

Introduction to database design

KBL chapter 4
(cursory reading: 4.6, and UML examples)

Rasmus Pagh

Some figures are borrowed from the ppt slides from the book used in the course, Database systems by Kiefer, Bernstein, Lewis
Copyright © 2006 Pearson, Addison-Wesley, all rights reserved.

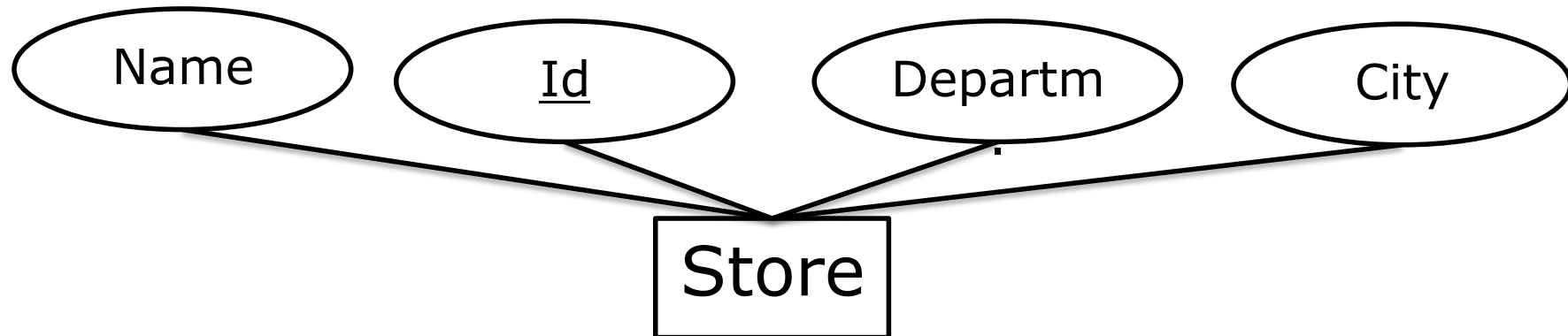


Today's lecture

- Entity-Relationship (E-R) model concepts
 - entities
 - relationships
 - cardinality constraints
 - entity type hierarchies
 - weak entities
- Translation to a schema in SQL
- Hand-in, part 1



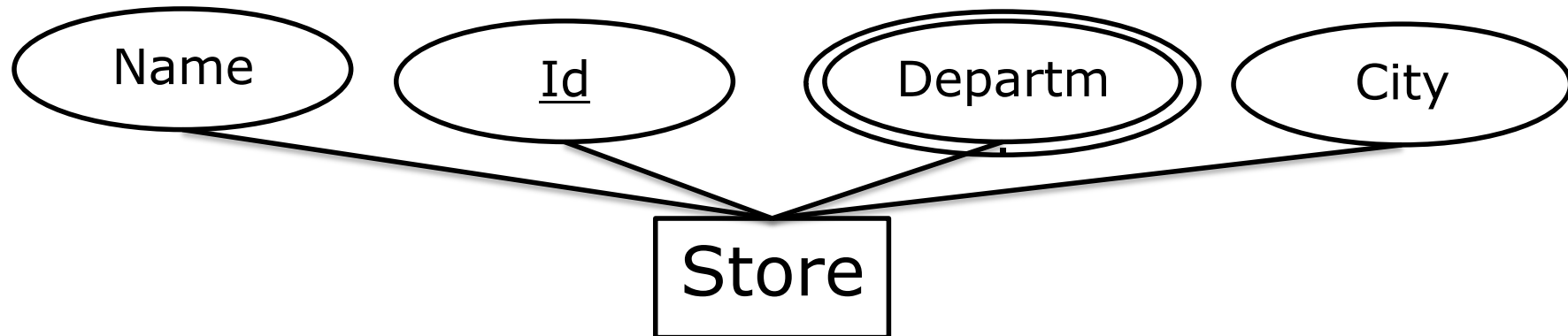
Entity



- An **entity** models an **object**
- An **entity type** (rectangle) models a class of entities
 - Described by **attributes** (ovals)
 - Can have a **key** (underlined)



Entity



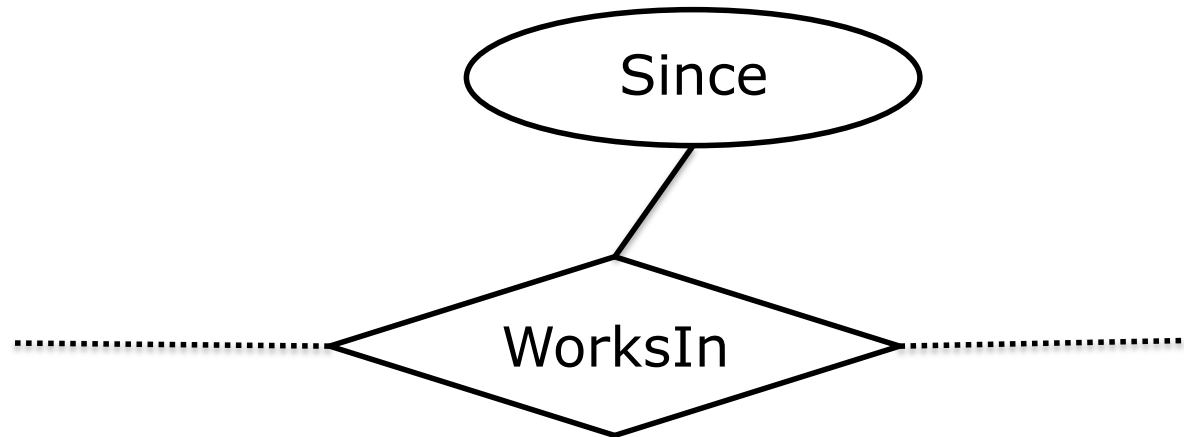
Contrary to the relational model, attributes can be **sets** (double oval).

