DATABASE MANAGEMENT SYSTEMS (IT 2040)

LECTURE 01- INTRODUCTION TO DBMS AND DATABASE DESIGN PROCESS

LECTURE CONTENT

- Introduction to databases, DBMS and their benefits
- Database design process
- Requirement analysis
- Conceptual modelling using EER diagrams
- Design traps

LEARNING OUTCOMES

- Explain what is a database and a DBMS.
- Identify situations where using a database would be beneficial.
- Explain the database design process.
- Draw a EER diagram for a given scenario.

DATABASE DESIGN PROCESS

- There are six main phases of the process to develop a database
  - Requirement collection and analysis
  - Conceptual database design
  - Logical database design
  - Schema refinement
  - Physical database design
  - Security design

REQUIREMENT COLLECTION AND ANALYSIS

- The purpose of the phase is to collect and analyze the expectations of the users & the intended uses of the database.
- The process would include interviewing clients and analyzing documents such as files used to record data and reports to be generated.
- At the end of the requirement collection, the database developer should identify any unclear or incomplete requirements, redundant information and eliminate them.

REQUIREMENT ANALYSIS (CONTD.)

- Aspects to consider include
  - What data is to be stored in the database?
  - What applications are to be built?
  - What operations have to be performed?

CONCEPTUAL DATABASE DESIGN

- The result of the requirement analysis step is a concisely written set of users' requirements.
- Once, this step is completed, the next step is to create a **conceptual database schema** for the database, using a high-level conceptual data model.
- This step is called **conceptual database design**.
- **Entity-Relationship (ER) model** is a high-level conceptual data model.

ER MODEL - ENTITIES & ATTRIBUTES

- You already know these !
  - Check the handout for the definitions and examples for entities, different types of attributes and keys.
- Select an important entity in a context you are familiar. Add simple attributes, a multivalued attribute, a composite attribute and a key to the entity you identified.
  - Try not to use the attributes shown in the handouts.
- Exchange what you have drawn with your peer. What have they written?

ER MODEL - BINARY RELATIONSHIPS

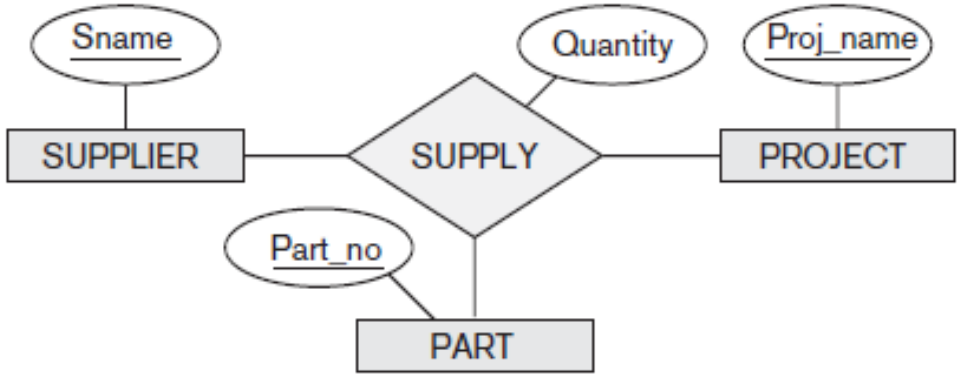
- You already know these too !
  - Check the handout for the definitions and examples for different cardinalities in binary relationships such as 1:1, 1:N and M:N
  - Now draw examples for each cardinality ratio above associated with binary relationships.
  - Exchange what you have drawn with your peer. What has he/she drawn?

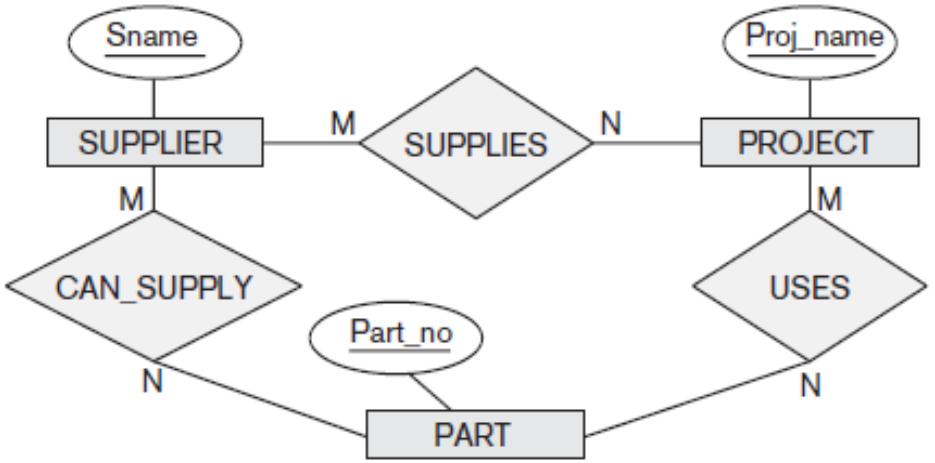
ER MODEL - WEAK ENTITIES

- Weak entities are entities that cannot be uniquely identified alone in a domain.
- Following restrictions must hold with relevance to weak entities
  - The owner entity set and the weak entity set must participate in a one-to-many relationship set (one owner entity is associated with one or more weak entities, but each weak entity has a single owner). This relationship set is called the **identifying relationship set** of the weak entity set.
  - The weak entity set must have total participation in the identifying relationship set.
- Can you think of an weak entity in a domain you know?

