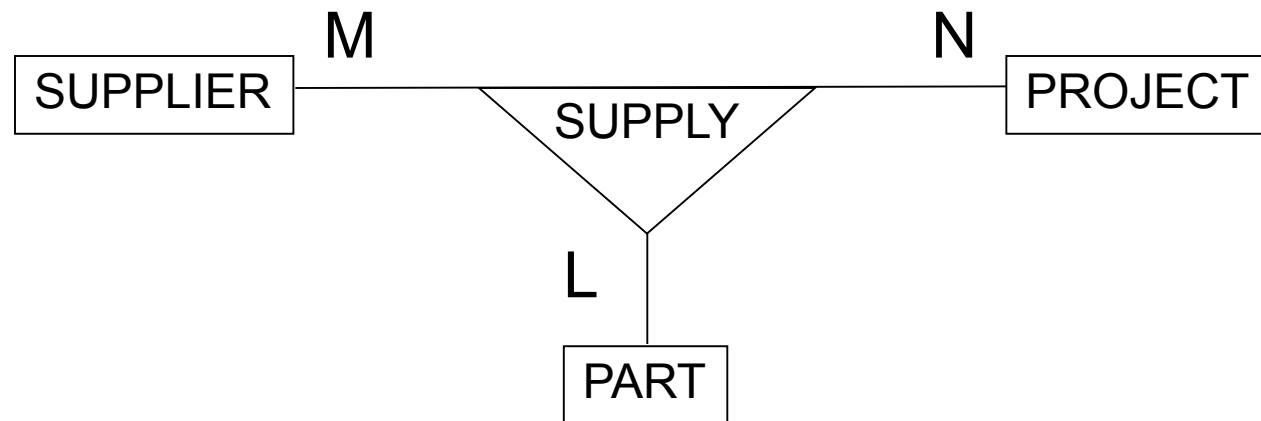
- Many-to-many





## Connectivity (cont.)

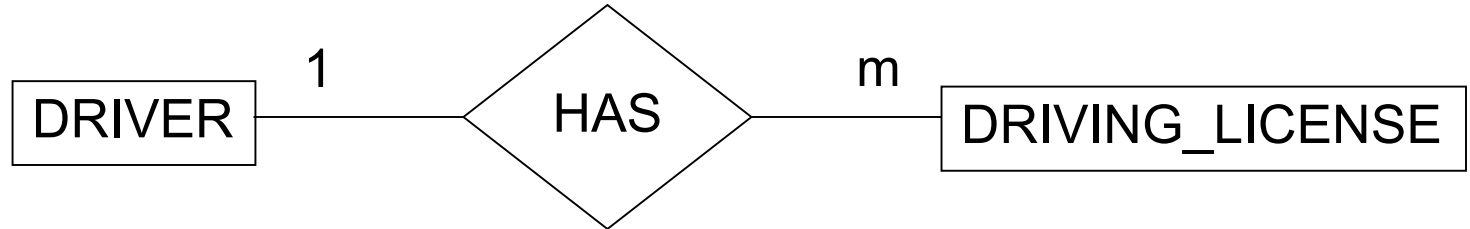
- Many-to-many-to-many



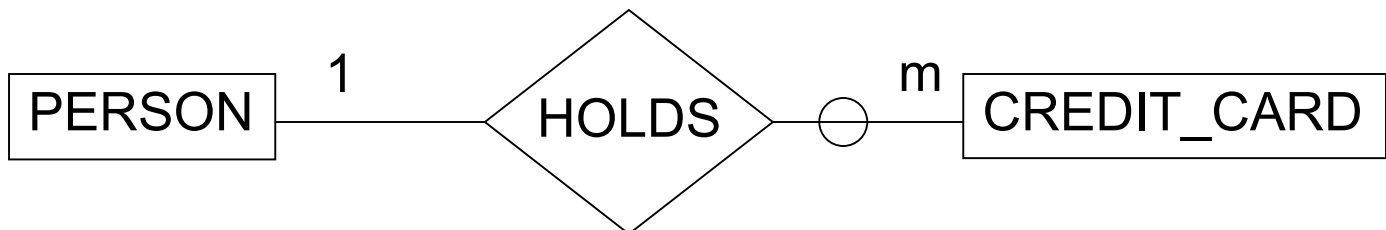
- In ternary relationship, there can also be 1-1-1, 1-1-m, 1-m-n, m-1-1, m-1-n, etc.