A **schema** for an entity type

- the name,
- attributes (single or set valued) + domains,
- key constraints.



Relationships

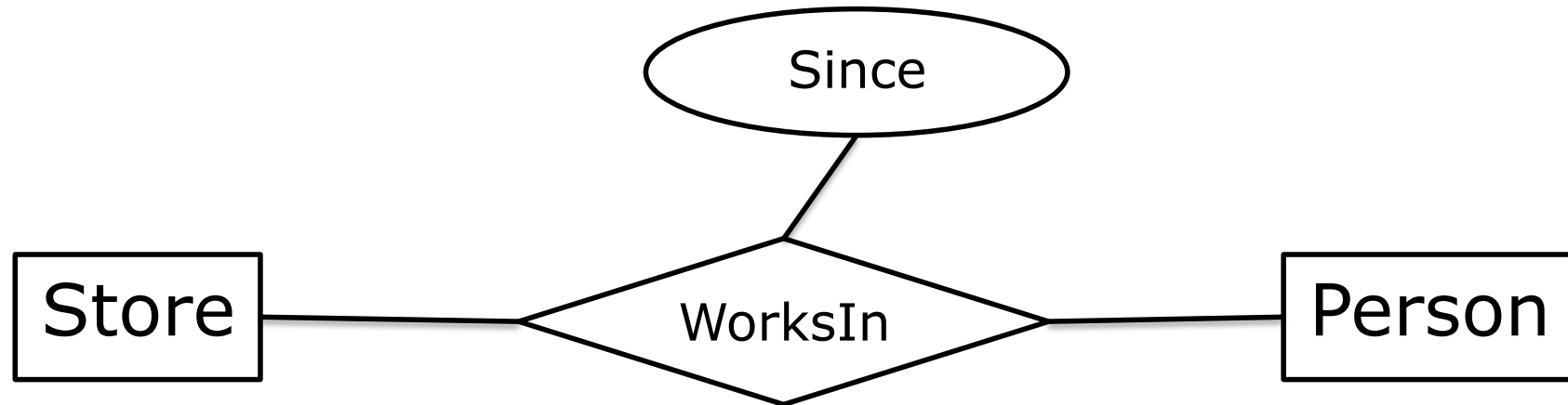


A relationship type (diamond) models the mechanism that **relates** entities to each other.

A relationship type may have **attributes**.



Relationships

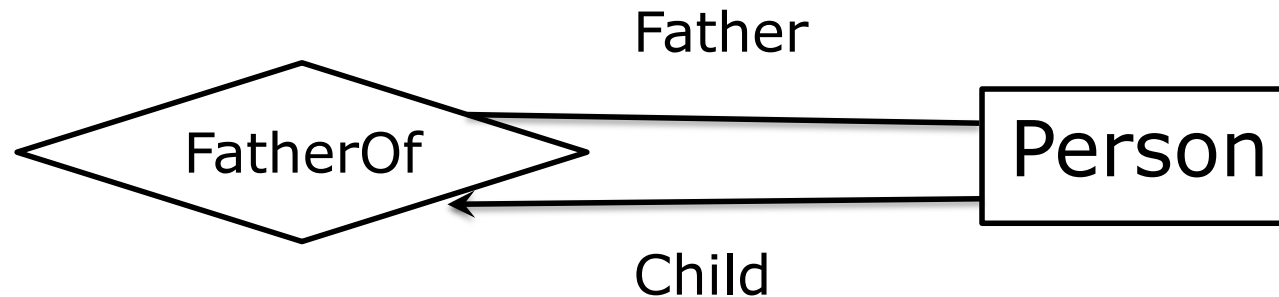


A relationship is connected to one, two or more entity types, called **roles**.

A role specifies **how** an entity participates in a relationship.



Relationships

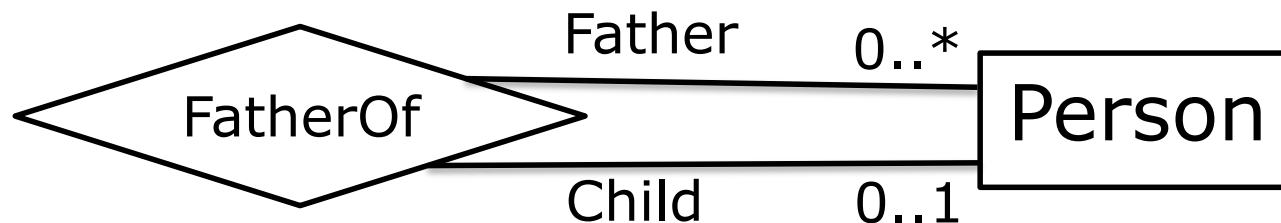


One entity can participate in **several** roles in a relationship.



