The Entity-Relationship Model (Part I)

R &G - Chapter 2

Steps in Database Design

1. **Requirements Analysis**
   - user needs (non-professional clients)
2. **Conceptual Design**
   - high level description
3. **Logical Design**
   - translate ER into DBMS data model
4. **Schema Refinement**
   - Remove redundancy, normalization

Design a database

Design a GOOD database

Today's Class

1. Requirements Analysis
   – user needs; what must database do?
2. Conceptual Design
   – high level description

Conceptual Modeling: Informal Explanations

- ER model: (E -*entities*, R – *relationships)*

Informally

- Entities are usually (relevant) <u>nouns</u>
  - "Stevens has full-time and part-time students."
  - "Stevens has around two hundreds faculty members."
  - **Question**: what are the nouns/entities in the examples above?

- Relationships are statements about 2 or more objects.
  - Often, <u>verbs</u>.
  - e.g., "a prof teaches a course"
  - **Question**: what is the verb/relationship in the example?

Conceptual Modeling: Informal Explanations (Cont.)

- Besides entities & relations, what else?
  - What information about entities & relationships should be stored in the database?
  - What rules (called *integrity constraints*) should hold on these entities and relationships?

- In relational databases, this is generally encoded in an **Entity-Relationship (ER) Diagram**

# Today's Class

Basic concepts of ER diagram
- Entity set
- Key of entity set
- Relationship set

Entity and Entity Set

- *Entity:* Real-world object, distinguishable from other objects.
  - An entity is described using a set of *attributes*.

- *Entity set*: A collection of similar entities. E.g., all employees.
  - All entities in an entity set have the same set of attributes.
 - Each entity set has a *key* that can uniquely identify each entity.

Draw An Entity Set in ER Diagram

- **Entity set**
  - Draw in a rectangle;
  - Name of entity set: either singular or plural format is fine.

- **Attributes of entity set**
  - Draw in ovals, one attribute per oval;
  - Ovals are connected with the entity set rectangle with solid lines

- **Key of entity set:**
  - Underlined in ER diagram (more will be discussed later)
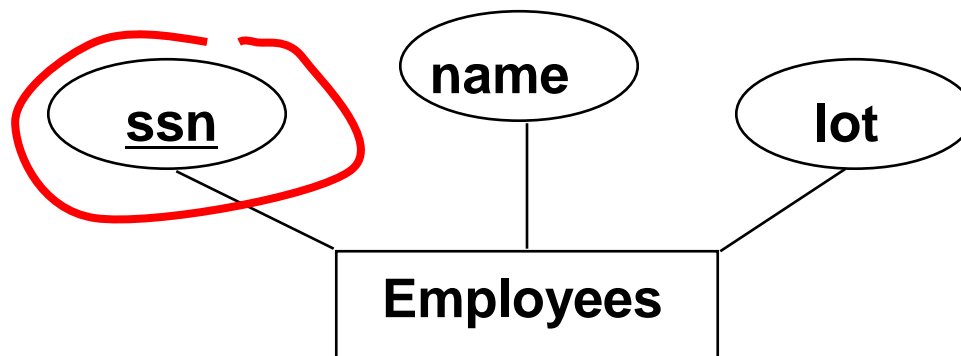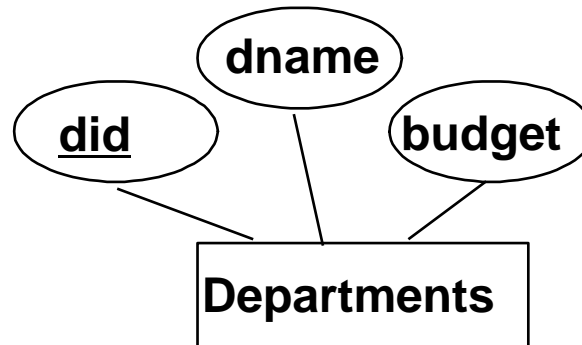
Draw An Entity Set in ER Diagram (Cont.)