ER MODEL - TERNARY RELATIONSHIPS

- A ternary relationship is when three entities participate in the relationship.
- When to use ternary vs binary?





- Ternary relationships could be used when there is a common attribute that needs all three entities together, (ex: quantity)
- Include the ternary relationship *plus* one or more of the binary relationships, if they represent different meanings and if all are needed

ACTIVITY

- Draw an ER diagram for the scenario below.
  - A Library is organized into several sections such as fiction, children and technology. Each section has a name and a number(unique) and its headed by a head librarian.
  - Each book belong to a section and has a title, authors, ISBN, year and a publisher.
  - A book may have several copies. Each copy is identified by an access number.
  - For each copy borrowed, current borrower and due date should be tracked.
  - Members have a membership number(unique), an address and a phone number.
  - Members can borrow 5 books and could put hold request on a book
  - Librarian has a name, id number(unique), phone and an address.

EER MODEL

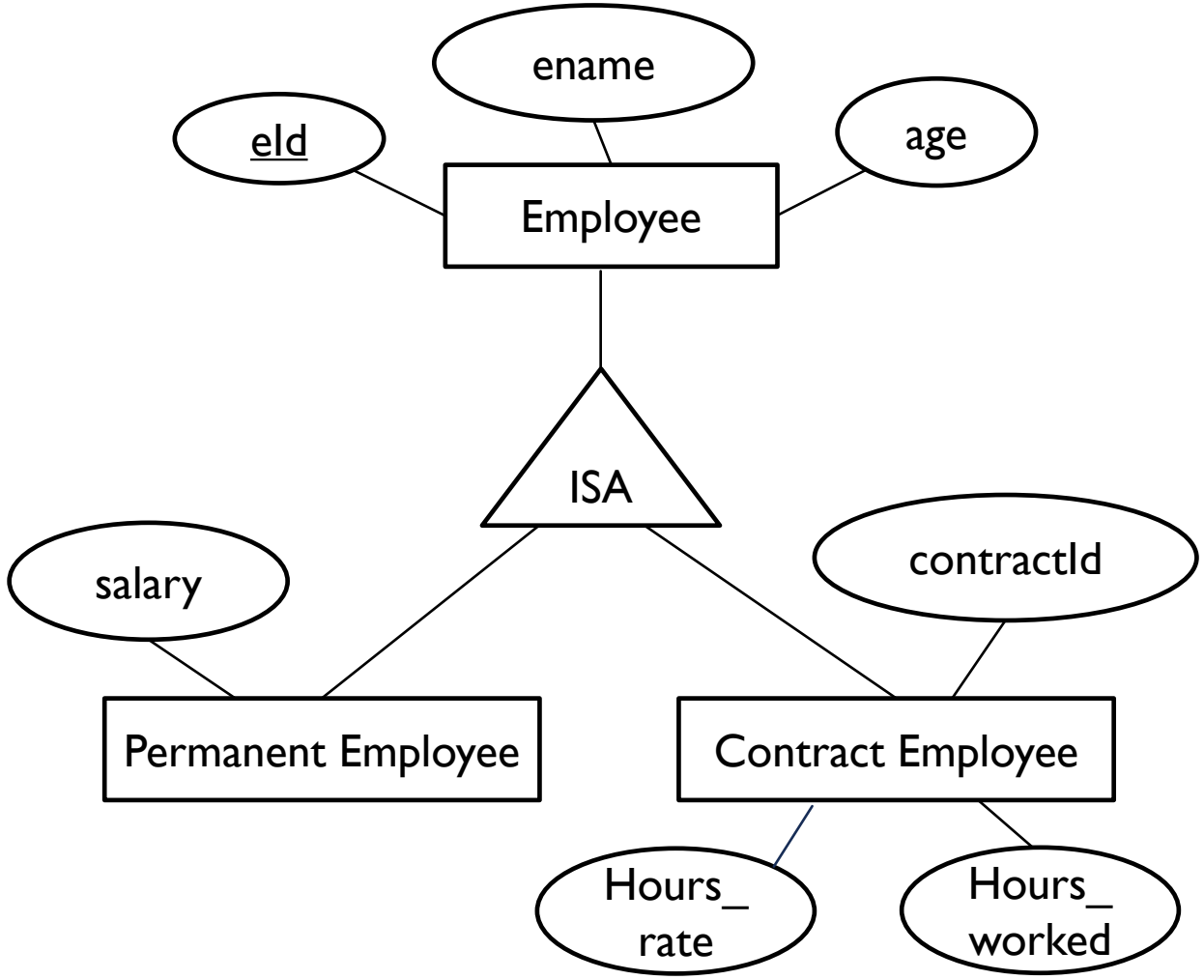
- ER model we discussed so far has been enhanced by adding several new concepts leading to the development of the EER model.
- An important extension included in the EER model is the specialization and generalization concepts.
  - Specialization is the process of defining a set of subclasses of an entity type.
    - Employee & permanent employee
  - Generalization is the process of identifying commonalities between entity types and grouping them as super-classes.