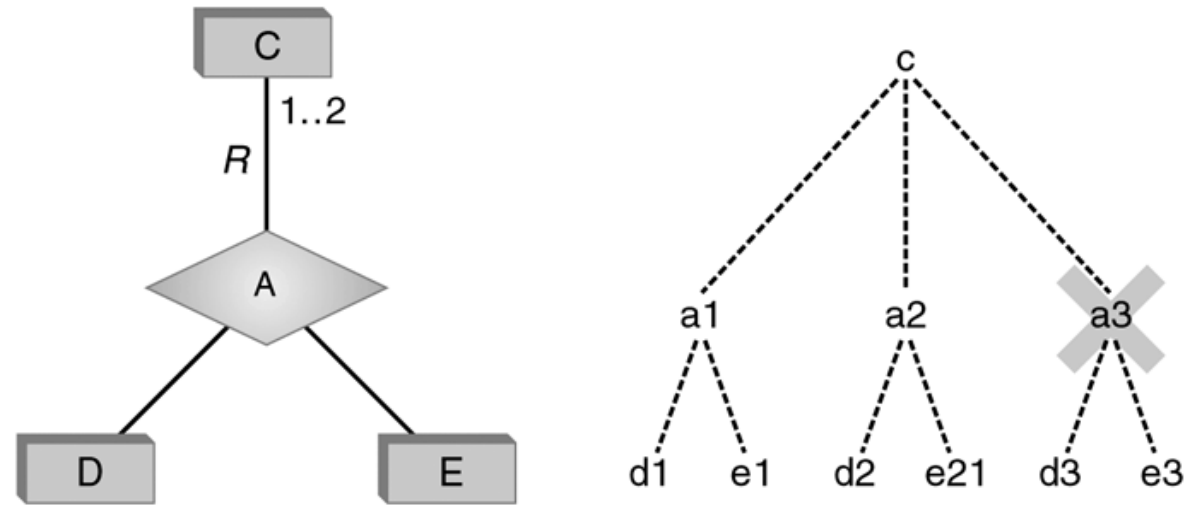
Cardinality constraints

- Constrains the number of times an entity can participate in a relationship in a specific role.
- Has the form *min..max*.
- Can be expressed explicitly or by different types of arrows



Cardinality constraints

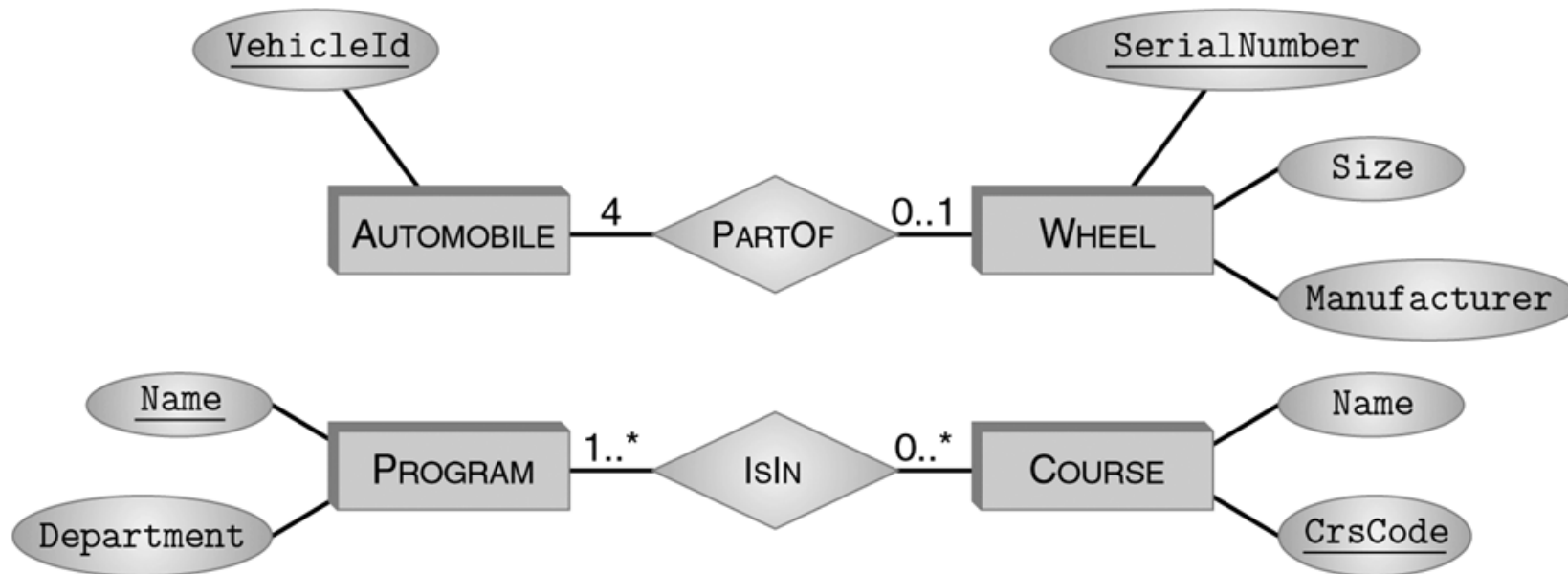
FIGURE 4.3 Cardinality in the E-R model.