- **No attribute sharing among multiple entity sets**
  - Each entity set has its own attributes
  - Different entity sets can have the attributes of the same name
    - E.g., both students and staff have the attribute name
  - Attributes (even with the same name) are never shared among entity sets in ER diagram.
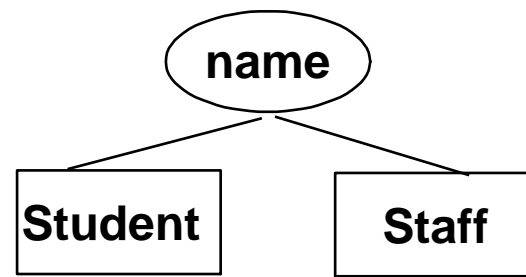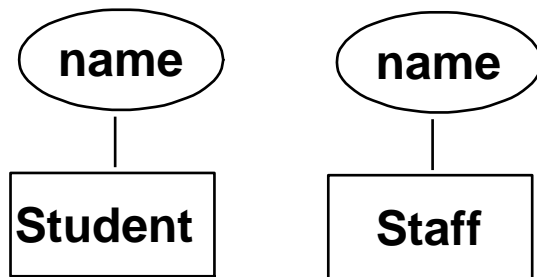
Exercise: ER

- What is the entity set?
- What are the attributes?
- What is the key?

Exercise: ER Design (1)

- The supermarket ShopRite! hires you to design a database for its product inventory. Each product has an ID, its description, category, price, and expiry date.

- **Question:**
  - What is the entity set?
  - What are the attributes of the entity set?
  - Draw the ER diagram.

Exercise: ER Design (2)

- Assume you want to design a student directory (a prototype of Facebook II)
- The main concepts you need to describe include
  - Users (ID, age, gender, school)
  - Friend & enemy relations among users
- **Questions**
  - Draw the ER diagram (*only the Entity set for now; relations for later*).

Today's Class

Basic concepts of ER diagram

•Entity set

•Key of entity set

•Relationship set

Three Types of Keys for Entity Set

- **Superkey**
- **Key**
- **Primary key**

Superkey

- A **superkey** = a set of attributes which, taken collectively, identify *uniquely* an entity in an entity set
  - All attributes of an entity set E compositely is a superkey of E.
   - E.g. *(did, dname, budget)* together is a superkey of the entity set *Departments*
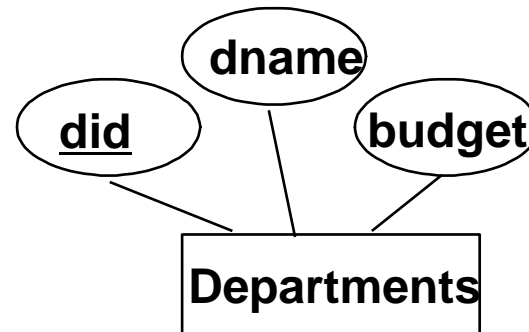
    - A subset of superkey may NOT be a superkey
      - E.g., (budget) cannot be a superkey as different departments can have the same budget

15

Key

A **key** (or *candidate* key) = a superkey for which no proper subset is a superkey (i.e., **minimal** superkey)

- Each entity set must have at least one key!
- E.g. (did) is a key of the entity set *Departments*
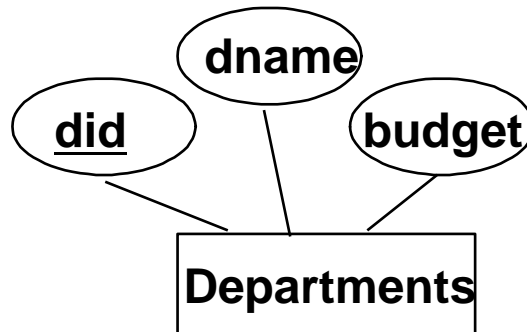
- Any superset of a key is a superkey but not a key (because it is not minimal)
  - E.g., (did,dname) is a superkey but not a key

Singleton and Composite Keys

- Singleton key: the key only consists of one attribute
  - E.g., (SSN), (Student ID)

- Composite key: the key consists of more than one attribute
  - E.g. No students in the same class have identical names
  - Key: (Name, Class)

- Each key should be included in parentheses, regardless singleton or composite.