# Existence

- **Mandatory:** all entity instances must participate in the relationship.



- **Optional:** some instances of entity may not participate in the relationship. Indicated by small circle on relationship arm.

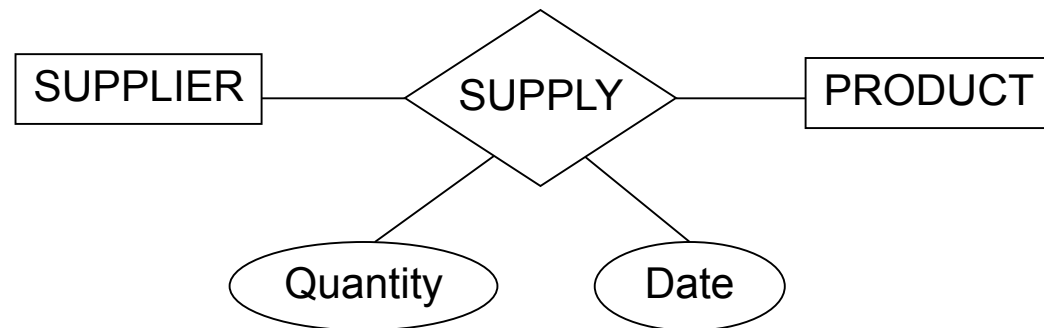
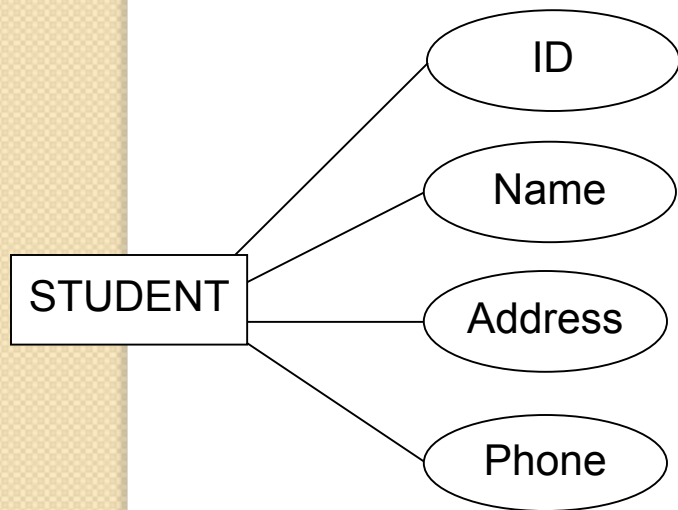


# 3.Attribute

- A characteristic of entity or relationship that provide descriptive detail about it.
- Created by means of aggregation abstraction.
- An occurrence of an attribute is attribute value.
- Examples
  - Attributes of entity STUDENT may include ID, name, address, and phone number.
  - Attributes of relationship SUPPLY that connects entities SUPPLIER and PRODUCT can be Quantity and Date.

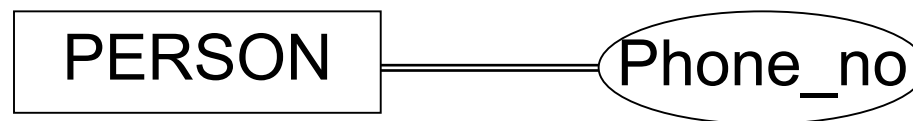
### 3.Attribute (cont.)

- Represented graphically by oval-shape symbol containing first-capital-letter name.



### 3.Attribute (cont.)

- **Single-valued attribute** is a simple attribute, which has a single value.
- **Multivalued attribute** has multiple values and represented graphically by using double lines linking between multivalued attribute to aggregated entity.