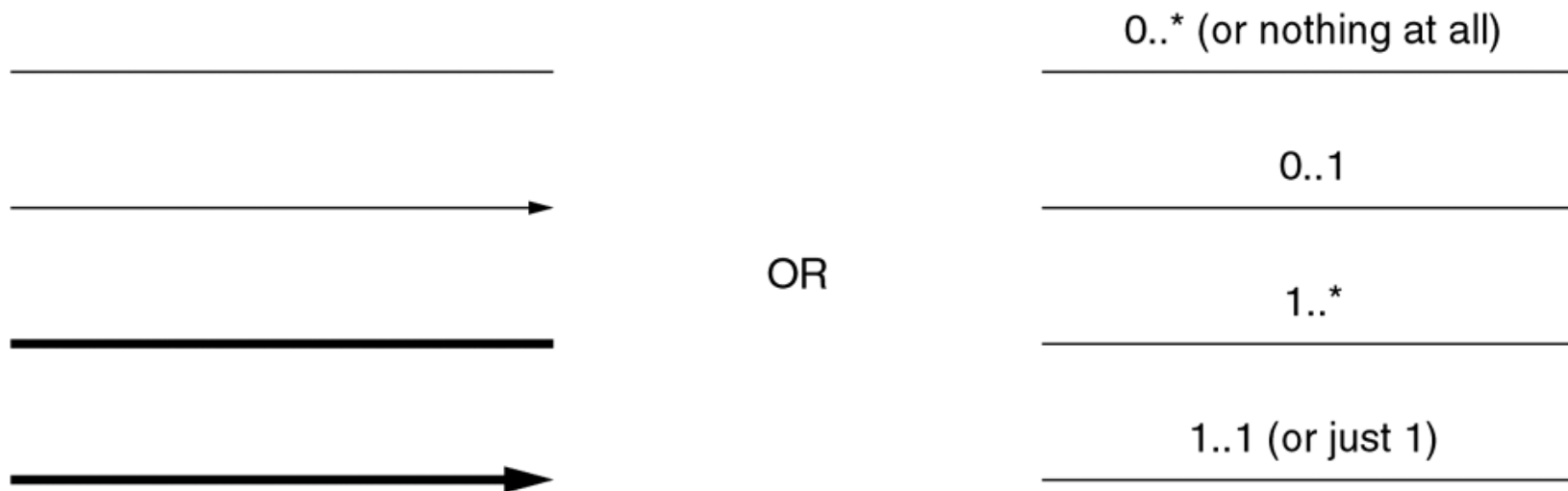
- NB! Different from UML semantics!