Primary Key

- There can be more than one key for an entity set
  - E.g., each department has a unique *did* as well as a unique *dname*

- A *primary* key = the key chosen as the principal means to identify entities in an entity set
  - ONLY the primary key is underlined in ER diagram
    - Singleton key: underline the single attribute in the key
    - Composite key: underline all the attributes in the key
    - E.g., (Name, Class) is the primary key of the entity set Student
  - Other keys do not need to be drawn in ER diagram
    - But these keys need to be specified in database design (will discuss later).

Exercise: Keys
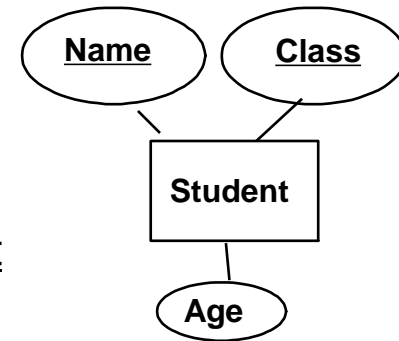

```
         name
   ssn            Parking_lot

        Employees
```
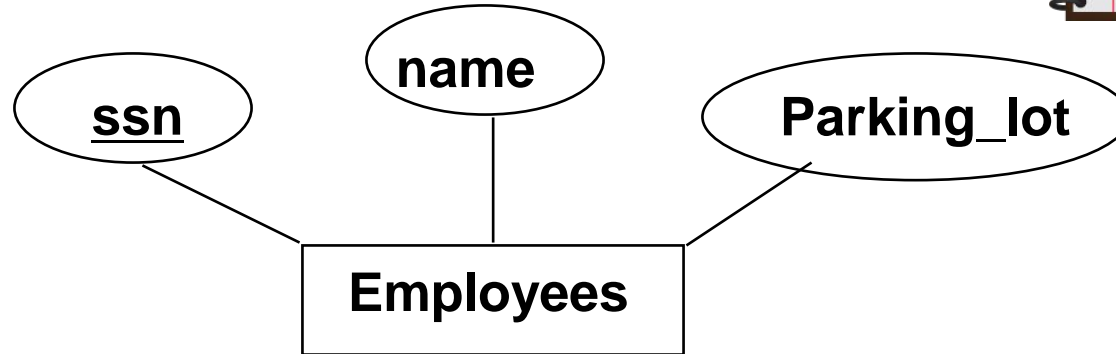
- Is (ssn) a superkey?
- Is (ssn) a key?
- Is (ssn, name, parking_lot) a superkey?
- Is (ssn, name, parking_lot) a key?

Last Lecture

Basic concepts of ER models

- Entities and entity set
- Keys
  - *Superkey*: a set of attributes that can uniquely identify each entity
  - *Key*: minimal superkey
    - It must be a superkey: it can uniquely identify each entity
    - It must be minimal: none of its subset can be a superkey
  - *Primary key*
   - The key that is picked to be underlined in ER diagram

Exercise: Key and Primary Key

Consider the following information of the students:
- Each student has a unique student ID;
- Each student has a unique locker number;
- Some students have the same name;
- Some students have the same birth date;
- No student have the same combination of name and birth date;