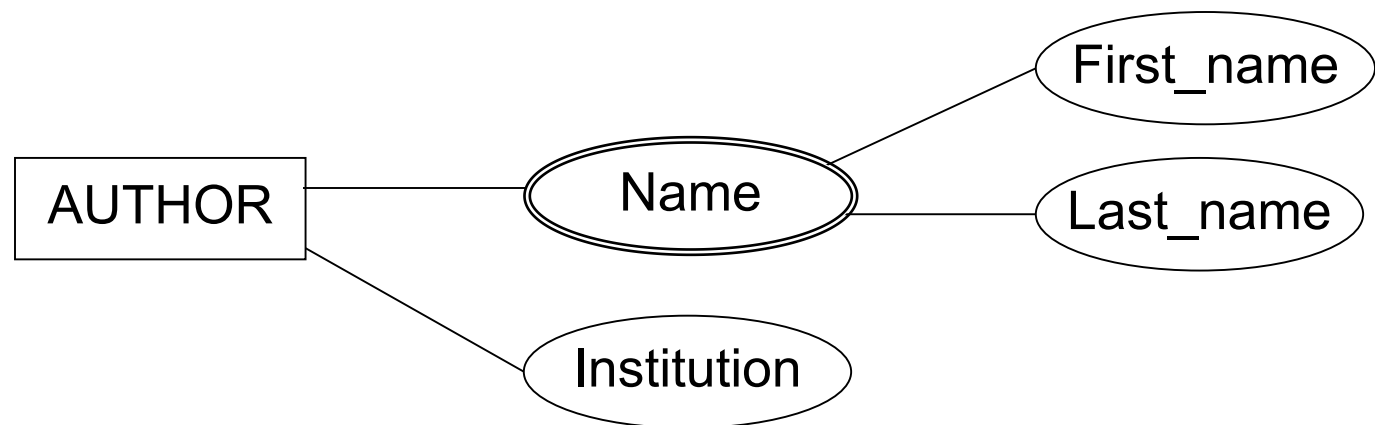


### 3.Attribute (cont.)

- Every entity must have a primary key attribute (underlined).

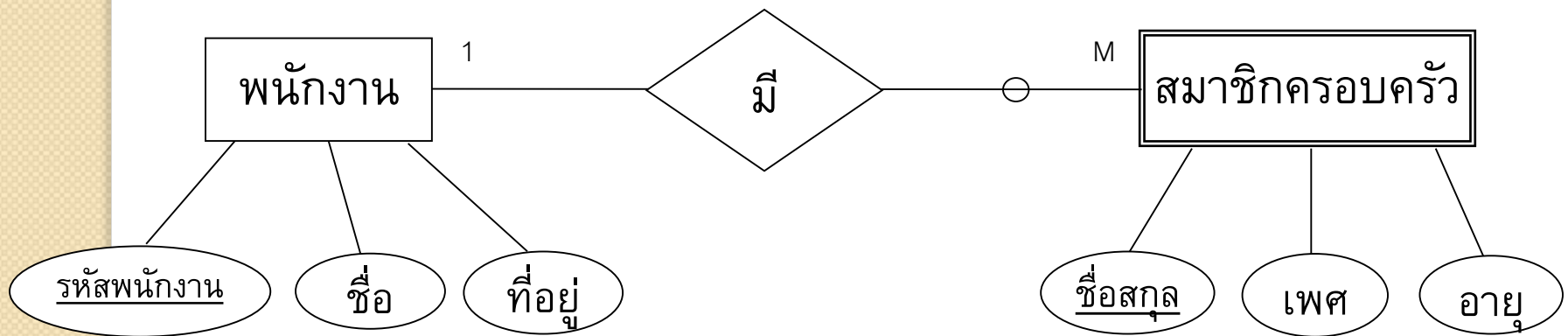


- **Composite attribute** consists of multiple attributes.