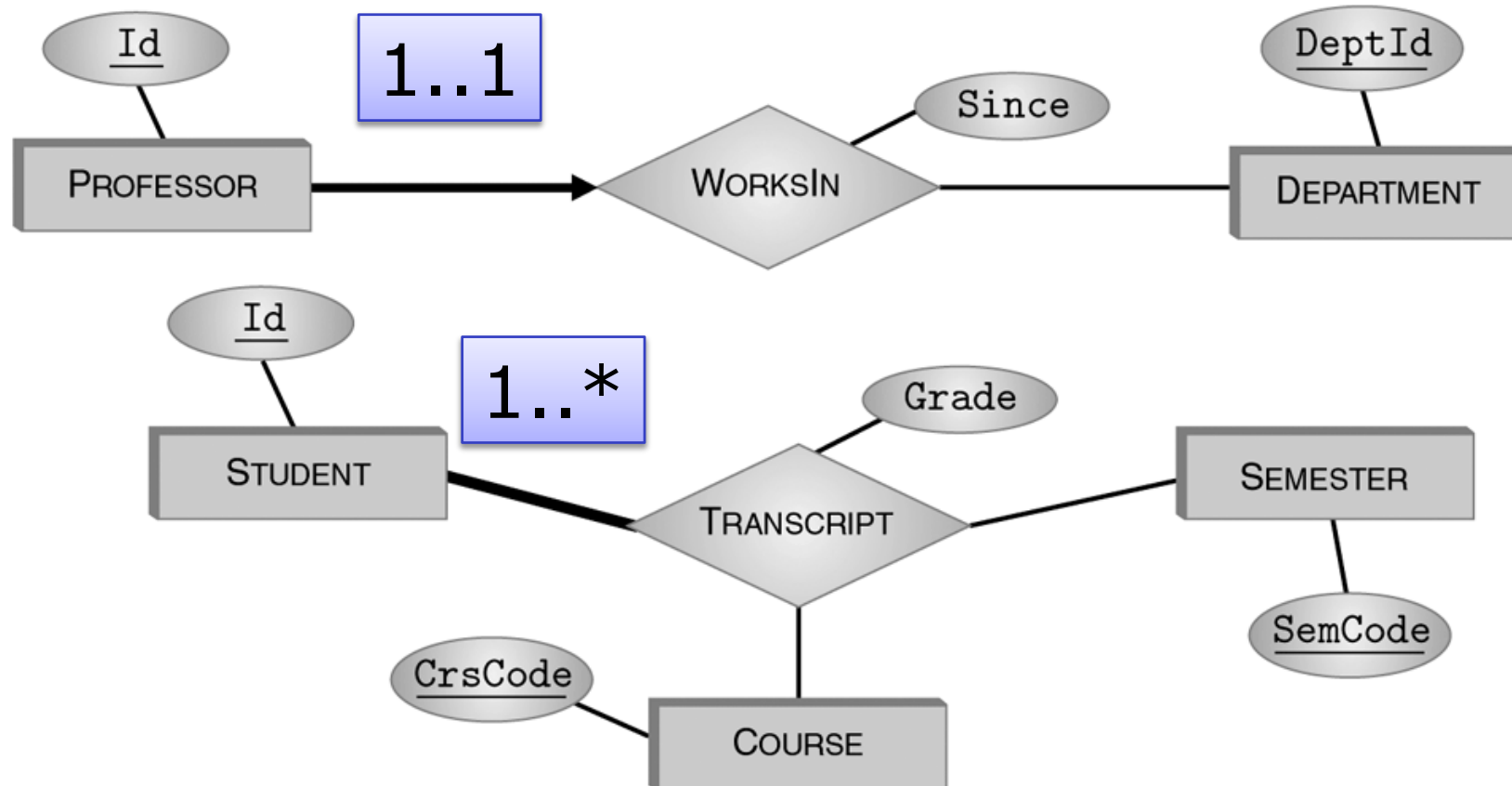
Example



Alternative cardinality notation



Participation constraints



When $\text{min} \geq 1$ a cardinality constraint is called a *participation constraint*.



Case study

- For each researcher, his/her name, year of birth, and current position (if any).
- For each institution, its name, country, and inauguration year.
- For each institution, the names of its schools (*e.g. School of Law, School of Business, School of Computer Science,...*). A school belongs to exactly one institution.
- An employment history, including information on all employments (start and end date, position, and what school).
- Information about co-authorships, i.e., which researchers have co-authored a research paper. The titles of common research papers should also be stored.
- For each researcher, information on his/her highest degree (BSc, MSc or PhD), including who was the main supervisor, and at what school.
- For each professor, information on what research projects (title, start date, and end date) he/she is involved in, and the total amount of grant money for which he/she was the main applicant.



Problem session

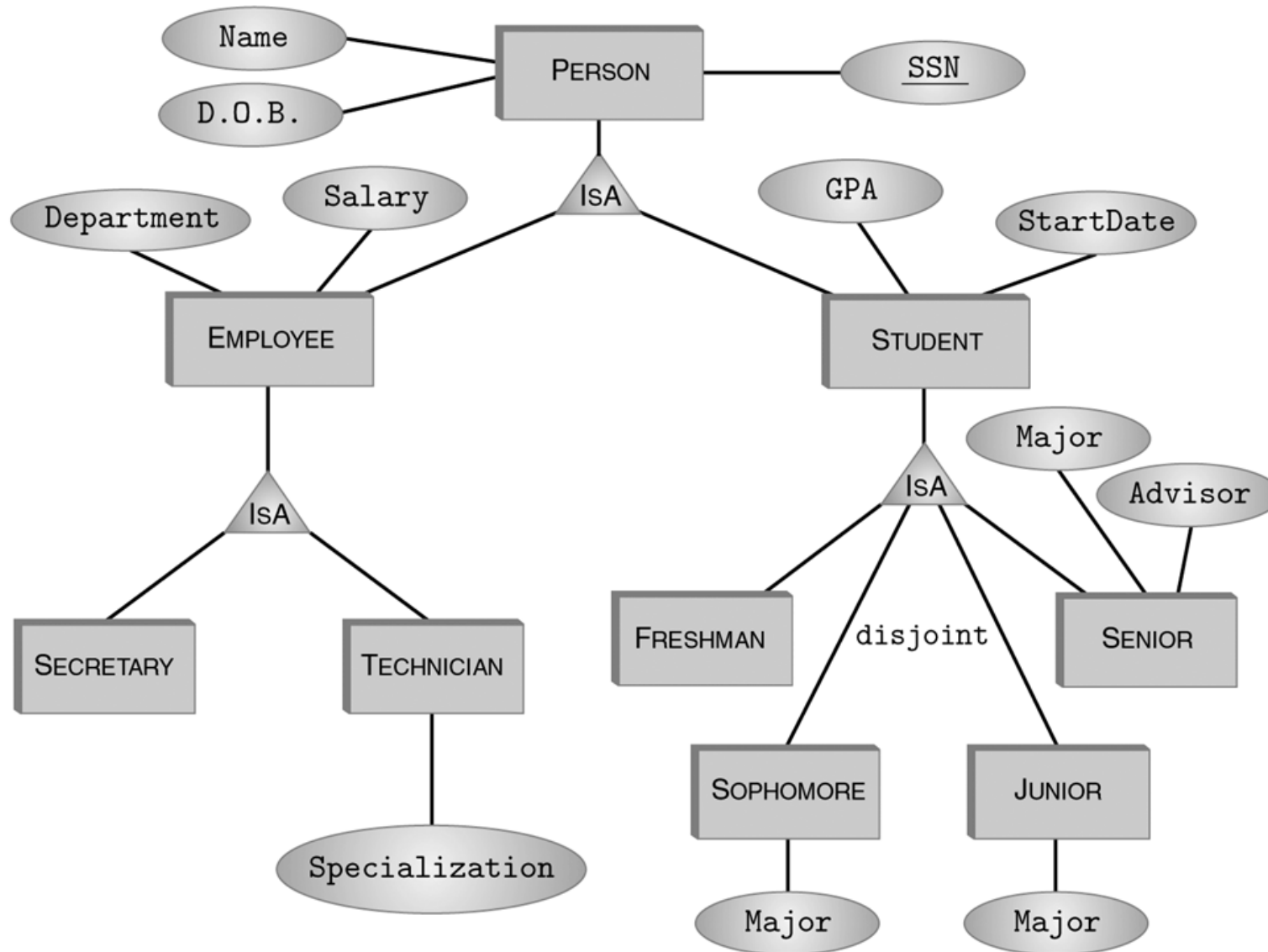
- Create an E-R diagram for the course manning database we discussed last week. Basic attributes summarized:

Line	Course	Semester	Group	Teacher	Contribution	ECTS
------	--------	----------	-------	---------	--------------	------

- Try to formulate the assumptions you make, e.g.:
 - Can a course be on several study lines?
 - Is there a limit to the number of teachers on a course?



Type hierarchy (IsA)

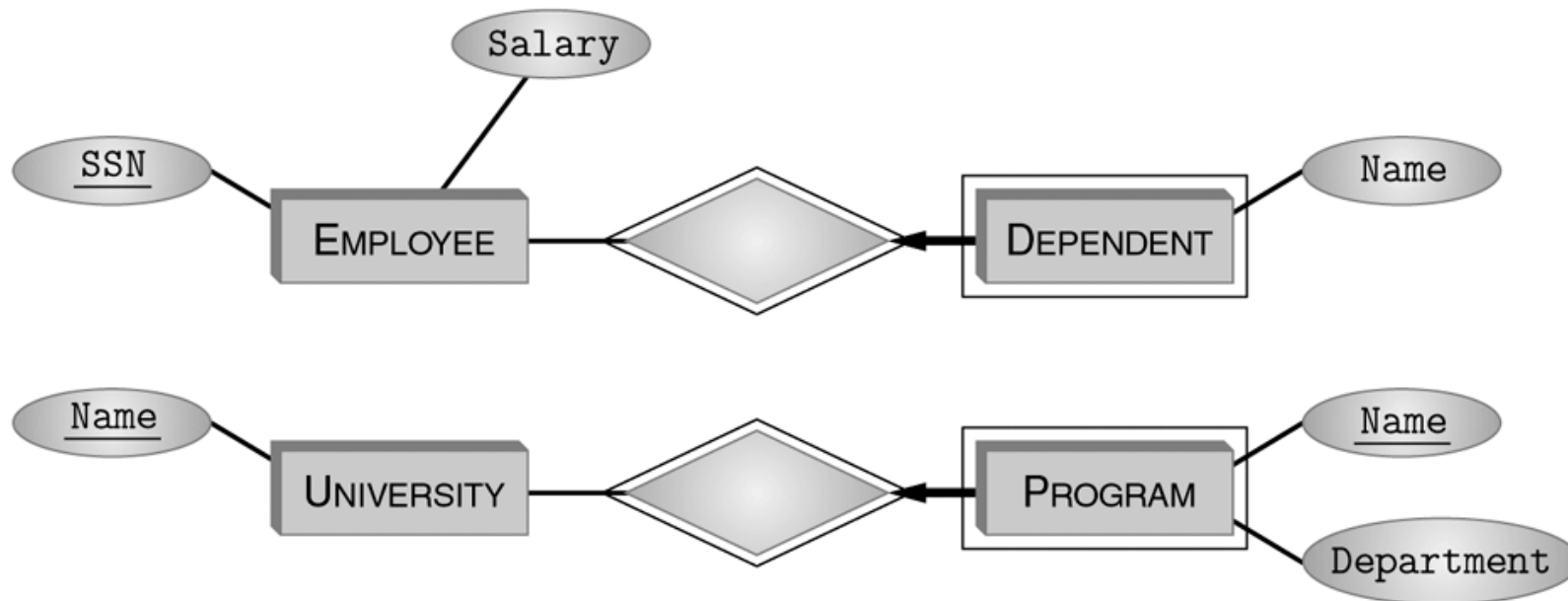


Type hierarchy (IsA)

- When an entity type A is a **supertype** of entity type B it means that an instance of B is also an instance of A.
- A **classification hierarchy** where attributes are **inherited**.
- **IsA relationship** is a relationship describing a type hierarchy.
- **Covering constraint**: All instances of a super type are also instances of some subtype.
- **Disjointness constraint**: Every instance of a supertype is only an instance of one subtype