Questions:
- What are the keys of the *students* entity?
- What is the primary key of the entity?
  - Multiple correct answers
- Draw the primary key in the ER diagram
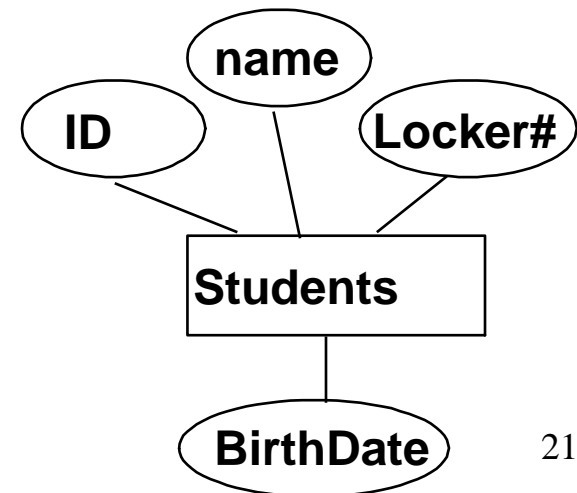
Today's Class

Basic concepts of ER diagram

• Entity set

• Key of entity set

• Relationship set

Relationships and Relationship Set

- ***Relationship*: Association among two or more entities.**
  - E.g., Alan *works in* Pharmacy department. Alice *works in* HR department

- ***Relationship Set*: Collection of similar relationships.**
  - An n-ary relationship set R relates n entity sets $E_1 \ldots E_n$; each relationship in R involves entities e_1 in E_1, ..., e_n in E_n
  - Relationship sets can have their own attributes.

Draw An Relationship Set in ER Diagram

- **Relationship set**
  - Draw in a diamond;
- **Attributes of relationship set**
  - Draw in ovals, one attribute per oval;
  - Ovals are connected with the relationship set diamond by solid lines
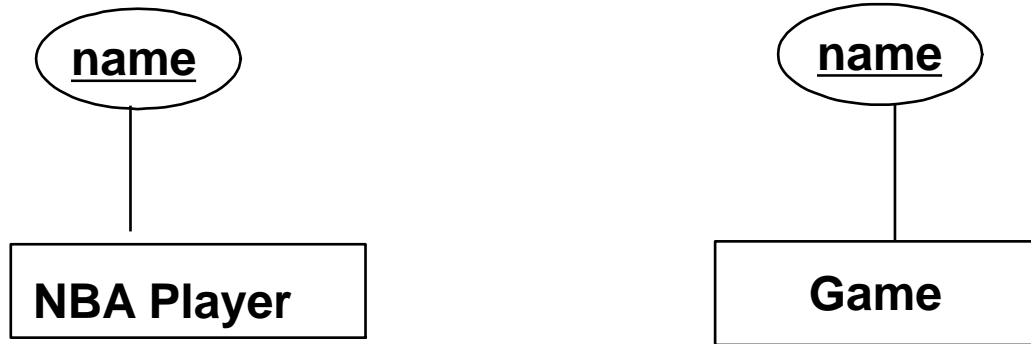- **Relationship set does NOT have keys**

