



# Recap

- Check syllabus
- Sign on Piazza  
([piazza.com/iastate/spring2020/coms363](https://piazza.com/iastate/spring2020/coms363))
- Course improvement project: call for participants

# Database History

- What do you visualize when you hear the term “database”?
- Spreadsheet like Excel table (1982), google sheet (2006)?
- Several spreadsheets?

# Why not spreadsheet?

- You have a registration sheet
- Your instructors also have registration sheets
- You change yours
- Your instructors should be “informed”
- Management is a problem

# Why not spreadsheet?

- Size matters
  - Out of memory
- Security features
  - You can't see other students' records

# Data Model History

- Navigational (60's)
  - linked list of free-form records
  - Hard to implement, but if you can implement it well, it has high performance
- Relational (80's, proposed in 1970)
  - Remain dominant
  - "table" of fixed-length records, with each table used for a different type of entity
- Post-relational
  - Object-oriented (90's)
  - NoSQL (late 2000)

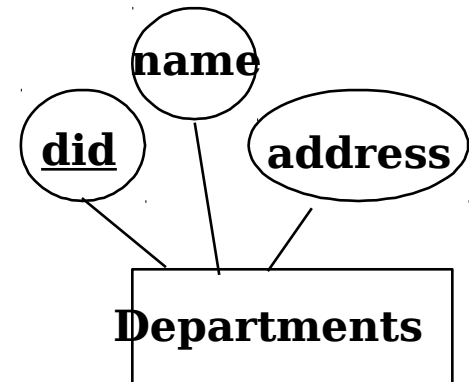
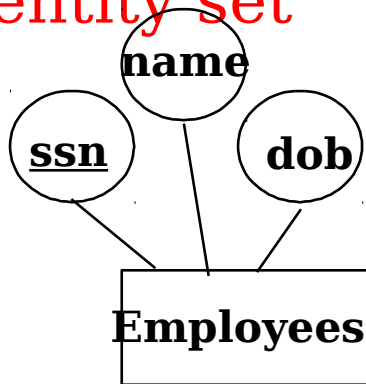
# Database Design

- Before we can design a database to represent a mini-world, we need to understand what this mini-world is about
- If we can get these two questions answered, we have a rather good understanding
  1. What **entities** does this mini-world have?
  2. What **relationships** exist among these entities?

**Entity-Relationship (ER) model** is a high level conceptual data model

# Basic terms and notations

- An **entity** is a uniquely identifiable object that exists on its own (e.g., an employee, a department)
- Entities have **attributes**
- An attribute is a **key attribute** if its values are distinct for each individual entity.
- Similar entities (having same attributes) form an **entity set**

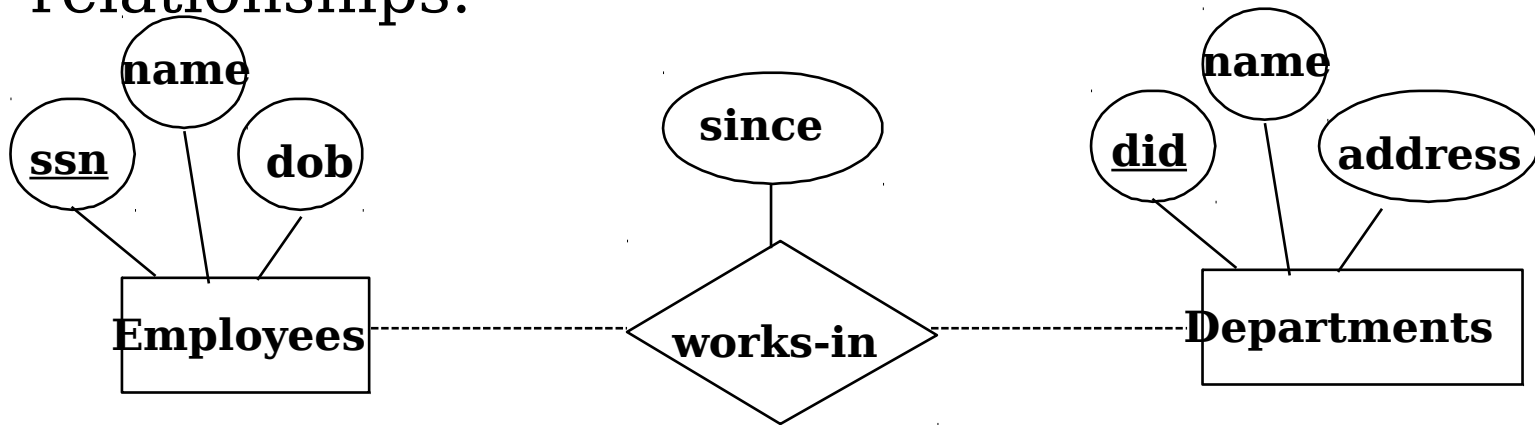


## Notation

- an entity set : a rectangle
- an attribute: a circle and a line connected to a rectangle
- a key attribute: underlined

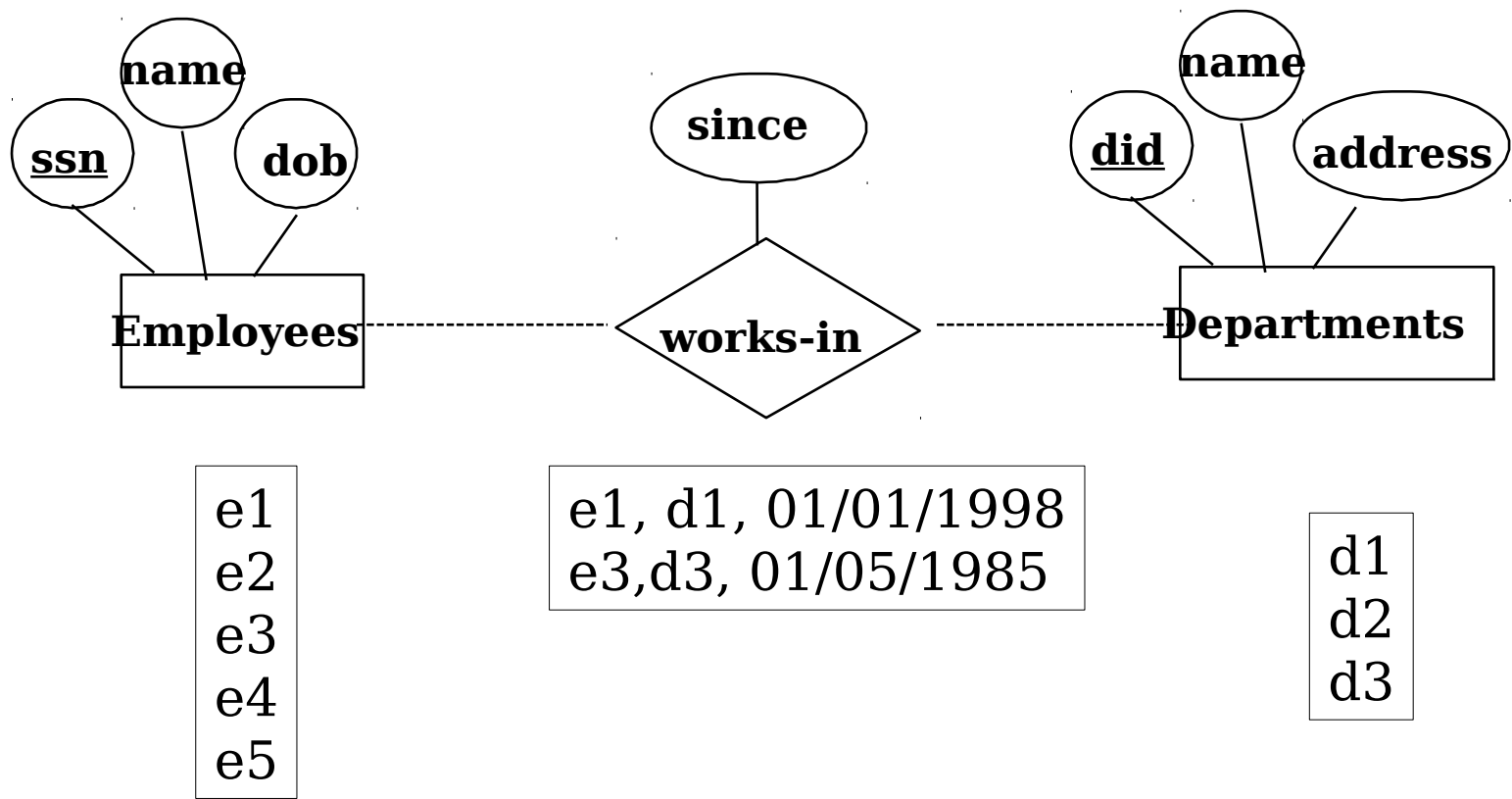


- **Relationship:** Association among two or more entities.
  - Example: Attishoo works in the Pharmacy department.
- **Relationship Set:** Collection of similar relationships.