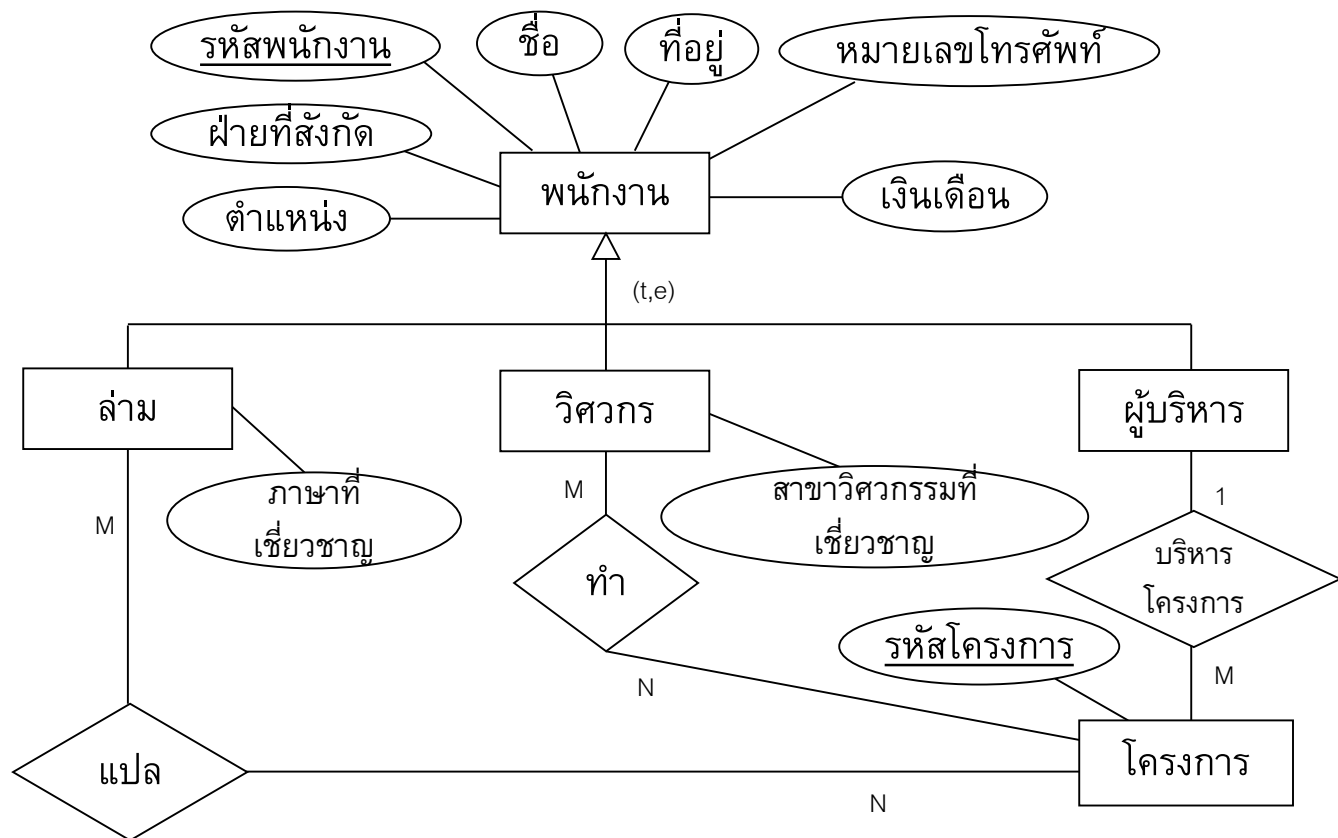
## Entity (cont.)

- **Strong entity** has a its own primary key.
- **Weak entity** must co-exist with an identifying strong entity, and needs external identifier to compose its own primary key. Weak entity is represented by double-border rectangle.



# Enhanced ER (EER) Model

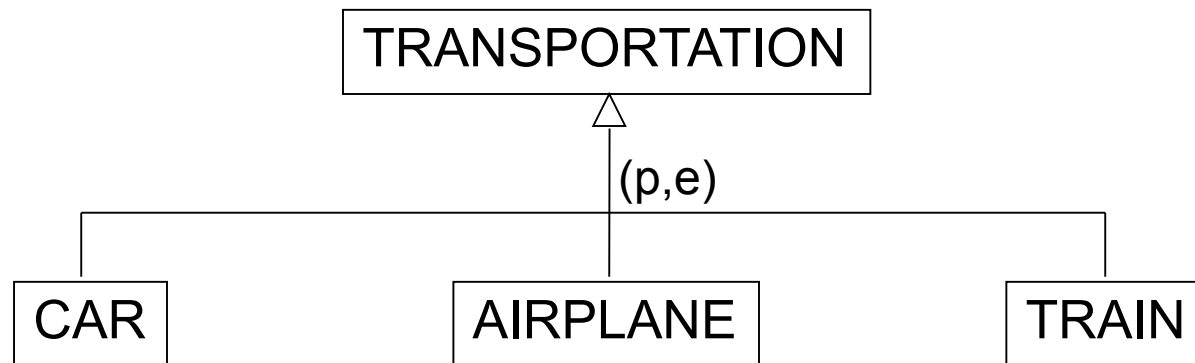
- A model that extends ER model by incorporating object-oriented concepts such as generalization and multiple inheritance.