24
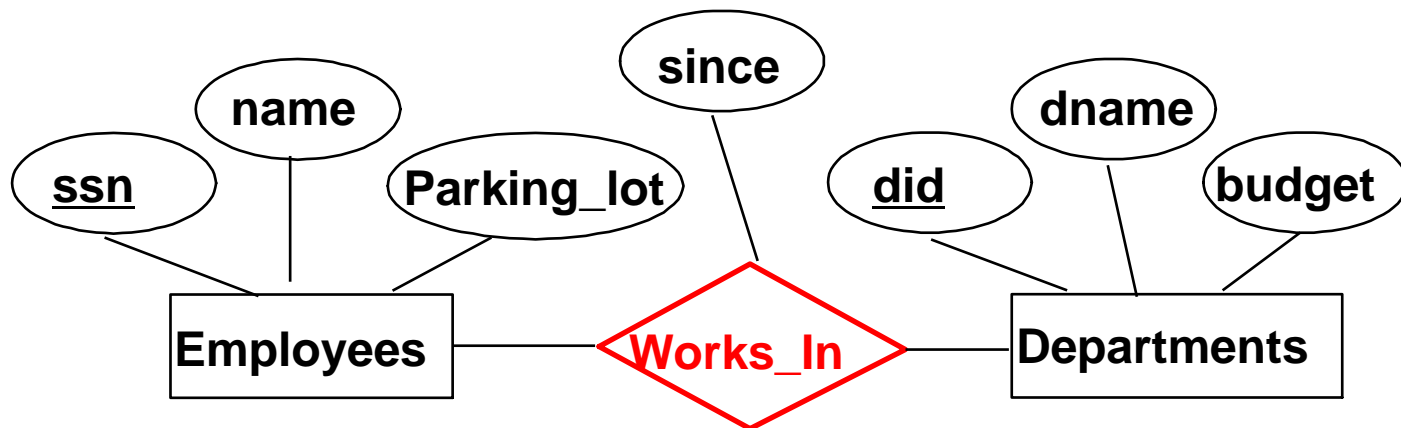
Degree of A Relationship

- **Degree : the number of participating entity sets.**
  - Degree 1: unary (1 diamond connecting with 1 rectangle)
  - Degree 2: binary (1 diamond connecting with 2 rectangles)
  - Degree 3: ternary (1 diamond connecting with 3 rectangles)
  - Degree n: n-ary (1 diamond connecting with n rectangles)
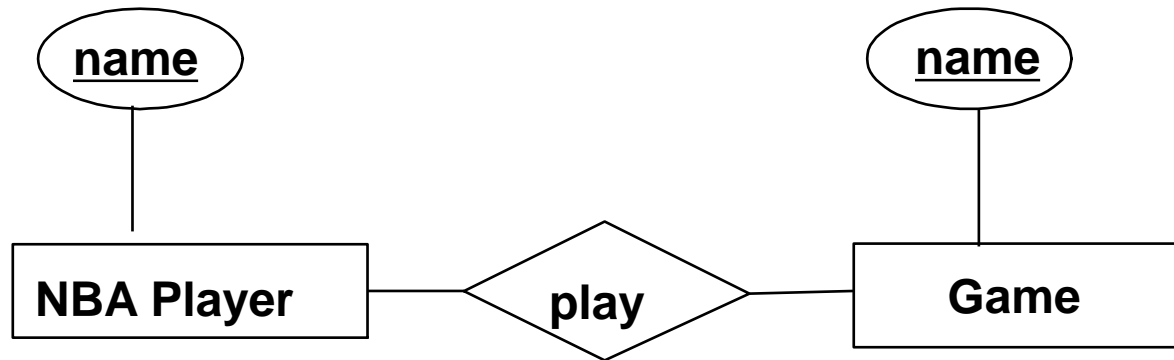- **Binary relationships are very common and widely used.**