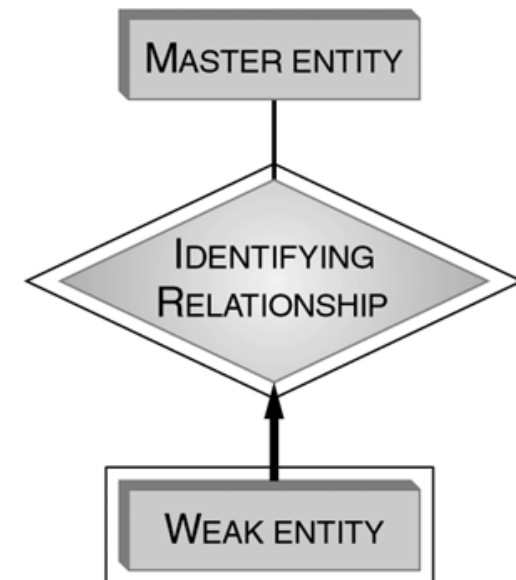


Weak entity types

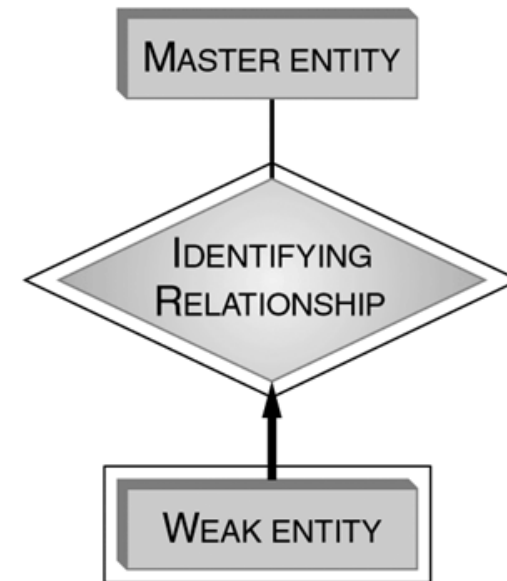
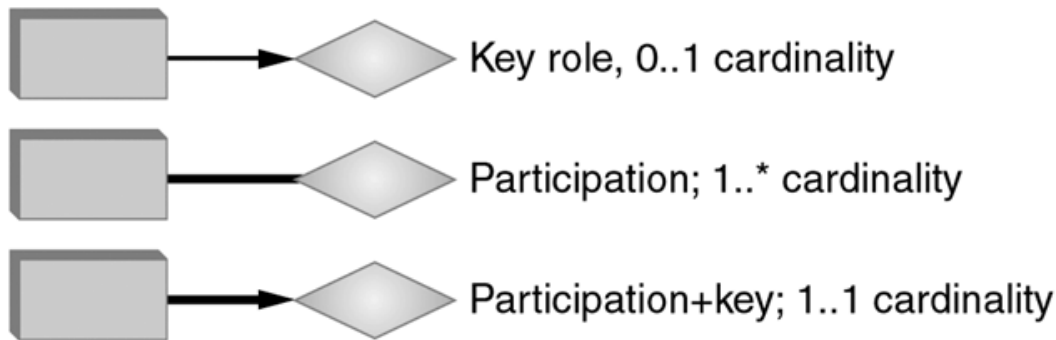
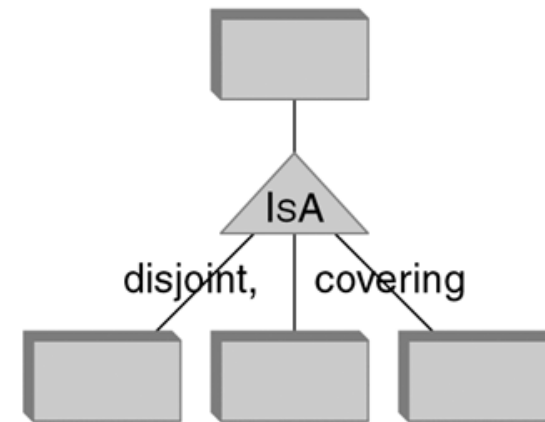
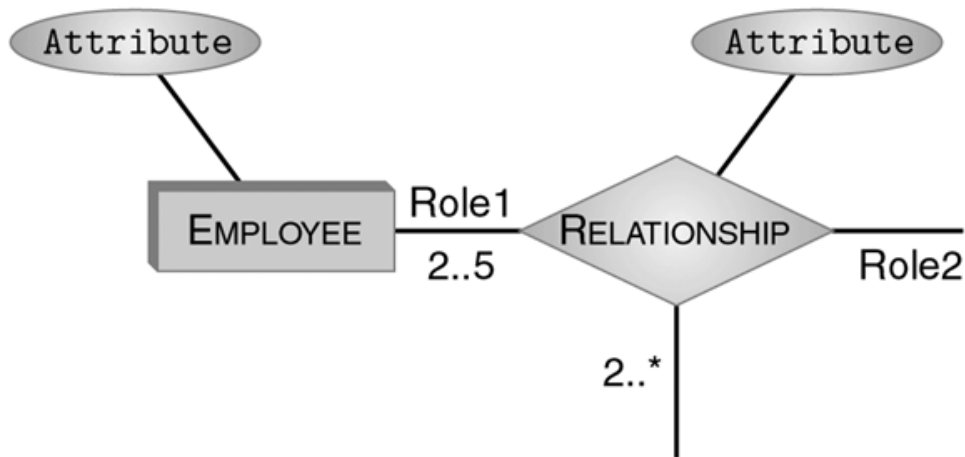


Weak entity types (or exclusive part-of)

- Definition:
A **weak entity** needs a foreign key to be uniquely identified.
- In the E-R model the foreign key must be modeled as a relationship. This is called an **identifying relationship**.