EER MODEL - ISA RELATIONSHIPS

- In many cases an entity type has numerous subgroupings or subtypes of its entities that are meaningful and need to be represented explicitly because of their significance to the database application.
  - Ex: the entities that are members of the EMPLOYEE entity type may be distinguished further into contract employees and permanent employees
- Such subtypes could be represented in EER diagrams using 'ISA' relationships

EER MODEL - ISA RELATIONSHIPS (CONTD.)

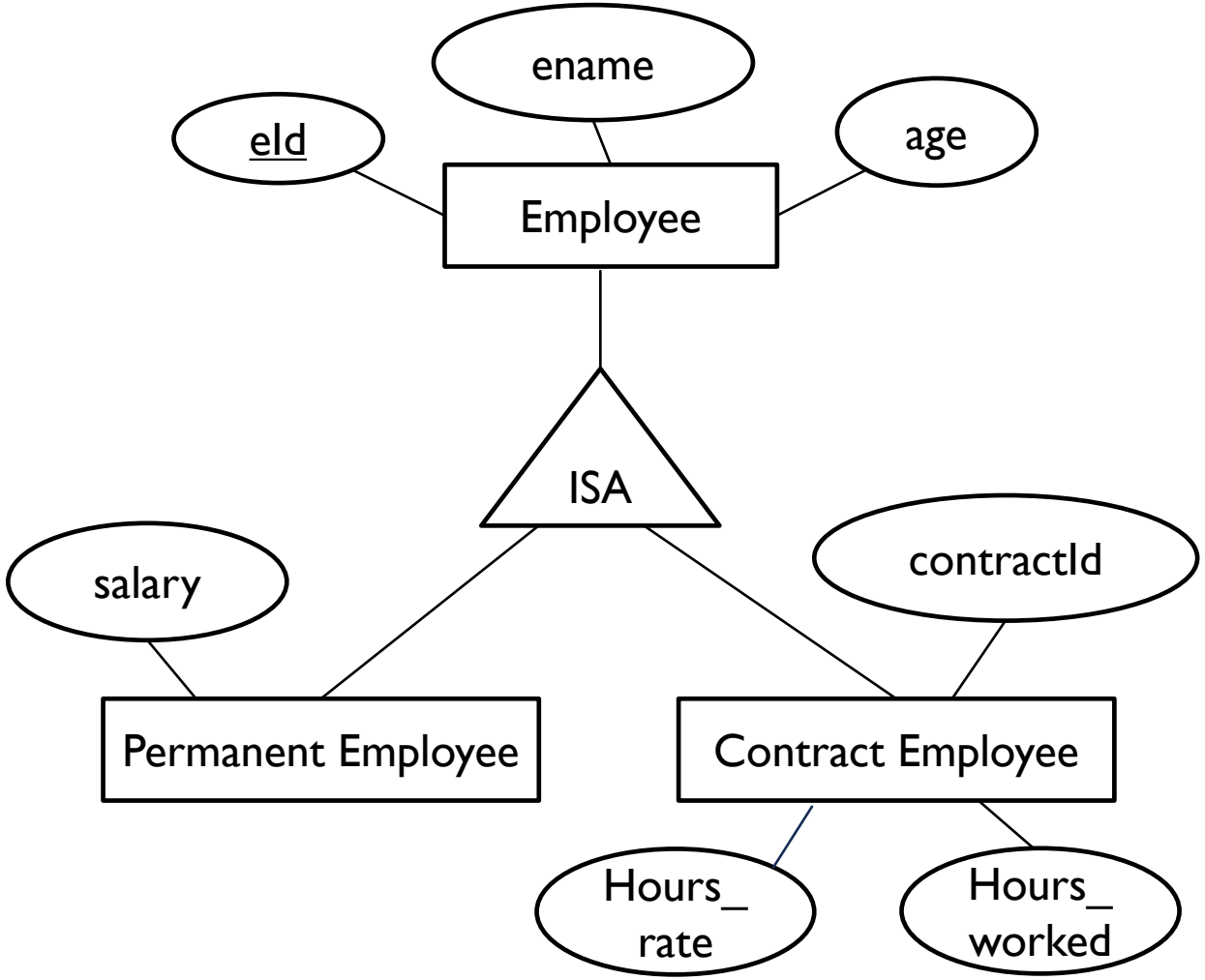
- Note that the subclasses may have their own attributes and relationships.
- Every entity in the subclass is also an employee entity and have all the have all of the attributes of Employees entity.
- Thus, attributes of the permanent employee include all attributes of employee entity and those of permanent employee.