Exercise: Relationship sets

```
  ( name )                    ( name )
     |                           |
┌─────────────┐           ┌─────────────┐
│ NBA Player  │           │    Game     │
└─────────────┘           └─────────────┘
```
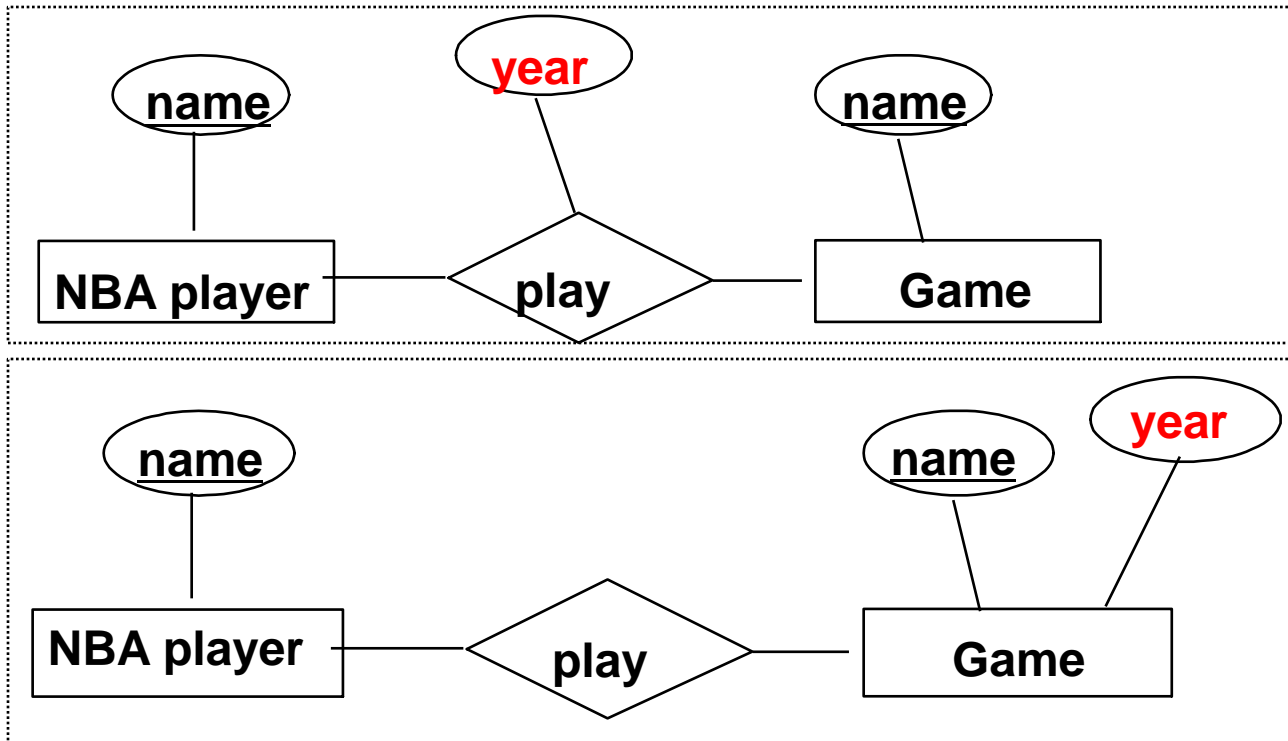
- Given two entity sets *NBA Player* and *Game*, add the relationship set between them, which describes the fact of *which NBA players play in which games*.

Add Attributes to ER Diagram

- How to represent the "year" info for each game that the players played?
  - E.g., "Stephen Curry plays NBA Finals for 2019 and 2022".
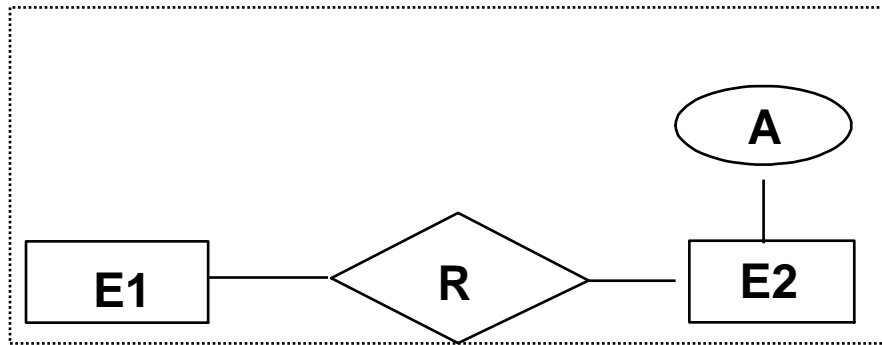
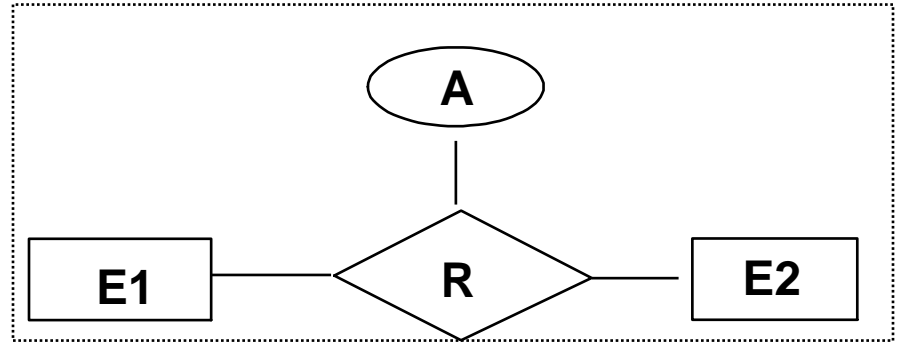# Two Possible Solutions