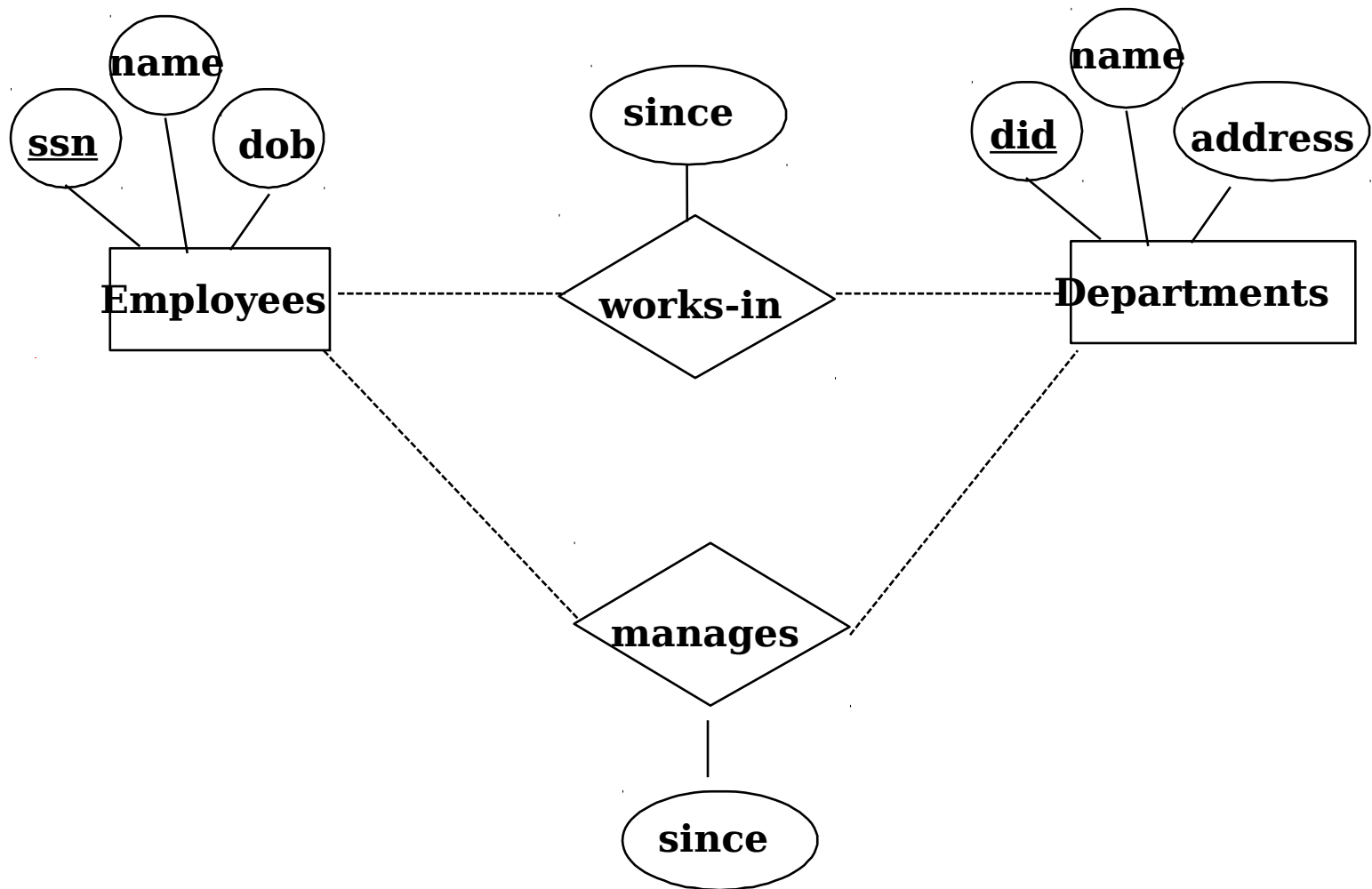


## Notation

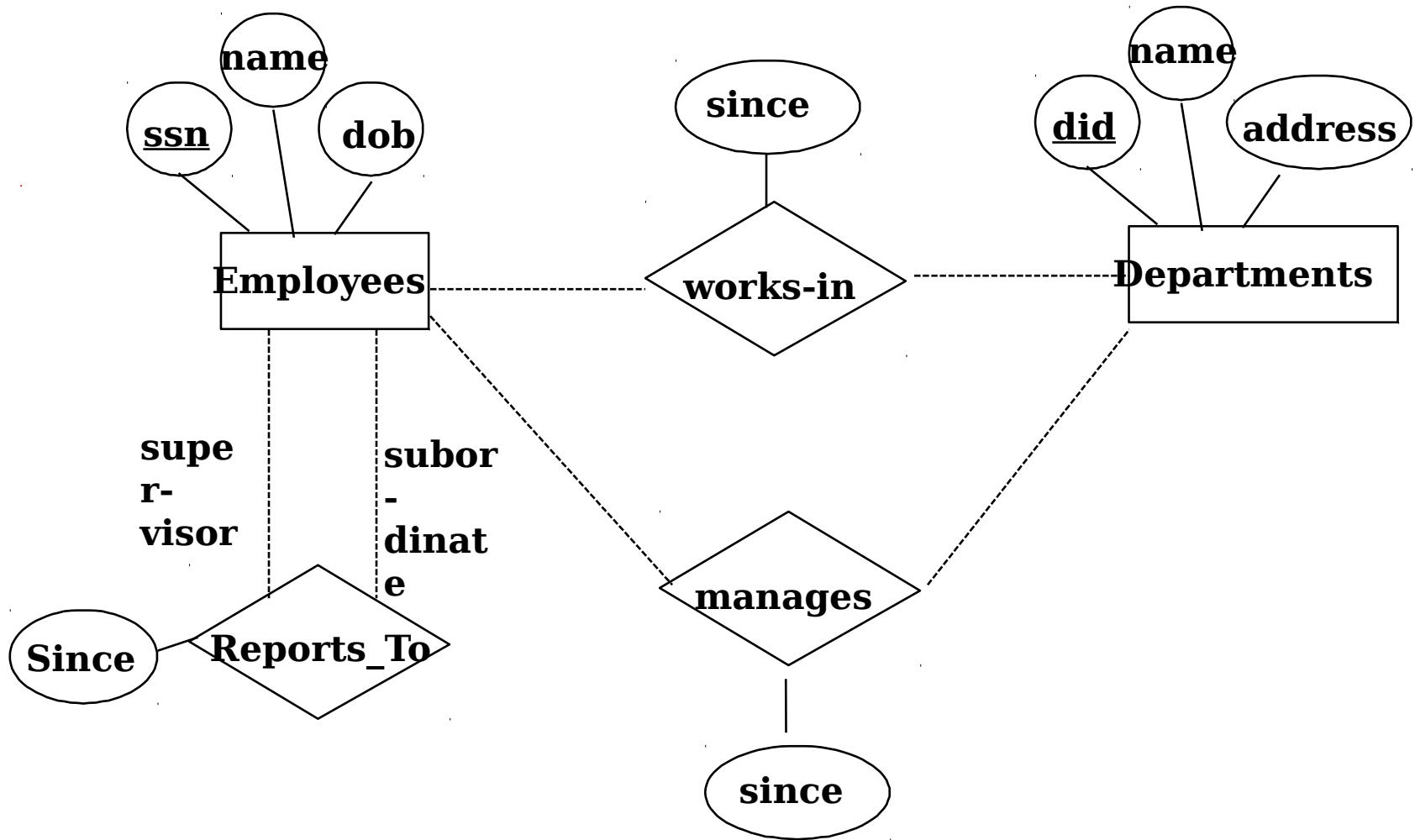
- a relationship set is represented by a diamond
  - Every participating entity is connected to it by a dotted edge
  - Every attribute is also connected to it by an edge
- an attribute: a circle and a line connected to a rectangle



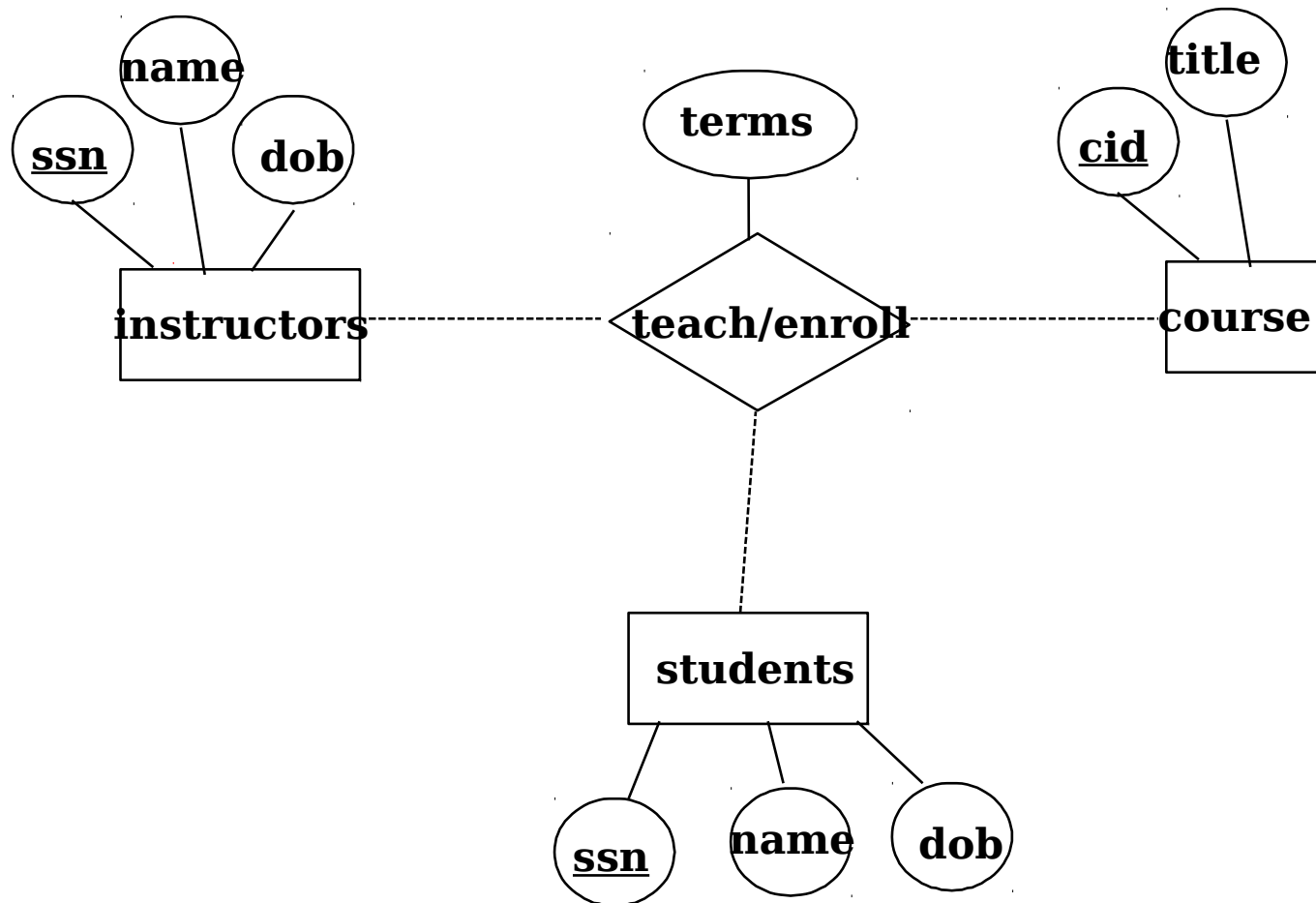
two entity sets may have different relationships