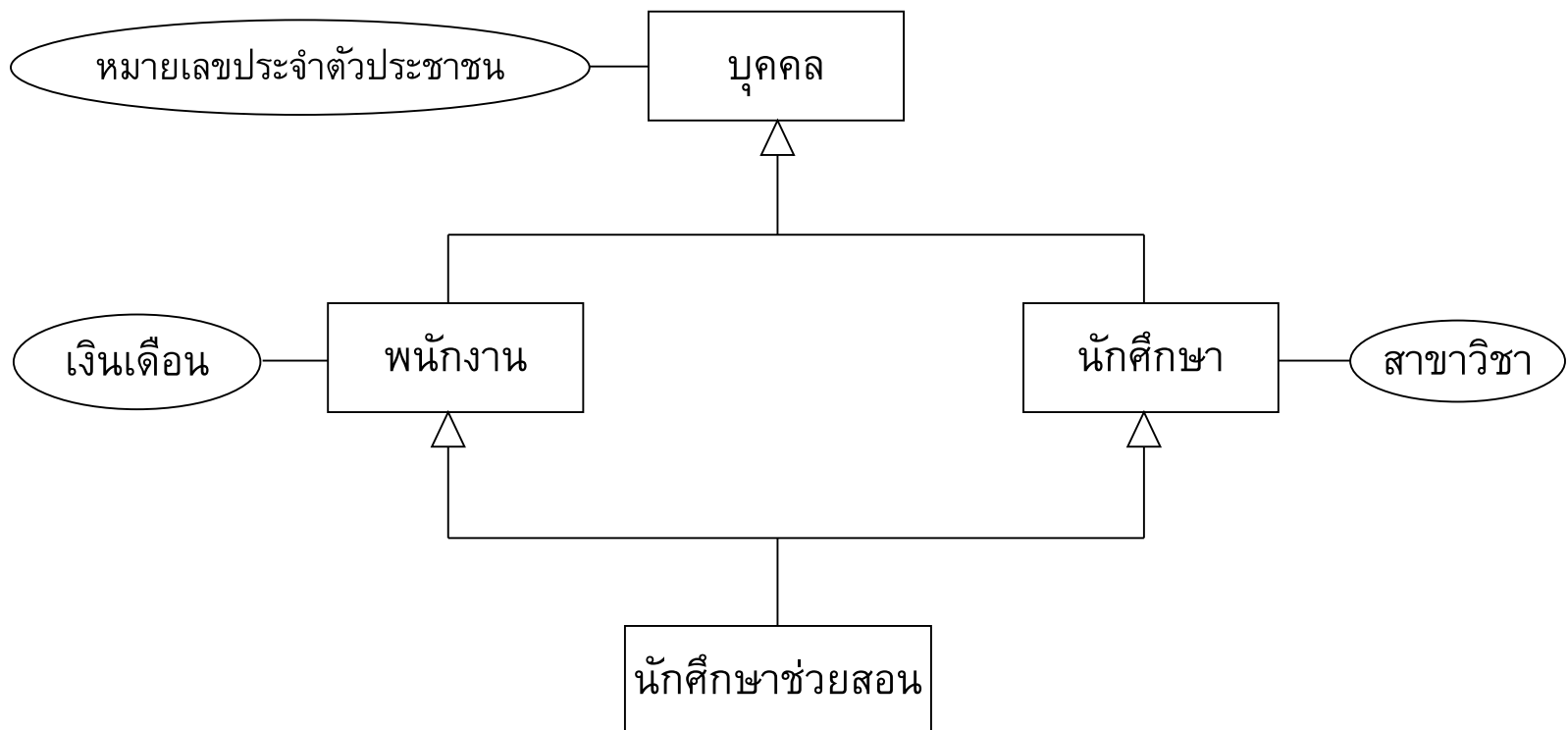
## EER Model (cont.)

- **Single inheritance** represents generalization abstraction between each pair of entities.
  - Use triangle-head arrow symbol with associated coverage in parenthesis indicating that child entities inherit all attributes from its parent entity.



## EER Model (cont.)

- **Multiple inheritance** symbol has multiple arrow heads indicating that a child entity inherits all attributes from its all parent entities.



# ER Modeling Software Tool

- Several CASE tools are available for different variants of ER models.
- This course prefers the original ER model.
  - Microsoft Visio and add-on stencil
    - [http://cs.winona.edu/faculty/NAnderson/385/ER Diagram.zip](http://cs.winona.edu/faculty/NAnderson/385/ER%20Diagram.zip)

# Data Dictionary

- A meta-data providing text description of all primitive constructs in an ER model.
  - Prevents readers from misunderstanding ER model.
  - Provides additional information not appearing in ER model such as constraints.
- Always created in conjunction with an ER model.