Schema A

Schema B

- Which one is correct? Should the attribute *year* be added to the relationship set *play* (Schema A) or the entity set *Game* (Schema B)?
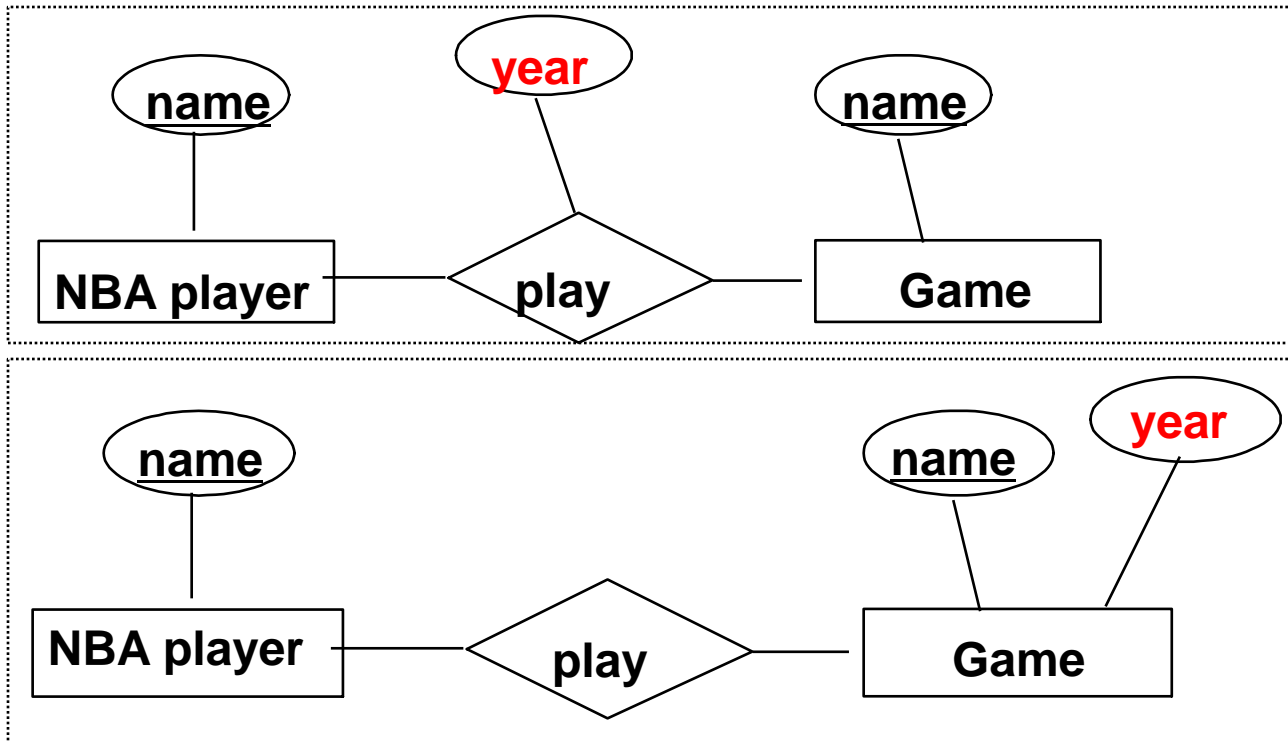
Where to Put Attributes?