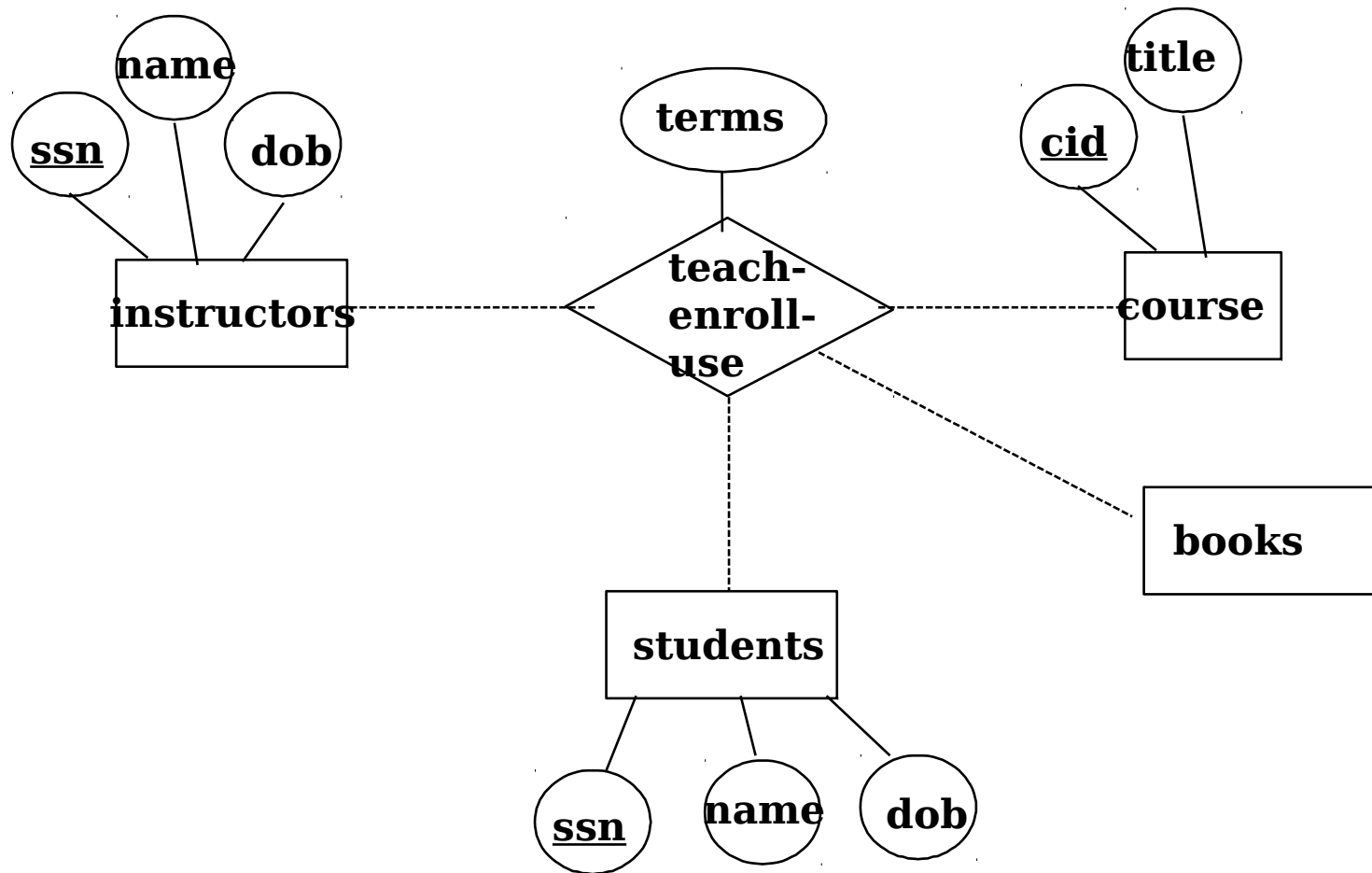
Entities in a same set can participate in some relationship



more than two entity sets may participate in a same relationship set



more than two entity sets may participate in a same relationship set



So far, only two notations E and R, but they are so powerful in describing a “world”. When designing a database, think about these two questions

- 1) What entities it has
- 2) What relationships exist among these entities

Example: A database for university

- 1) E: faculty, students, courses, buildings, vehicles, research projects, departments, parking lots ...
- 2) R: A faculty teaches some students in buildings...

- The two questions are fundamental, but we often need more information and therefore need to ask more questions
- Question 3: Is there any participation constraint on relationships?
  - **Uni-participation**: An entity can participate in a relationship at most once
  - **Total-participation**: Every entity in an entity set must participate in a relationship
  - **A combination of both**: Every entity in an entity set must participate in a relationship but only once

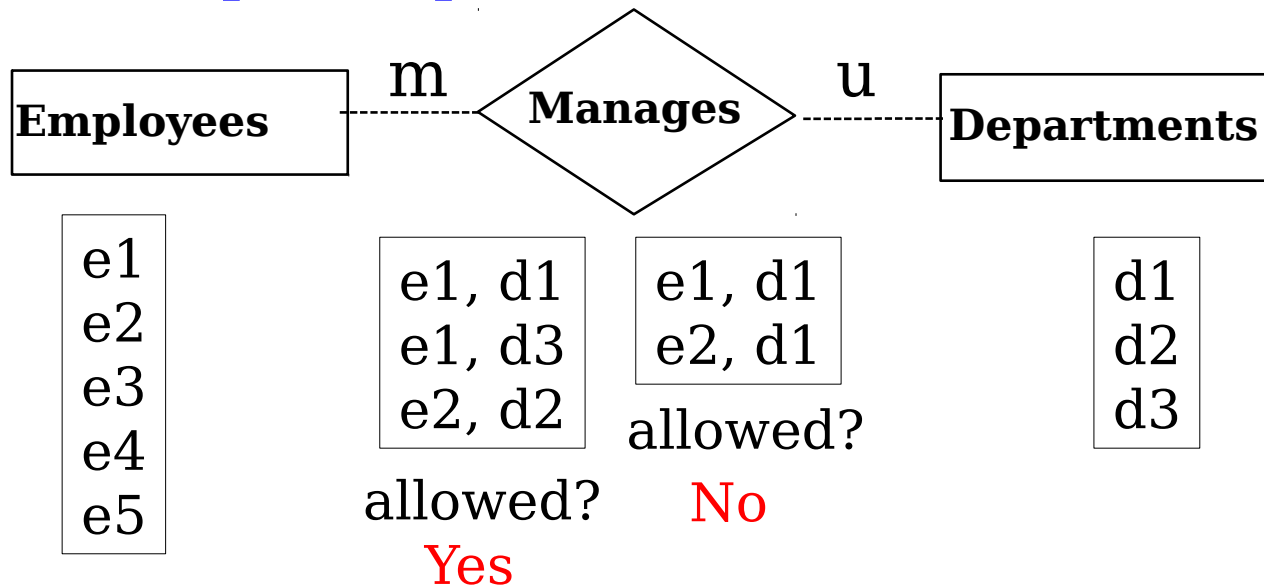


## Uni-participation (also called **key constraint**)

- Each entity in an entity set can participate in at most one relationship in a relationship set
- Notation: u
- If there is no such constraint, then it is called **multi-participation**, denoted as m

# Uni-participation (also called **key constraint**)

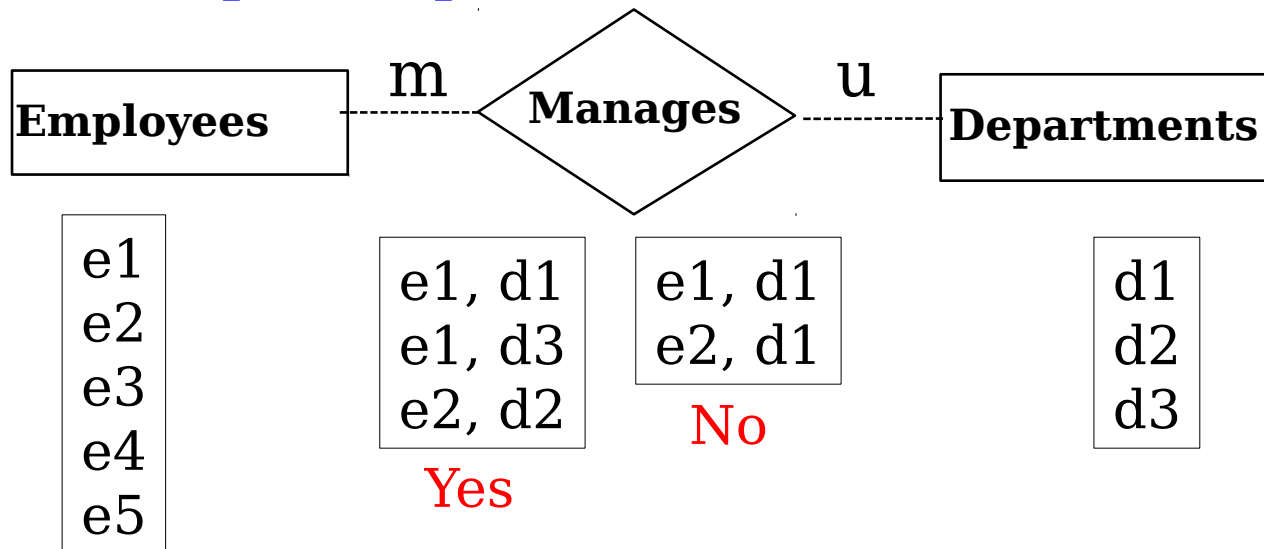
- Each entity in an entity set can participate in at most one relationship in a relationship set
- Notation: u
- If there is no such constraint, then it is called **multi-participation**, denoted as m



What does this mean?

# Uni-participation (also called **key constraint**)

- Each entity in an entity set can participate in at most one relationship in a relationship set
- Notation: u
- If there is no such constraint, then it is called **multi-participation**, denoted as m



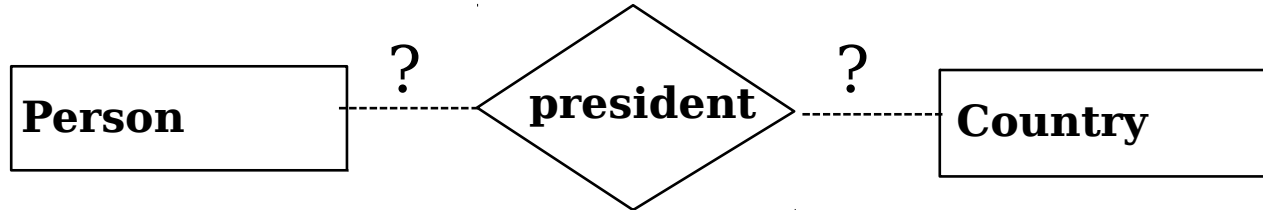
A department can participate in relationship “manages” at most one time

□ Every department can have at most one manager (but does not have to)

□ Every Employee can manage several departments (but does not have to)

# Can you think of another example?

- How about



- Each person can be president in at most one country?
- Each country can have at most one president (currently)?
- Each country can have several presidents (over history)?

# Total participation

- Every entity in an entity set must participate in a relation set
- Notation: solid line
- If there is no such constraint, then it is called **partial-participation**, denoted as dotted line



e1  
e2  
e3  
e4  
e5

e1, d1  
e1, d3  
allowed?  
**No**

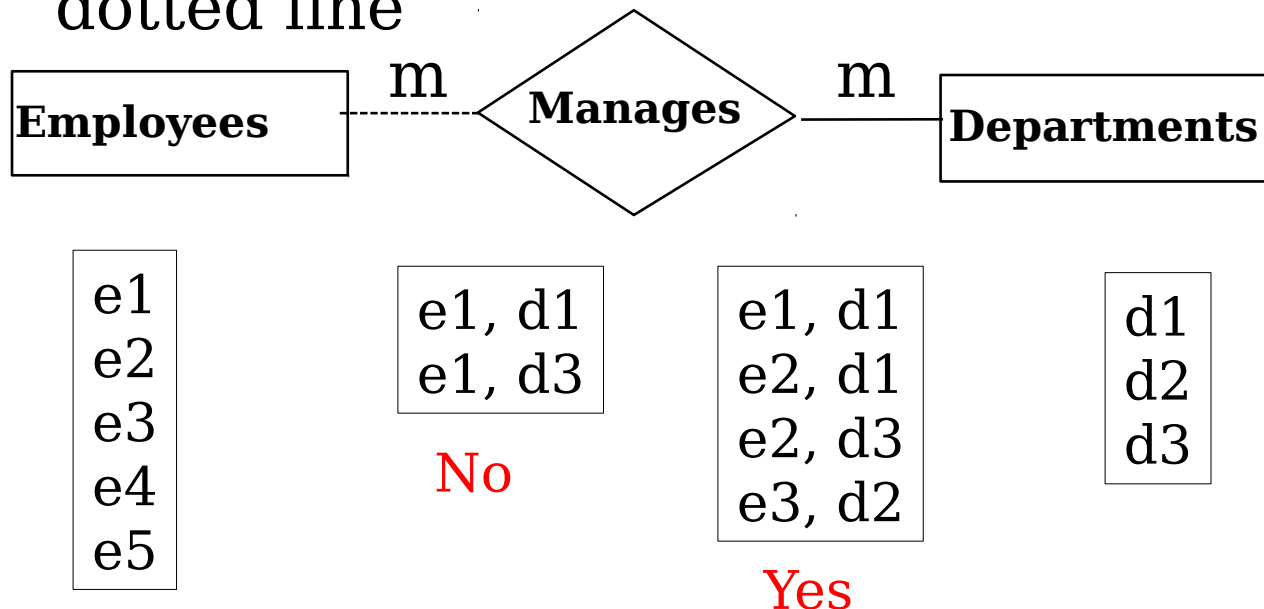
e1, d1  
e2, d1  
e2, d3  
e3, d2  
allowed?  
**Yes**

d1  
d2  
d3

What does this mean?