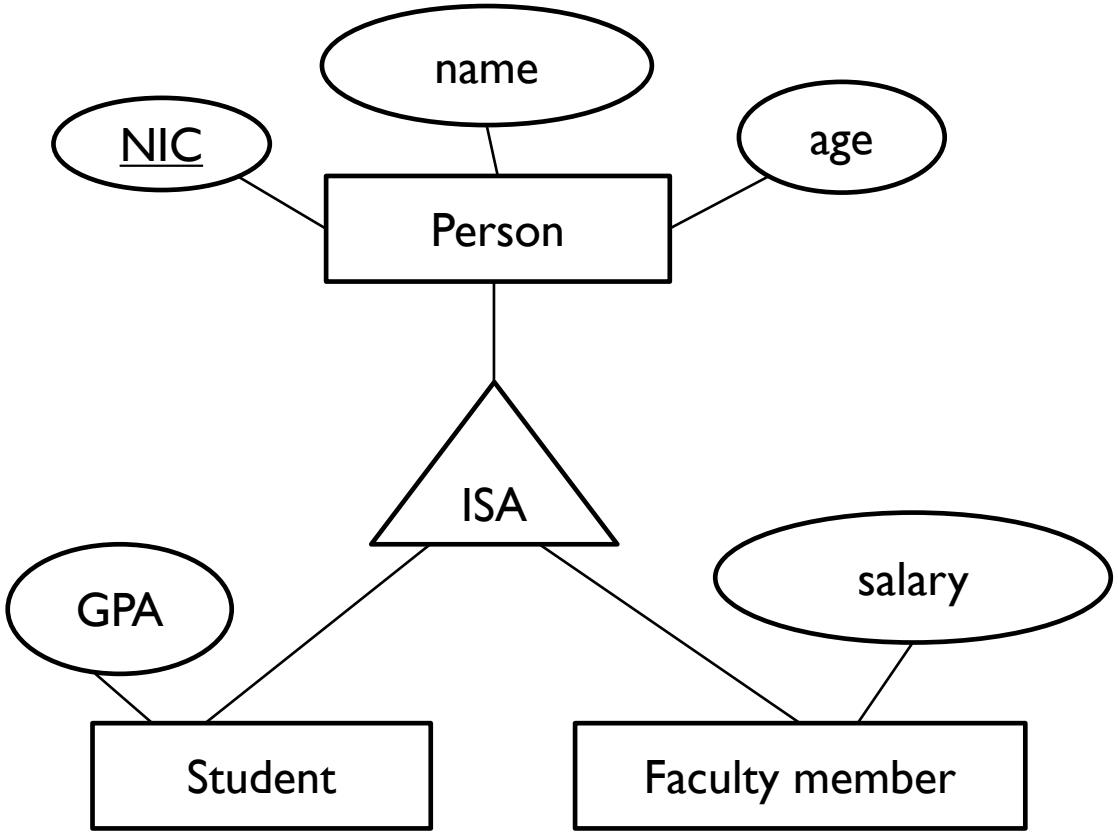


OVERLAPPING CONSTRAINT

- **Overlapping constraint** determine whether two subclasses are allowed to contain the same entity.
  - For example can an employee E be a permanent employee and a contract employee? Probably not. Therefore, the permanent employee subclass and the contract employee subclass are disjoint.
  - Can a person P in a university environment be a student and a faculty member at the same time. If it is so, we denote this by writing student overlaps faculty member. In absence on such a statement we assume that the sub classes are disjoint.

OVERLAPPING CONSTRAINT (CONTD.)