Schema 1
Attribute associate with Entity set

Schema 2
Attribute associated with Relationship set

- What's the difference between these 2 ER diagrams?
  - Schema 1: each E2 entity is associated with a single value of A
  - Schema 2: each (E1, E2) pair is associated with a single value of A
    - Each E2 entity can be associated with multiple A values, one for each (E1, E2) pair
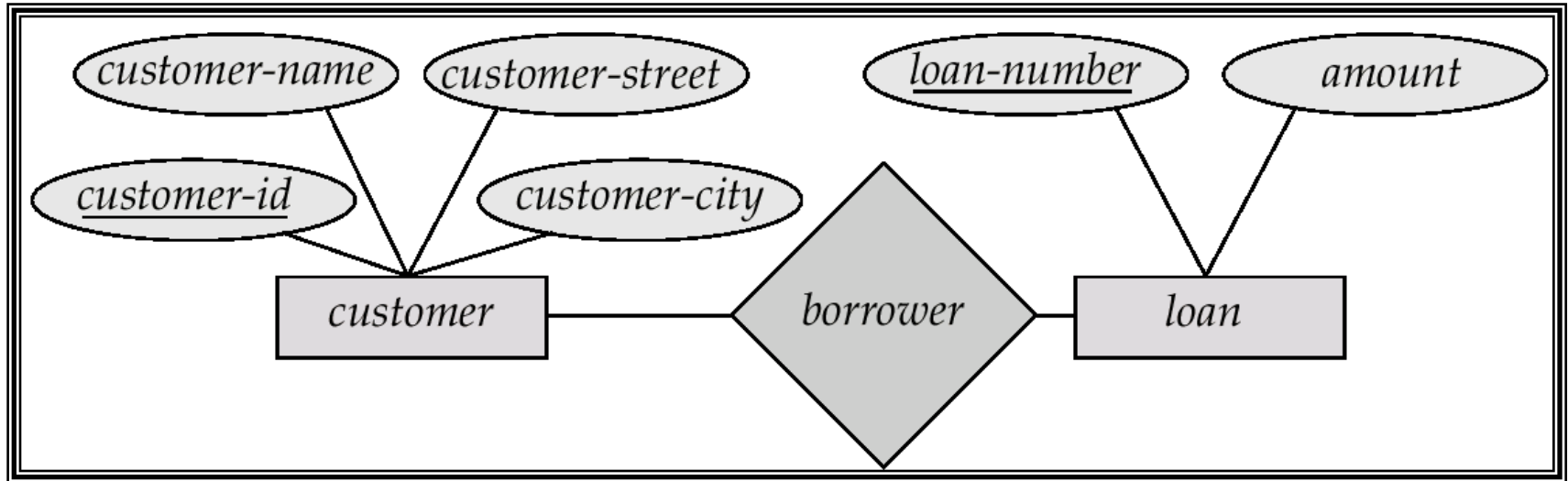
Attribute Exercise (Continue)

Schema A

name — NBA player — play — year — name — Game

Schema B

name — NBA player — play — name — Game — year

- Does Schema A allow the same player to play the same game for different years?
  - E.g., "Stephen Curry plays NBA Finals for 2019 and 2022"
- What about Schema B?

Summary: E-R Diagrams

- □ **Rectangles** represent entity sets.
- □ **Diamonds** represent relationship sets.
- □ **Lines** link attributes to entity sets and entity sets to relationship sets.
- □ **Ellipses** represent attributes of entity/relationship sets
- □ **Underline** indicates primary key attributes