# Total participation

- Every entity in an entity set must participate in a relation set
- Notation: solid line
- If there is no such constraint, then it is called **partial-participation**, denoted as dotted line

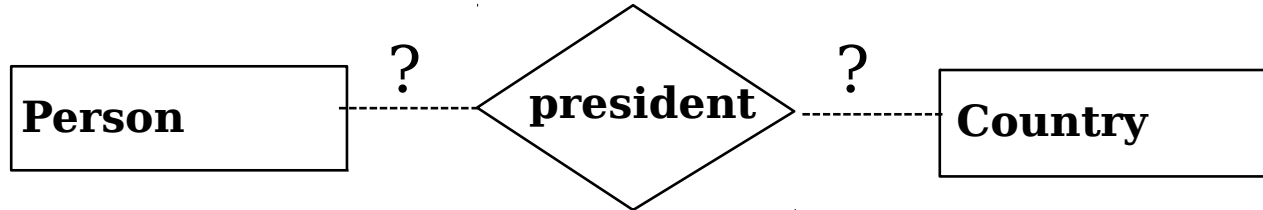


Every department must participate in relationship “manages”

□ Every department must have at least one manager

# Can you think of another example?

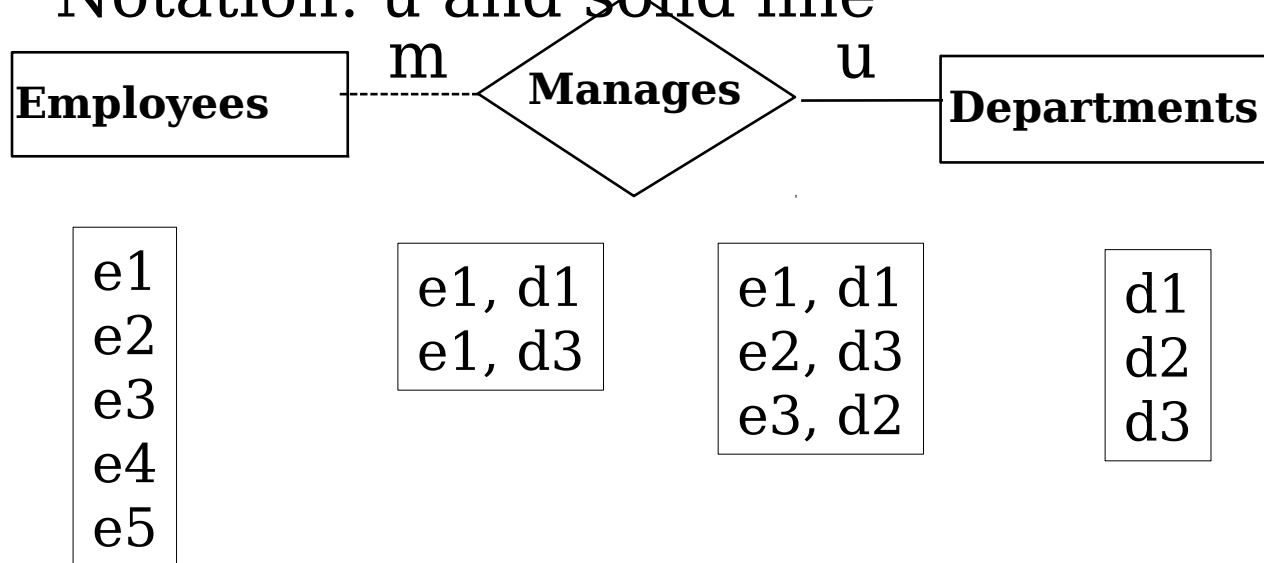
- How about



- Each person can be president in at most one country?
- Each country can have at most one president (currently)?
- Each country can have several presidents (over history)?

# Uni-participation and total participation

- Every entity in an entity set must participate in a relation set and participate only once
- Notation: u and solid line



which one allowed?

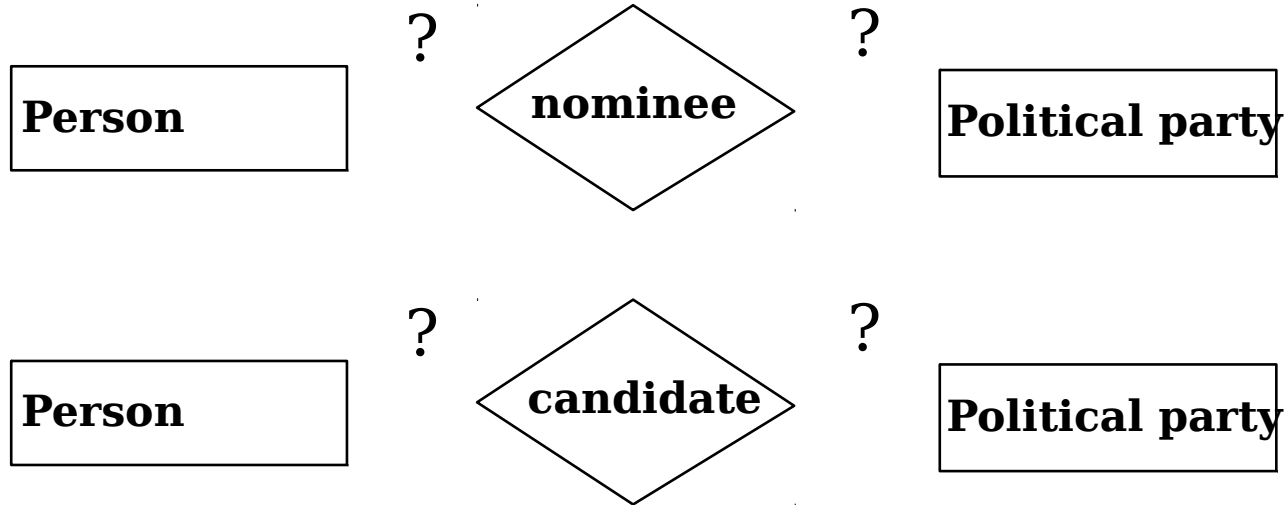
Every department must participate in relationship "manages" and only once

□ Every department must have one and only one manager



# Can you think of another example?

- How about

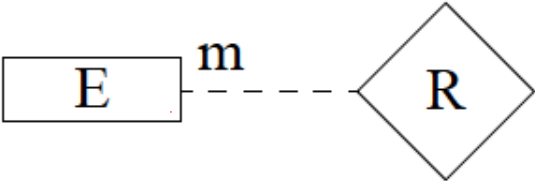
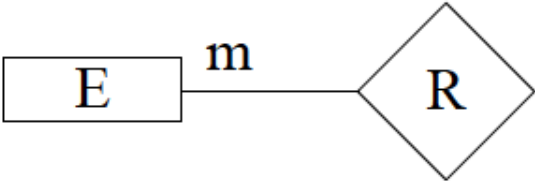
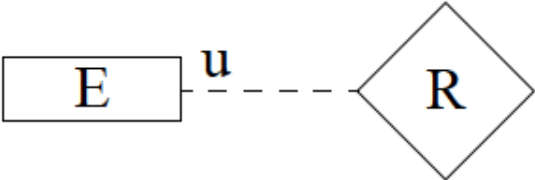
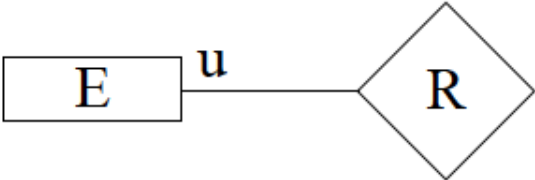


- Suppose only “Republican” and “Democratic” for “Political party”

Partial vs. total participation: shown by dotted vs. solid edges

Multi vs. uni participation: shown by labels “m” vs. “u”

“m” for multi or many; “u” for uni or unique

	Partial (Dotted)	Total (Solid)
Multi (“m”)	 <p>Key(E) cannot be key of R Some E may not occur in R</p>	 <p>Key(E) cannot be key of R Every E occurs in R</p>
Uni (“u”)	 <p>Key(E) is a key of R Some E may not occur in R</p>	 <p>Key(E) is a key of R Every E occurs in R</p>

## More examples



# A quick review

- What does E have?
- What does R have?
- Any constraints on R?
  - Uni-participation vs. multi-participation
  - Total-participation vs. partial-participation

# A quick review

- Q1: What does E have?
- Q2: What does R have?
- Q3: Any constraints on R?
  - Uni-participation vs. multi-participation
  - Total-participation vs. partial-participation

One more question, a special kind of relationship :

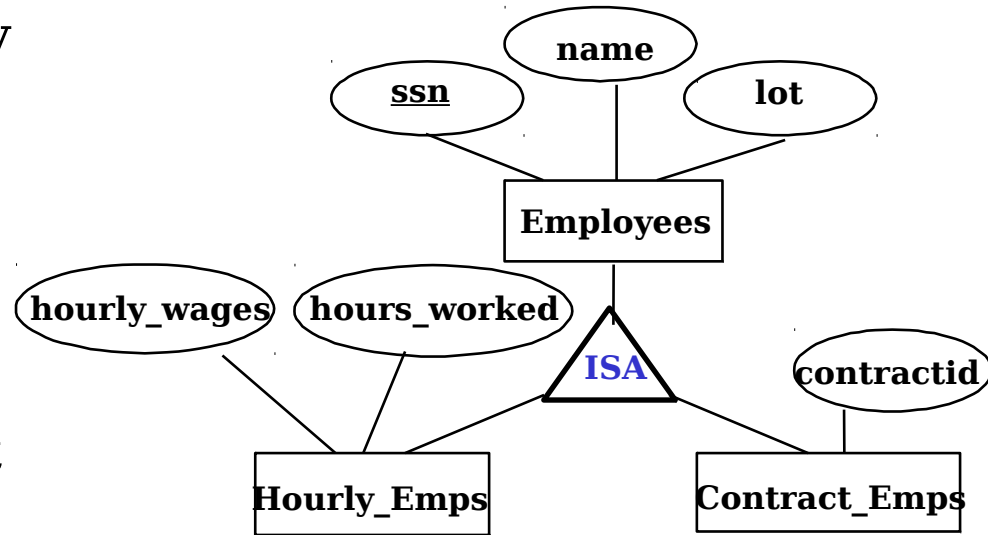
Q4: Is there any relationship that is ISA?