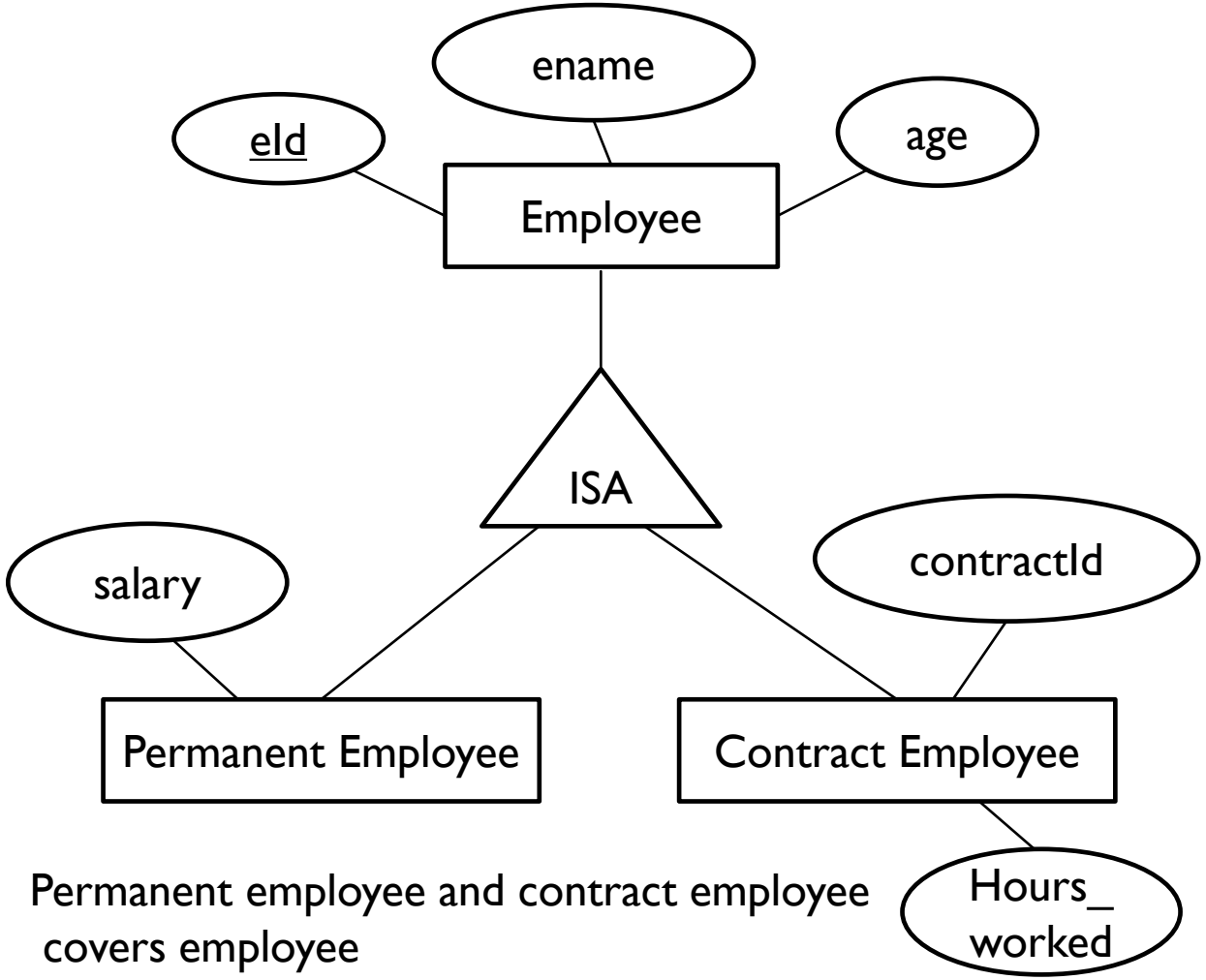


Student overlaps faculty member

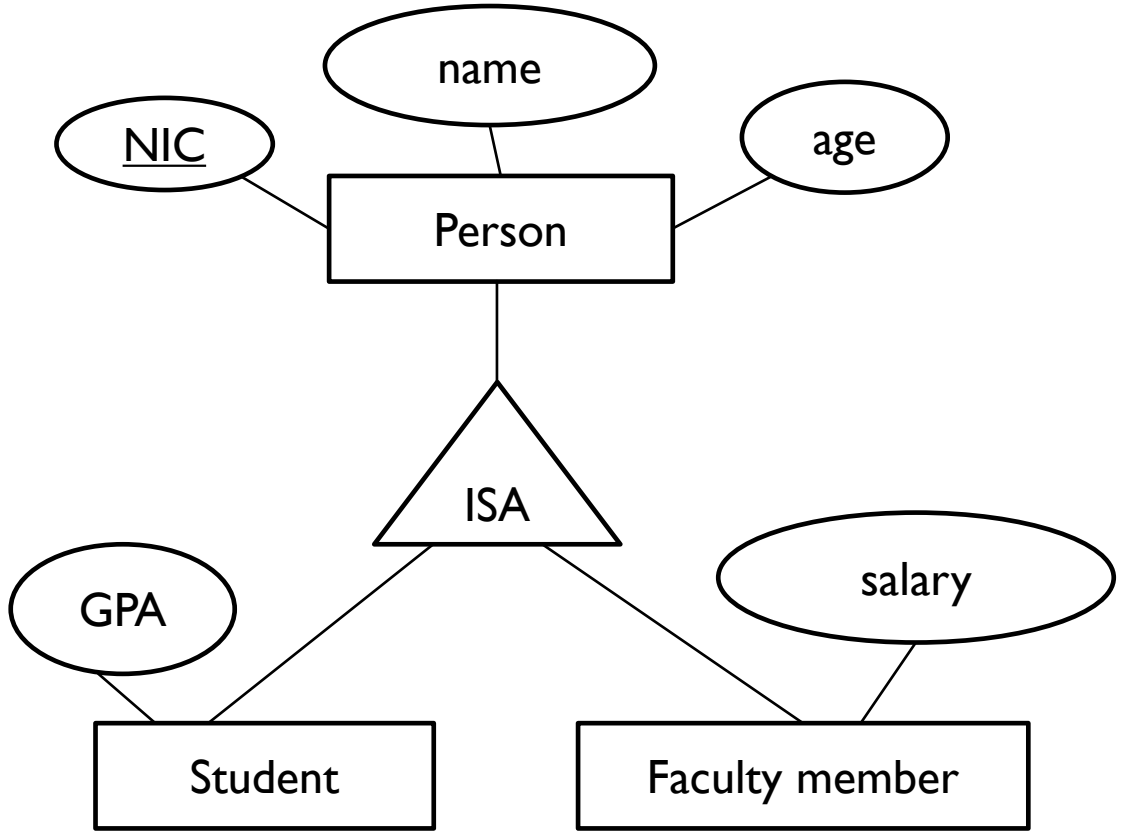
COVERING CONSTRAINT

- **Covering constraints** determine whether the entities in the subclasses collectively include all entities in the super class.
  - For example, does every employee entity e, belong to one of its subclasses (i.e. permanent employee or contract employee)? If so we denote this by writing permanent employee and contract employee covers employee.
  - Does every person p in a university environment belong to either student sub class or the faculty member sub class? Probably not. Therefore, there is no covering constraint associated with the hierarchy.
- Existence of a covering constraint is also know as having a total specialization.
- Absence of a covering constraint in a class hierarchy is known as partial specialization.