# Data Dictionary (cont.)

: For the ER model in the 5<sup>th</sup> previous slide.

เอนทิตีหรือความสัมพันธ์	ลักษณะประจำ	คำอธิบาย	เงื่อนไขบังคับ
พนักงาน		พนักงานของบริษัท	
	รหัสพนักงาน		กุญแจหลัก
	ชื่อ		ห้ามค่าว่าง
	ที่อยู่	ที่อยู่ติดต่อได้	ห้ามค่าว่าง
	หมายเลขโทรศัพท์	หมายเลขโทรศัพท์ติดต่อภายในหรือโทรศัพท์เคลื่อนที่	ห้ามซ้ำ, ห้ามค่าว่าง
	ฝ่ายที่สังกัด		ห้ามค่าว่าง, เขตข้อมูล = (ฝ่ายบริหาร, ฝ่ายสนับสนุน, ฝ่ายการผลิต)
	ตำแหน่ง		เขตข้อมูล = (ผู้จัดการบริษัท, หัวหน้าฝ่าย, เจ้าหน้าที่ปฏิบัติการ), ค่าโดยปริยาย = เจ้าหน้าที่ปฏิบัติการ
	เงินเดือน		>= 20,000 บาท
ล้าม		พนักงานแปลภาษาระหว่างการสนทนา	
	ภาษาที่เชี่ยวชาญ	ภาษาที่สองที่ชำนาญยกเว้นภาษาไทย	ห้ามค่าว่าง, เขตข้อมูล = (ภาษาอังกฤษ, ภาษาญี่ปุ่น)
วิศวกร		พนักงานทำให้เกิดผลโครงการ	
	สาขาวิศวกรรมที่เชี่ยวชาญ		ห้ามค่าว่าง, เขตข้อมูล = (เครื่องกล, ไฟฟ้า, ซอฟต์แวร์)
ผู้บริหาร		พนักงานระดับบริหารโครงการ	
โครงการ		โครงการพัฒนาเทคโนโลยีที่บริษัทรับจ้าง	
	รหัสโครงการ		กุญแจหลัก
แปล		รับผิดชอบหน้าที่ล้ามประจำโครงการ	
ทำ		ทำงานวิศวกรรมประจำโครงการ	
บริหาร		ควบคุมดูแลโครงการ	

# Exercises

1. อธิบายส่วนประกอบปฐมฐานเอนทิตีทั้งหมดของแบบจำลองข้อมูลอีอาร์พร้อมวาดภาพประกอบโดยใช้ชุดสัญลักษณ์ตามที่นำเสนอในหนังสือด้วยเครื่องมือวิซิโอ
2. อธิบายส่วนประกอบปฐมฐานความสัมพันธ์ของแบบจำลองข้อมูลอีอาร์และส่วนประกอบปฐมฐานของแบบจำลองอีอีอาร์พร้อมวาดภาพประกอบโดยใช้ชุดสัญลักษณ์ตามที่นำเสนอในหนังสือด้วยเครื่องมือวิซิโอ
3. อธิบายส่วนประกอบปฐมฐานลักษณะประจำทั้งหมดของแบบจำลองข้อมูลอีอาร์พร้อมวาดภาพประกอบโดยใช้ชุดสัญลักษณ์ตามที่นำเสนอในหนังสือด้วยเครื่องมือวิซิโอ
4. ยกตัวอย่างพร้อมอธิบายเอนทิตีอ่อนแอเพิ่มเติมจากในหนังสืออีกหนึ่งตัวอย่าง
5. ยกตัวอย่างความสัมพันธ์แบบทวิภาคที่มีภาวะการเชื่อมต่อแบบหนึ่งต่อหนึ่ง, หนึ่งต่อมาก และมากต่อมาก อย่างละหนึ่งตัวอย่าง
6. ยกตัวอย่างความสัมพันธ์แบบทวิภาคที่มีการดำรงอยู่แบบเงื่อนไขบังคับการมีส่วนร่วมเชิงเลือกได้มาสามตัวอย่าง
7. ยกตัวอย่างความสัมพันธ์แบบไตรภาคในโลกความเป็นจริง ที่มีภาวะการเชื่อมต่อแบบหนึ่งต่อหนึ่งต่อหนึ่ง, มากต่อหนึ่งต่อหนึ่ง, มากต่อมากต่อหนึ่ง และมากต่อมากต่อมาก อย่างละหนึ่งตัวอย่าง