# ISA (`is a')

## Hierarchies

- If we declare A **ISA** B, every A entity is also considered to be a B entity.

- Reasons for using ISA:
  - To add attributes specific to a subclass.
  - To identify entities that participate in a relationship.



- **Overlap constraints:**

- Can Joe be an Hourly\_Emps as well as a Contract\_Emps entity? (Allowed/disallowed)
  - Default value: no overlap;
  - Otherwise, write Hourly\_Emps OVERLAPS Contract\_emps

- **Covering constraints:**

- Does every Employees entity also have to be an Hourly\_Emps or a Contract\_Emps entity? (Yes/no)
  - Default value: no;
  - Otherwise write Hourly\_Emps and Contract\_Emps COVER Employees

- ER design is *subjective*. There are often many ways to model a given scenario!
- Analyzing alternatives can be tricky, especially for a large enterprise. Common choices include:
  - Entity vs. attribute, entity vs. relationship, binary or n-ary relationship, whether or not to use ISA hierarchies, and whether or not to use aggregation.
- ER diagrams can use different notations, but use the same concept.
- Several software tools are available for creating ER diagrams: IBM Rational Rose, Microsoft Visio

# Key Concepts of ER Model

- Entity set
- Relationship set
- Constraints
  - uni-constraint
  - Total participation constraint
- A special kind of relationship: ISA

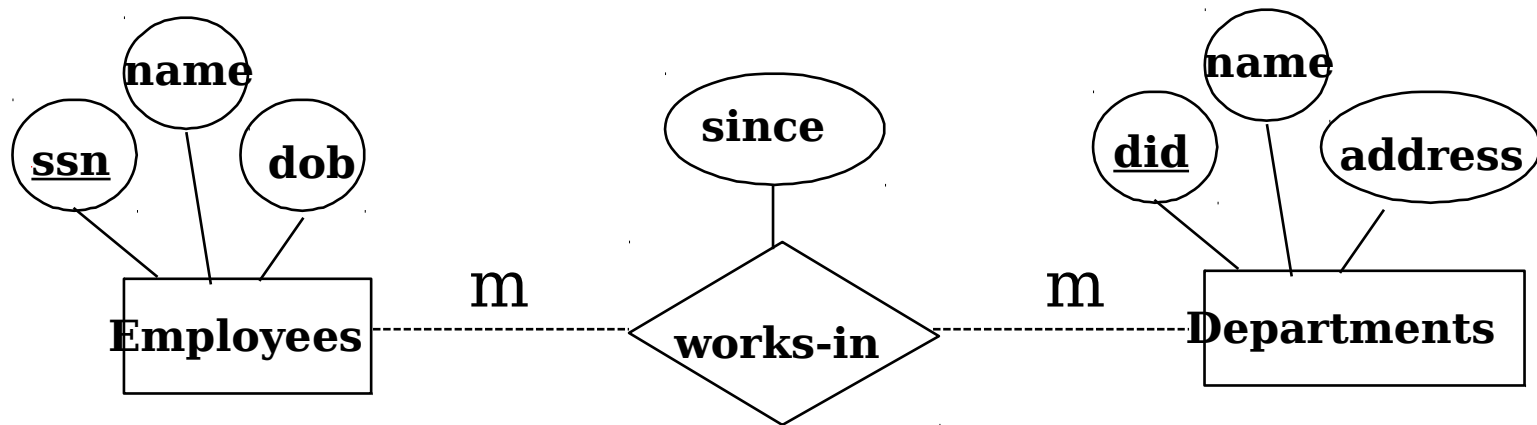
*Question to think*

*Is there anything that cannot be described by ER model?*

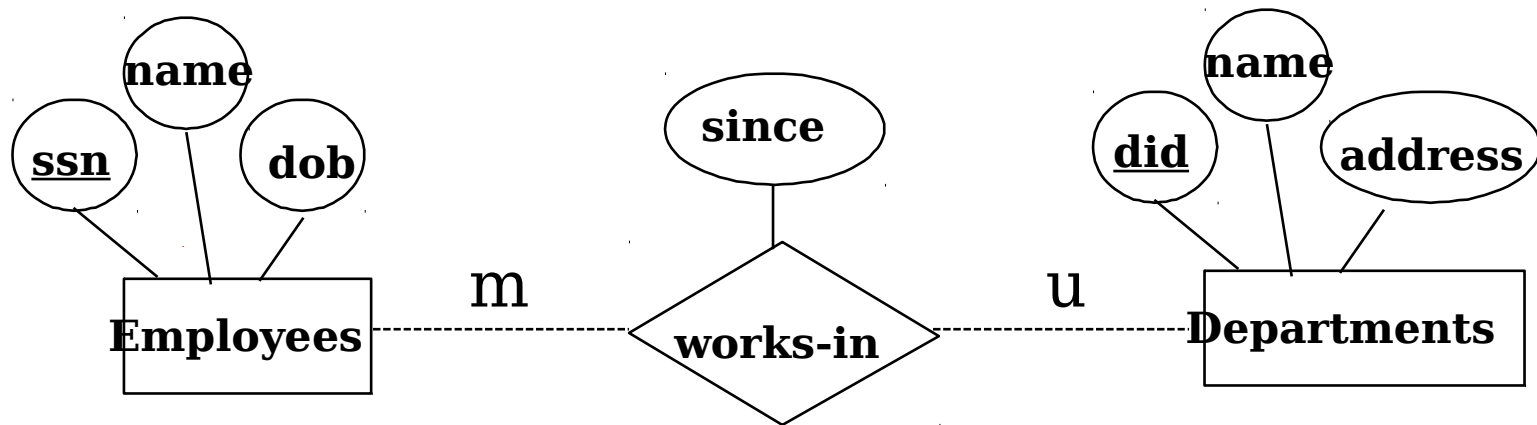


# E-R Model Exercises

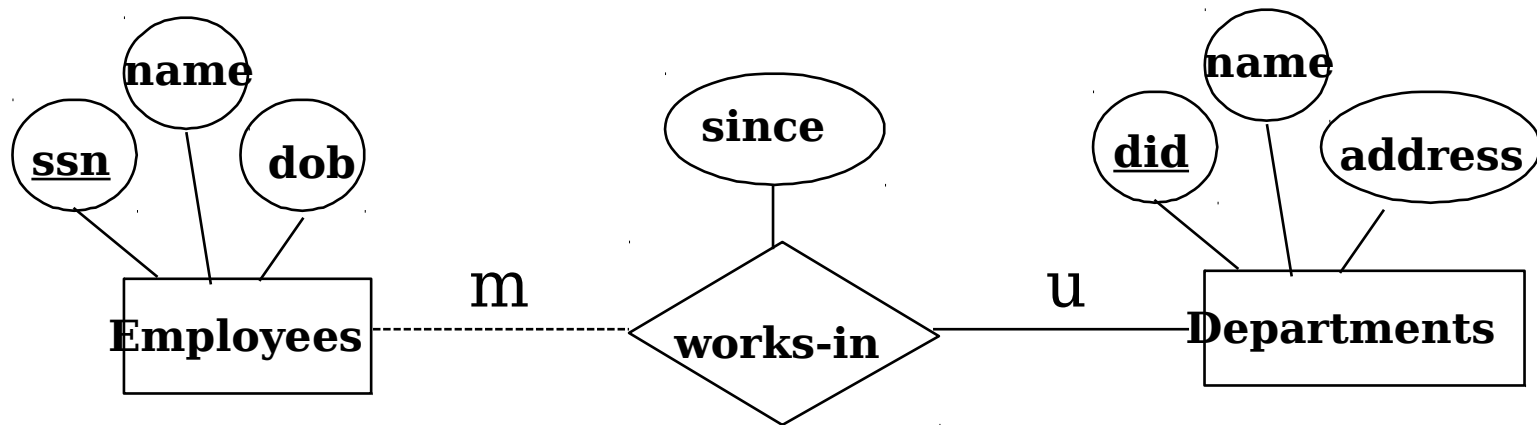
- Read an ER diagram
- Draw an ER diagram



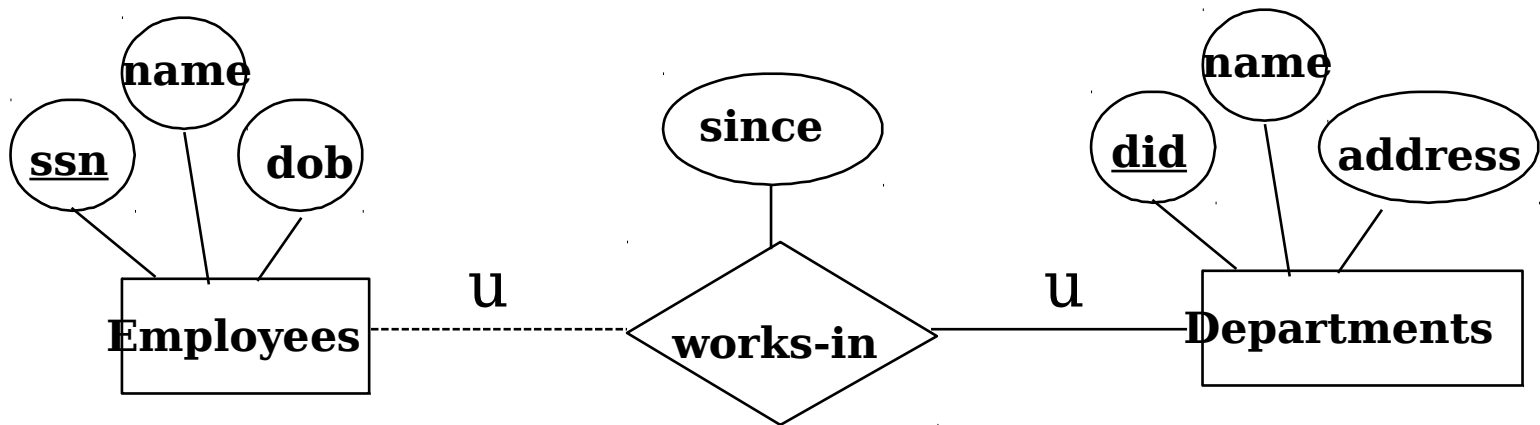
1. Two entity sets:
  - Employees and Departments
2. One relationship set
  - “works-in” between Employees and Departments
3. Constraint
  - An employee can participate in “works-in” one or more times, or does not participate at all
    - multi- and partial participation
  - A department can participate in “works-in” one or more times, or does not participate at all
    - multi- and partial participation



1. Two entity sets:
  - Employees and Departments
2. One relationship set
  - “works-in” between Employees and Departments
3. Constraint
  - An employee can participate in “works-in” one or more times, or does not participate at all
    - multi- and partial participation
  - A department can participate in “works-in” at most once
    - uni- and partial participation
    - Semantically: a department can have at most one employee



1. Two entity sets:
  - Employees and Departments
2. One relationship set
  - “works-in” between Employees and Departments
3. Constraint
  - An employee can participate in “works-in” one or more times, or does not participate at all
    - multi- and partial participation
  - Every department must participate in “works-in” once and only once
    - uni- and total participation
    - Semantically: a department can have at most one employee and only one employee



1. Two entity sets:
  - Employees and Departments
2. One relationship set
  - “works-in” between Employees and Departments
3. Constraint
  - An employee can participate in “works-in” at most one time
    - uni- and partial participation
    - Semantically: an employee can work in at most one department
  - Every department must participate in “works-in” once and only once
    - uni- and total participation
    - Semantically: a department can have at most one employee and only one employee

## Full answer

### 1. Two entity sets:

- Employees and Departments
- Employees have attributes ssn, name and dob. SSN is the key attribute
- Departments have attributes did, name and address. DID is the key attribute

### 2. One relationship set

- “works-in” between Employees and Departments
- “works-in” has attribute “since”.