COVERING CONSTRAINT (CONTD.)



Permanent employee and contract employee covers employee

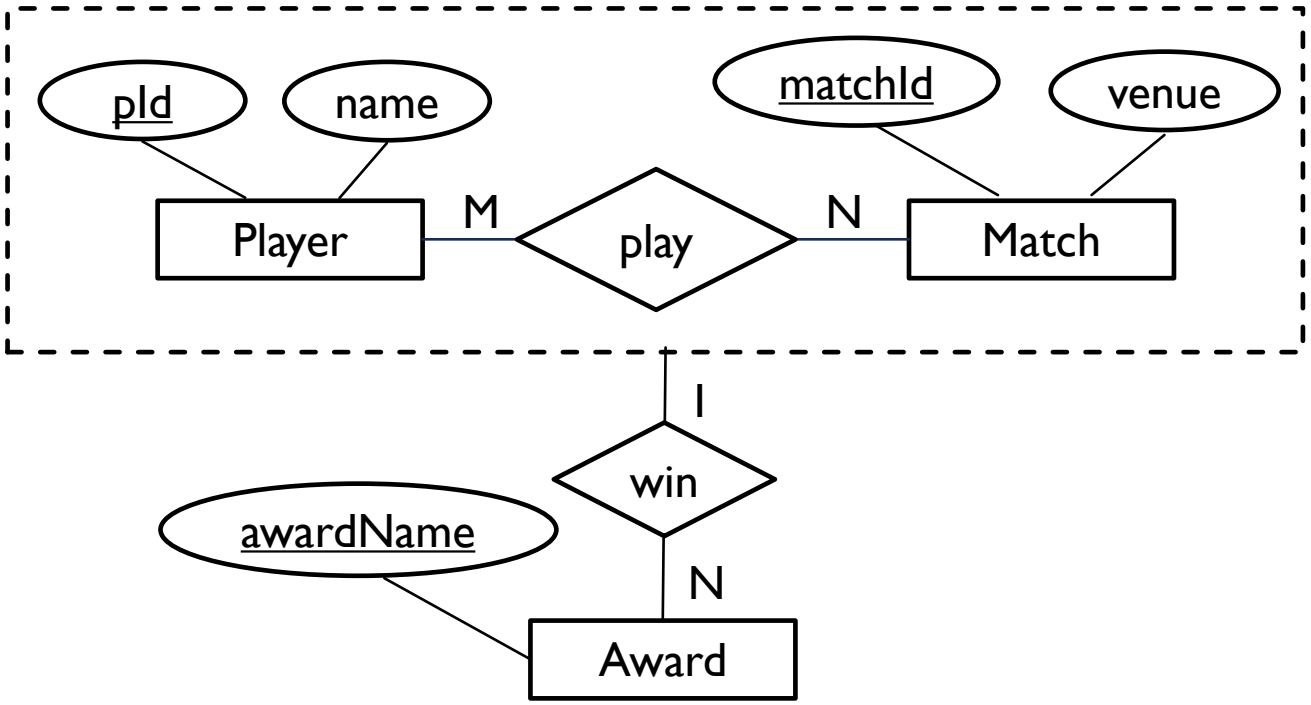


ACTIVITY

- In a blank paper draw two ISA relationships, one which is having a covering constraint and another which is having a overlapping constraint.
- Exchange what you have drawn with your peer. What have they drawn?
- Have you understood the concepts properly?

EER MODEL - AGGREGATION

- **Aggregation** allows us to indicate that a relationship as a above which is between a relationship set that participates in another relationship set.
- For example, a cricket players play in cricket matches. When he plays a match for his performance he may win awards.
- Note that, the difference between ternary and aggregation is that aggregation contain two independent relationships whereas in ternary relationship there is one.