### 3. Constraint

- An employee can participate in “works-in” at most one time
  - uni- and partial participation
  - Semantically: an employee can work in at most one department
- Every department must participate in “works-in” once and only once
  - uni- and total participation
  - Semantically: a department can have at most one employee and only one employee

# Drawing ER diagram

- What entity sets?
  - attributes, key attribute
- What relationship sets?
  - attributes
  - ISA?
- What constraints?
  - uni- vs. multi-participation
  - total vs. partial participation

# Exercise 1

- A company has a number of employees. The attributes of EMPLOYEE include Employee\_ID (identifier), Name, Address, and Birthdate.
- The company also has several projects. Attributes of PROJECT include Project\_ID (identifier), Project\_Name, and Start\_Date.
- Each Employee may be assigned to one or more projects, or may not be assigned to a project.
- A project must have at least one employee assigned, and may have any number of employees assigned.
- An employee's billing rate may vary by project, and the company wishes to record the applicable billing rate (Billing\_Rate) for each employee when assigned to a particular project.



# Exercise 1

- A company has a number of employees. The attributes of EMPLOYEE include Employee\_ID (identifier), Name, Address, and Birthdate.

# Exercise 1

- The company also has several projects. Attributes of PROJECT include Project\_ID (identifier), Project\_Name, and Start\_Date.

# Exercise 1

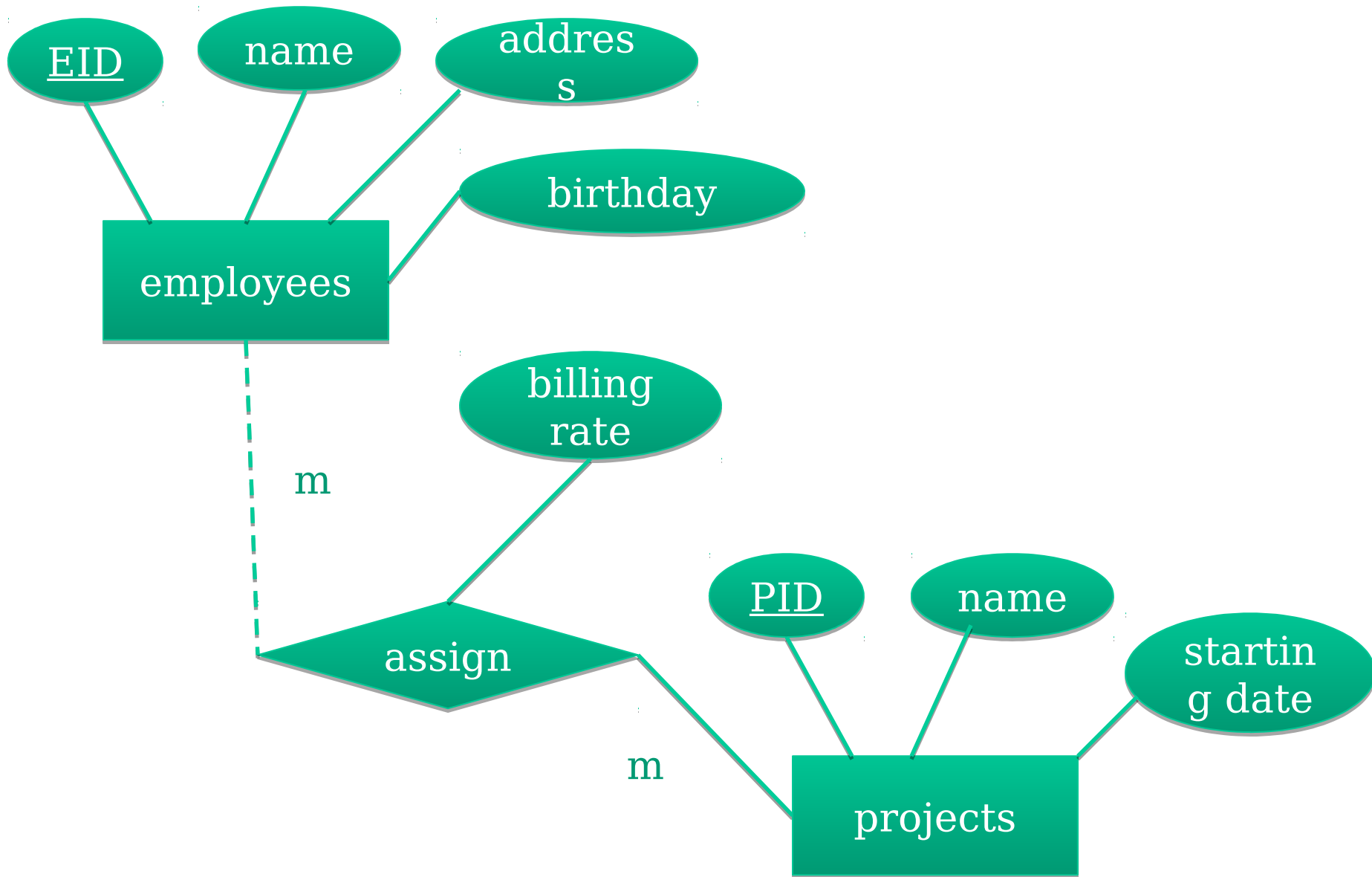
- Each Employee may be assigned to one or more projects, or may not be assigned to a project.

# Exercise 1

- A project must have at least one employee assigned, and may have any number of employees assigned.

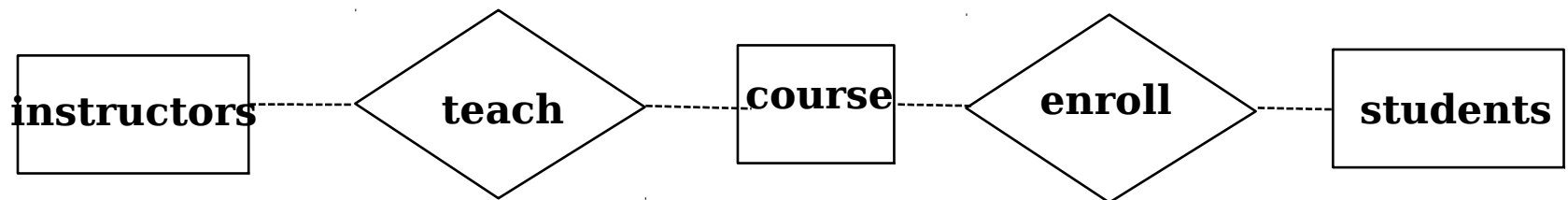
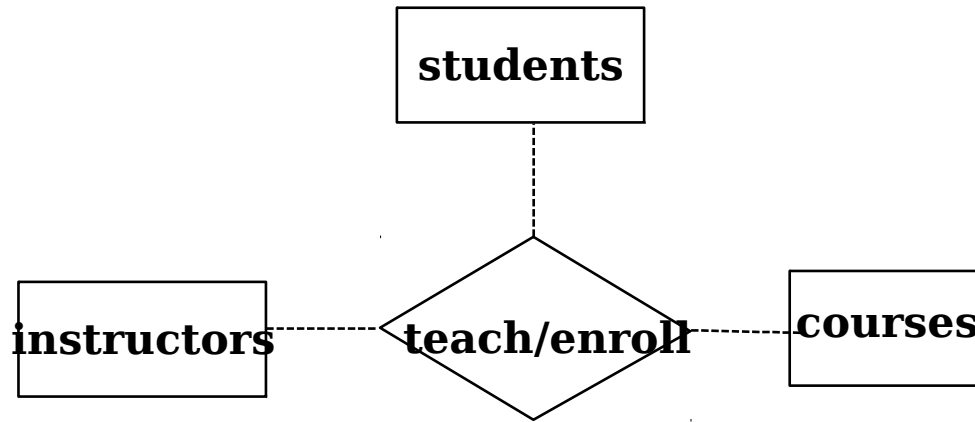
# Exercise 1

- An employee's billing rate may vary by project, and the company wishes to record the applicable billing rate (Billing\_Rate) for each employee when assigned to a particular project.

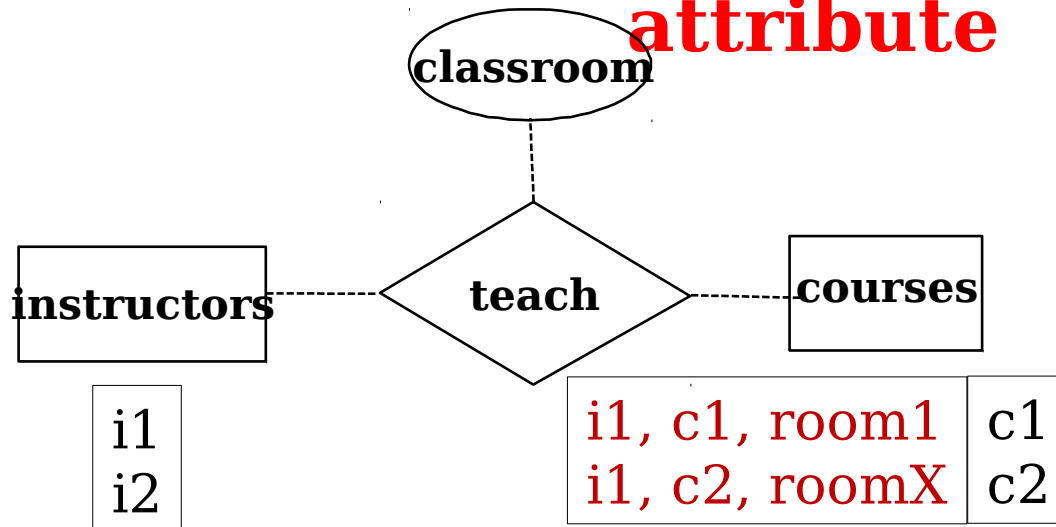


# Problem 1

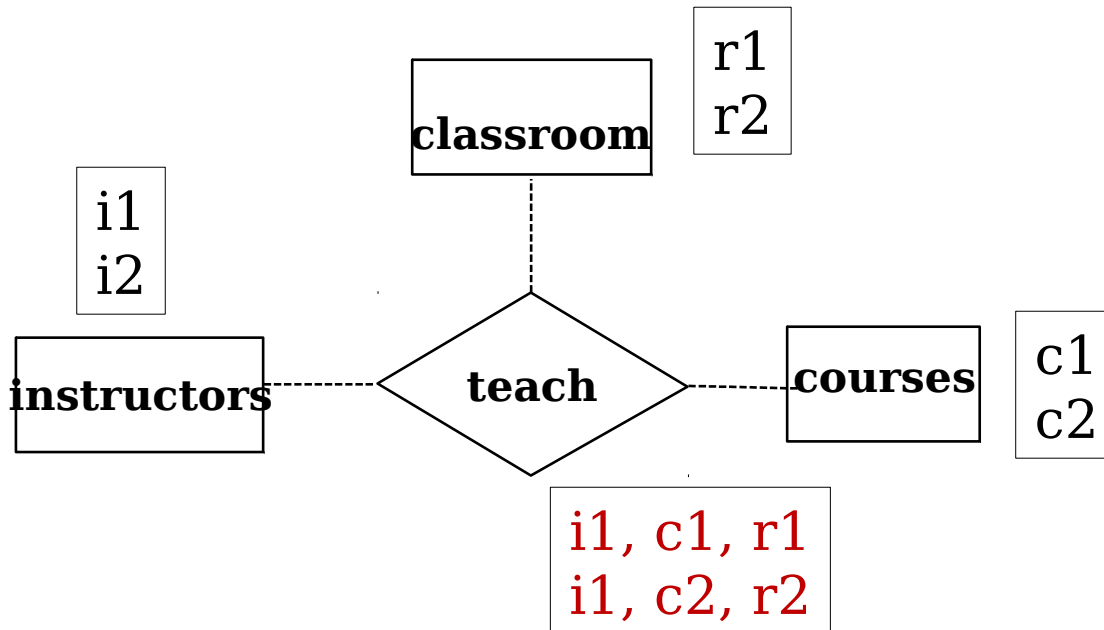
**Multiple entity sets  
participate  
in a same relationship  
set**



## Problem 2 Entity set or attribute



Here classroom is an attribute for relationship teach. It can have any value.

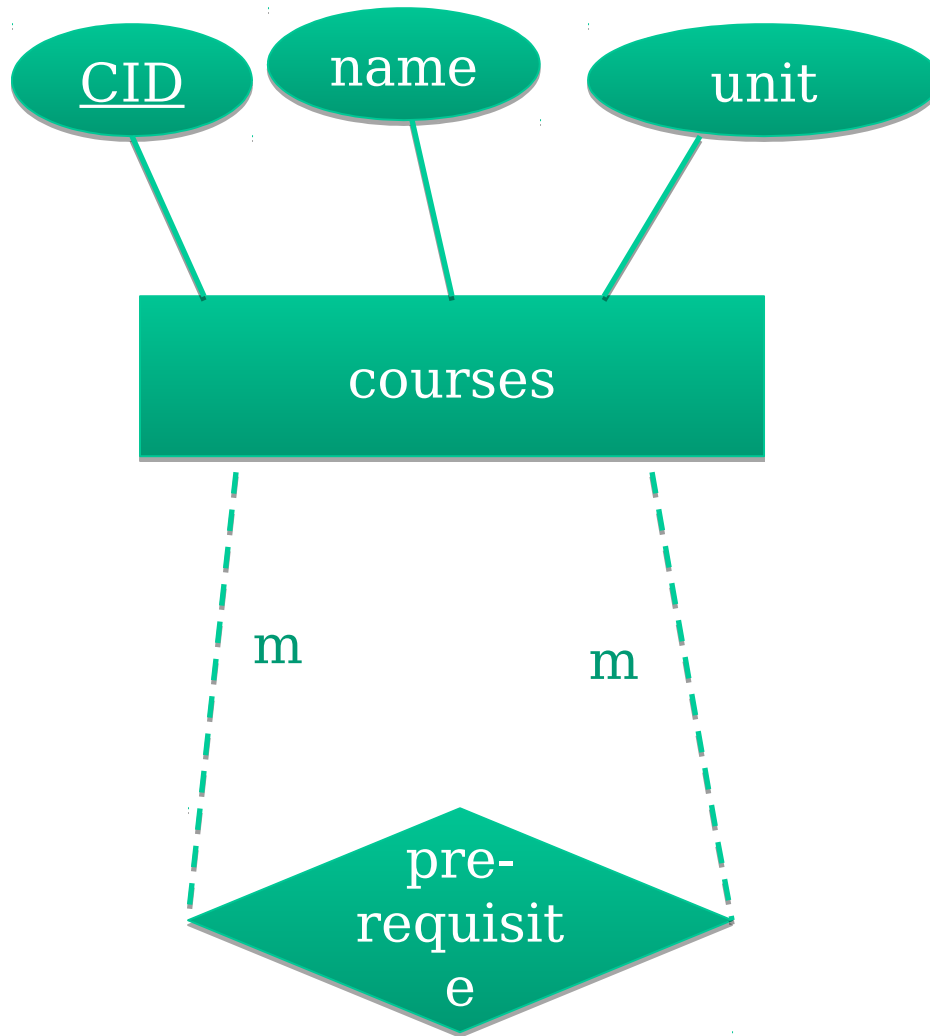


Here classroom is an entity set, which a number of members. Only these members can participate in relationship teach.



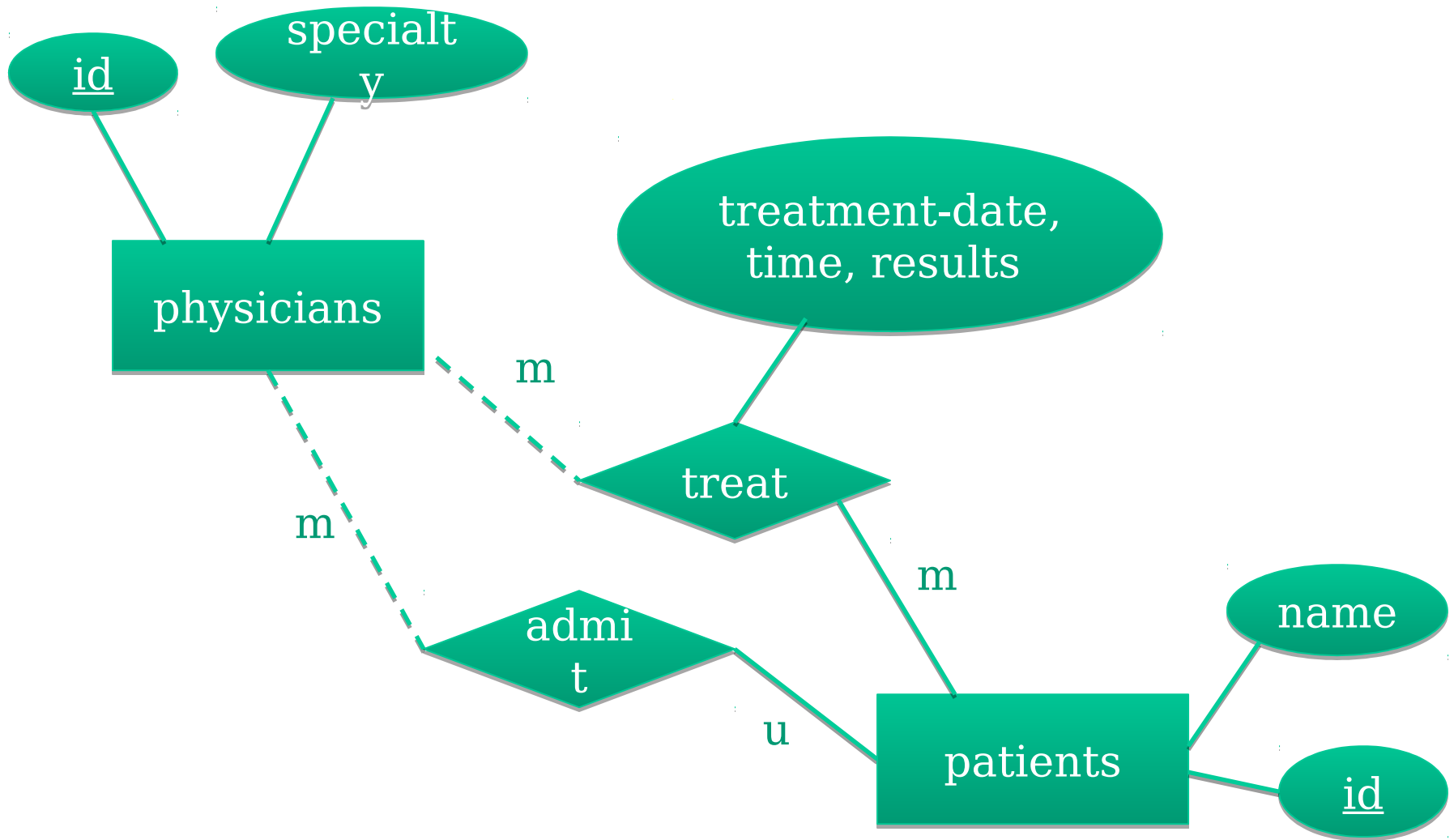
## Exercise 2

- A university has a large number of courses in its catalog. Attributes of COURSE include Course\_number (identifier), Course\_name, and Units.
- Each course may have one or more different courses as prerequisites, or may have no prerequisites. Similarly, a particular course may be a prerequisite for any number of courses, or may not be prerequisite for any other course.