ACTIVITY

- Can you think of an aggregation relationship in a domain familiar to you?
- Exchange what you have drawn with your peer. What have they drawn?
- Have you understood the concepts properly?

ACTIVITY

- Draw an E-ER diagram for the following requirements.
  - Students contain an id (unique), name and an address.
  - There are academic semesters containing an *semester id* (unique), *semester* and *year*.
  - There are courses offered during academic semesters. A course has a *number* (unique), *name* and *credits*.
  - Students make payments. A payment has *receipt number* (unique), *amount* and *date*.
  - Payments can be classified into Tuition (semester payment), Examination and other (Library fine, Printouts).
  - A Tuition payment is made for an academic semester
  - For other payments description should be stored
  - Students register for courses offered during a particular semester. The registered date must be stored in the database.

DESIGN TRAPS

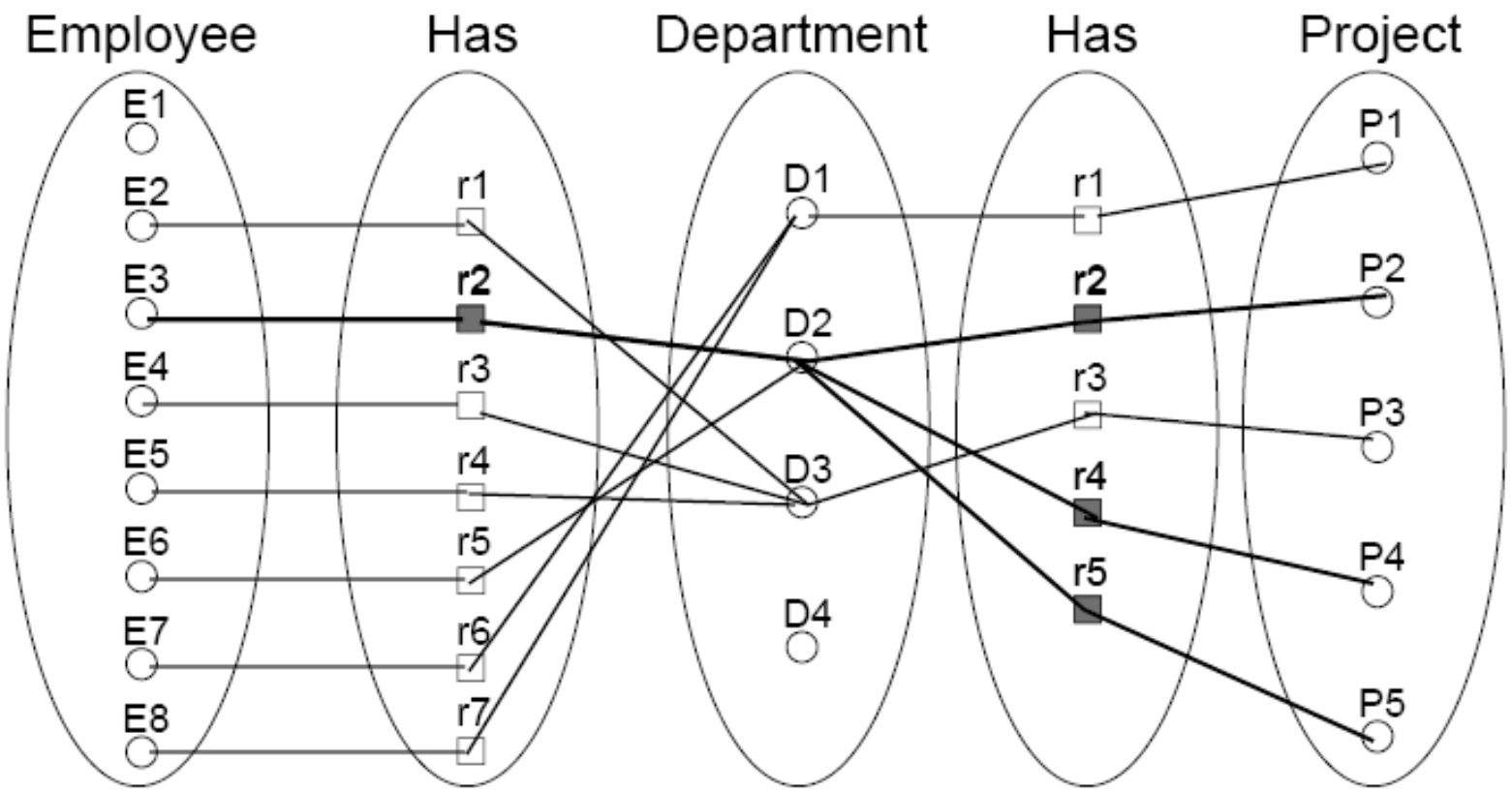
- There are several different "modeling traps" (called *connection traps*) that you can fall into when designing your ER model.
- Two connection traps that we will look at are:
  - Fan traps
  - Chasm traps

FAN TRAP

- A ***fan trap*** is when a model represents a relationship between entity types, but the pathway between certain entity instances is ambiguous.
  - Often occurs when two or more one-to-many relationships fan out (come from) the same entity type.
- Example: A department has multiple employees, a department has multiple projects, and each project has multiple employees.

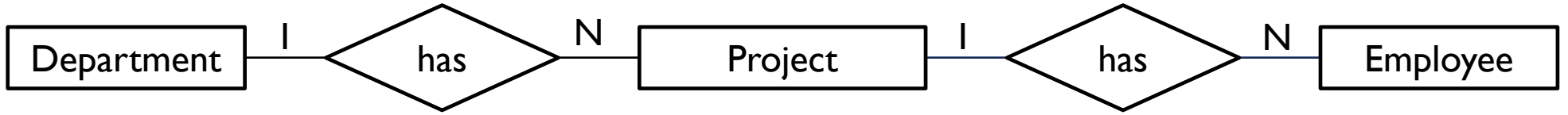
FAN TRAP (CONTD.)

- Which projects does employee E3 work on?



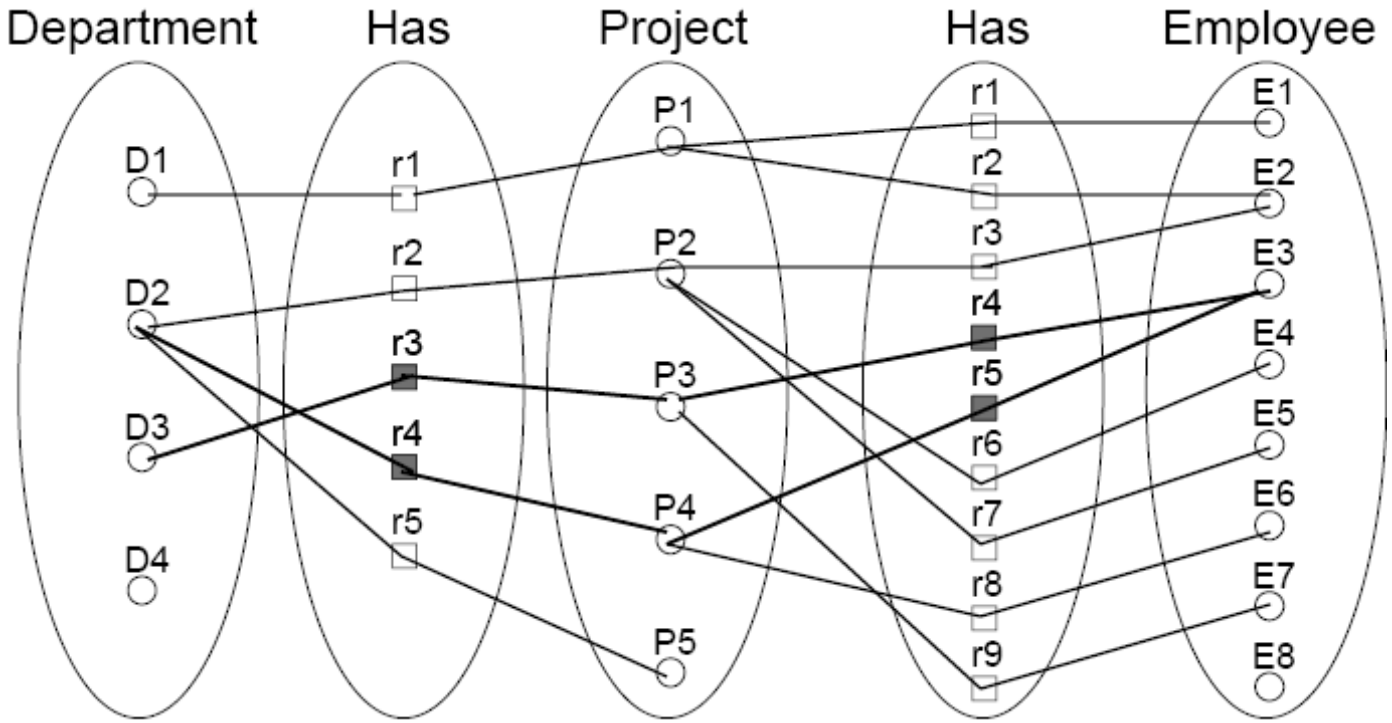
CHASM TRAP

- A ***chasm trap*** occurs when a model suggests that a relationship between entity types should be present, but the relationship does not actually exist. (*missing* relationship)
  - May occur when there is a path of optional relationships between entities.
- Example: A department has multiple employees, a department has multiple projects, and each project has multiple employees.



CHASM TRAP (CONTD.)

- Which department is employee E8 in?
- What are the employees of department D4?



WHAT YOU HAVE TO DO BY NEXT WEEK

- Try out the self-test questions on the course web.
- Try out tutorial and bring the answers to the class.
- Answer the questions at the end of chapter 2 of Database Management Systems by Ramakrishnan & Gehrke