# Exercise 3

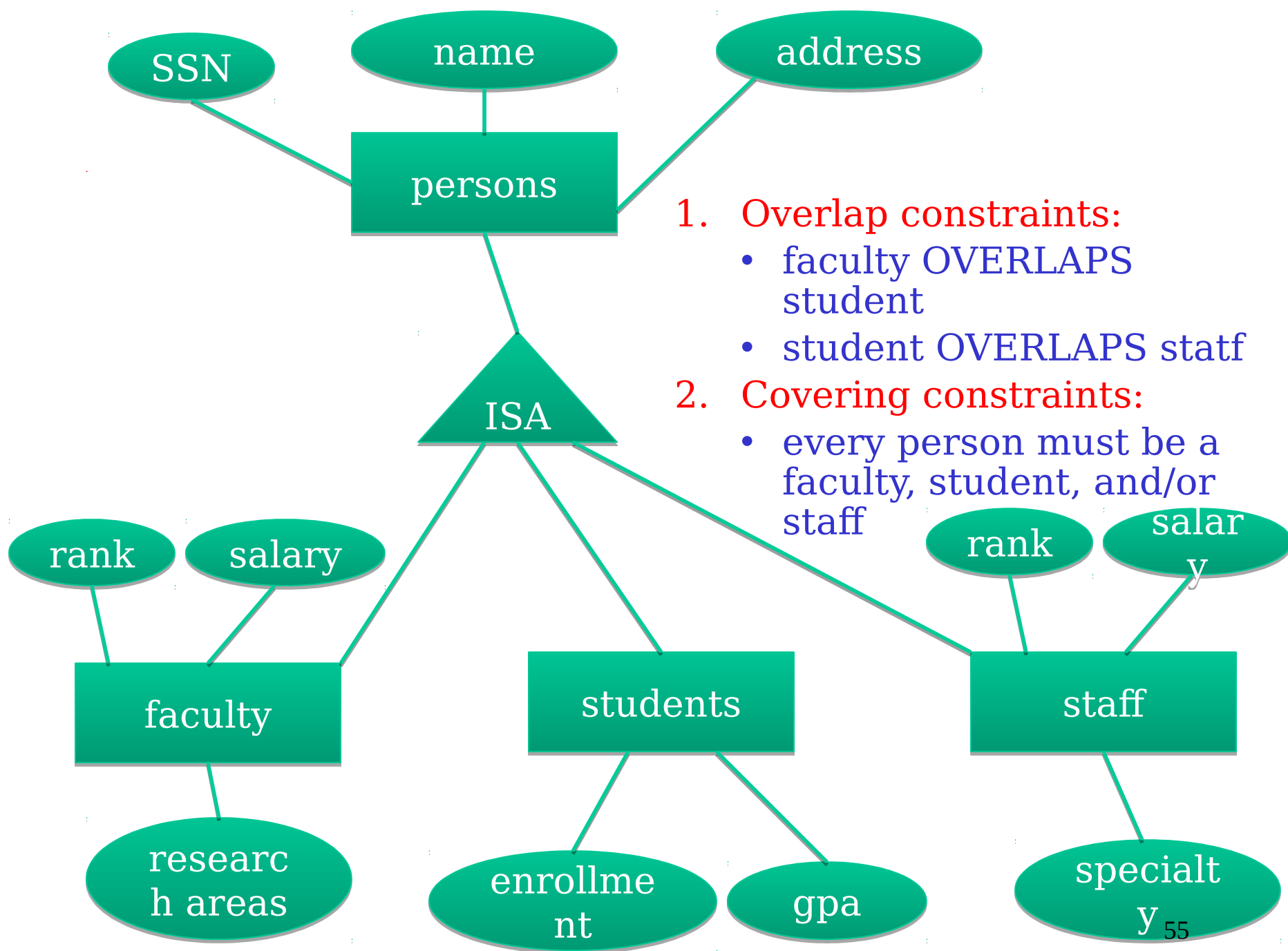
1. A hospital has a large number of registered physicians and patients.
2. Attributes of PHYSICIAN include Physician\_ID (identifier) and Specialty.
3. Attributes of Patients include Patient\_ID (identifier) and Patient\_Name.
4. Any patient who is admitted must have exactly one admitting physician. A physician may optionally admit any number of patients.
5. Once admitted, a given patient must be treated by at least one physician. A particular physician may treat any number of patients, or may not treat any patients.
6. Whenever a patient is treated by a physician, the hospital wishes to record the details of the treatment (Treatment\_Detail). Components of Treatment\_Detail include Date, Time, and Results.



# Exercise 4

- A university has a number of people
  - attributes: SSN, name, address
- A person can be a faculty, a student, a staff
  - faculty attributes: rank, salary, research areas, grants
  - student attributes: first enrollment date, gpa
  - staff attributes: rank, salary, specialty
- A person must be either a faculty, a student, and/or a staff
  - A student can be a staff and vice versa
  - A faculty can be a student and vice versa





# My shopping assistant

- I have a lot of things purchased from different stores
- I own several credit cards from some stores
- What do I want:
  - Get rid of very old, rarely used items
  - Keep track of what I have
  - Keep notes on what I need to buy
  - Keep track of my credit card usage
  - Keep a budget



## Exercise 5

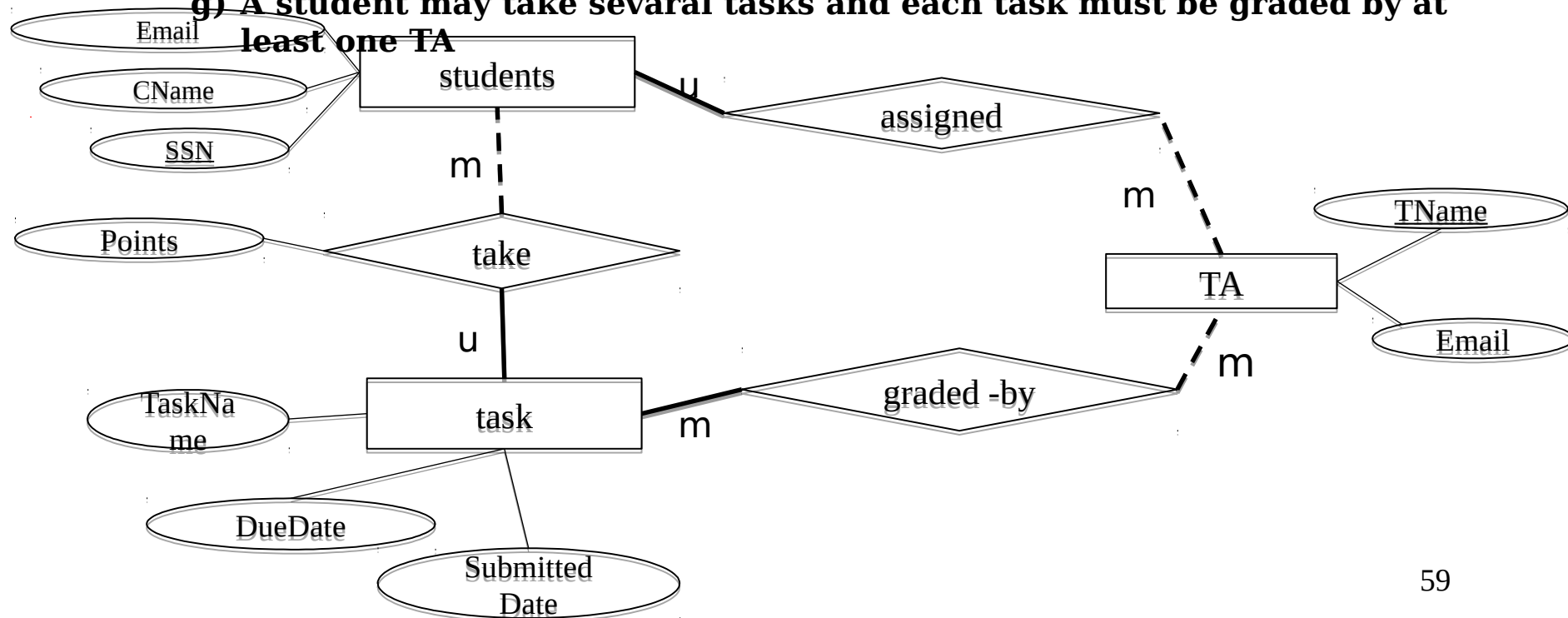
1.(10 points) Construct an E-R diagram for the instructor of COMS 363 to record the following information.

- 1) The instructor has a number of students and a number of teaching assistants (TAs)
- 2) Each student has SSN, CName, and Email
- 3) Each TA has a TName and Email
- 4) Each student is assigned with one and only one TA as his/her primary TA
- 5) A TA may or may not have any student
- 6) The course has a number of tasks (e.g., homework, project, exam), each of which has a TaskName, DueDate, SubmittedDate, and Points
- 7) A student may take several tasks and each task must be graded by at least one TA



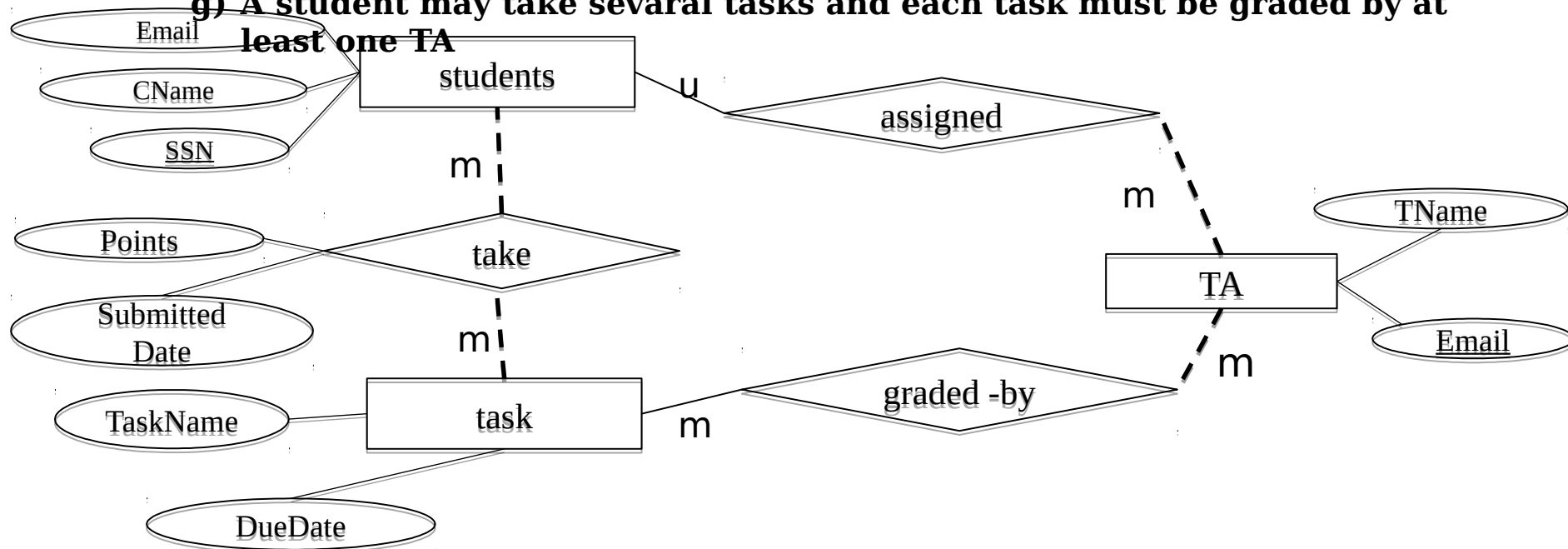
# Exercise 5

1. (10 points) Construct an E-R diagram for the instructor of COMS 561 to record the following information.
- a) The instructor has a number of students and a number of teaching assistants (TAs)
  - b) Each student has SSN, CName, and Email
  - c) Each TA has a TName and Email
  - d) Each student is assigned with one and only one TA as his/her primary TA
  - e) A TA may or may not have any student
  - f) The course has a number of tasks (e.g., homework, project, exam), each of which has a TaskName, DueDate, SubmittedDate, and Points
  - g) A student may take several tasks and each task must be graded by at least one TA



# Exercise 5 (design 1)

1. (10 points) Construct an E-R diagram for the instructor of COMS 561 to record the following information.
- a) The instructor has a number of students and a number of teaching assistants (TAs)
  - b) Each student has SSN, CName, and Email
  - c) Each TA has a TName and Email
  - d) Each student is assigned with one and only one TA as his/her primary TA
  - e) A TA may or may not have any student
  - f) The course has a number of tasks (e.g., homework, project, exam), each of which has a TaskName, DueDate, SubmittedDate, and Points
  - g) A student may take several tasks and each task must be graded by at least one TA



# Exercise 5

1. (10 points) Construct an E-R diagram for the instructor of COMS 561 to record the following information.

- a) The instructor has a number of students and a number of teaching assistants (TAs)
- b) Each student has SSN, CName, and Email
- c) Each TA has a TName and Email
- d) Each student is assigned with one and only one TA as his/her primary TA
- e) A TA may or may not have any student
- f) The course has a number of tasks (e.g., homework, project, exam), each of which has a TaskName, DueDate, SubmittedDate, and Points
- g) A student may take several tasks and each task must be graded by at least one TA