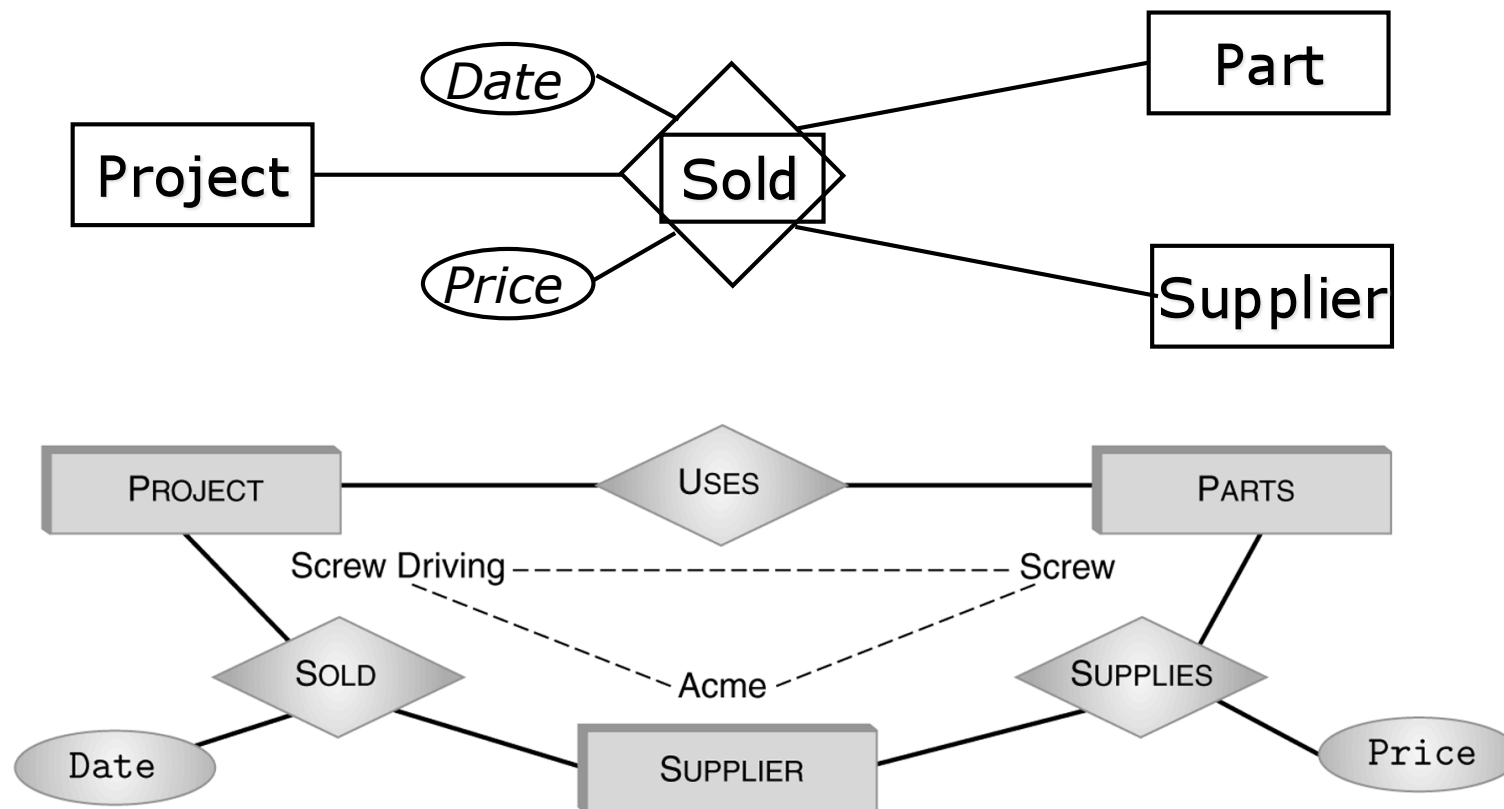


Summary



Problem session

- What is the difference, if any, between these E-R diagrams?



Translation of E-R diagrams to relational schemas

- Step-by-step description in the book for translating a E-R diagram to a schema in SQL
- Also a creative process
 - there is not always only one way of doing it, and
 - there may be things in the E-R diagram that cannot (easily) be expressed in SQL



Entities

(For now, ignore all set valued attributes)

- Entity type becomes a relation
- Attributes become attributes
- The key attributes become primary key



Relationships

- The relationship becomes a relation
- Attributes:
 - Attributes of the relationship
 - The primary keys of the roles involved in the relationship (declare NOT NULL)
- Primary key (composite):
 - Any key attributes of the relationship
 - The primary keys from the roles in the key
- Foreign key references:
 - To primary keys of the roles



Relationships

A very common special case:

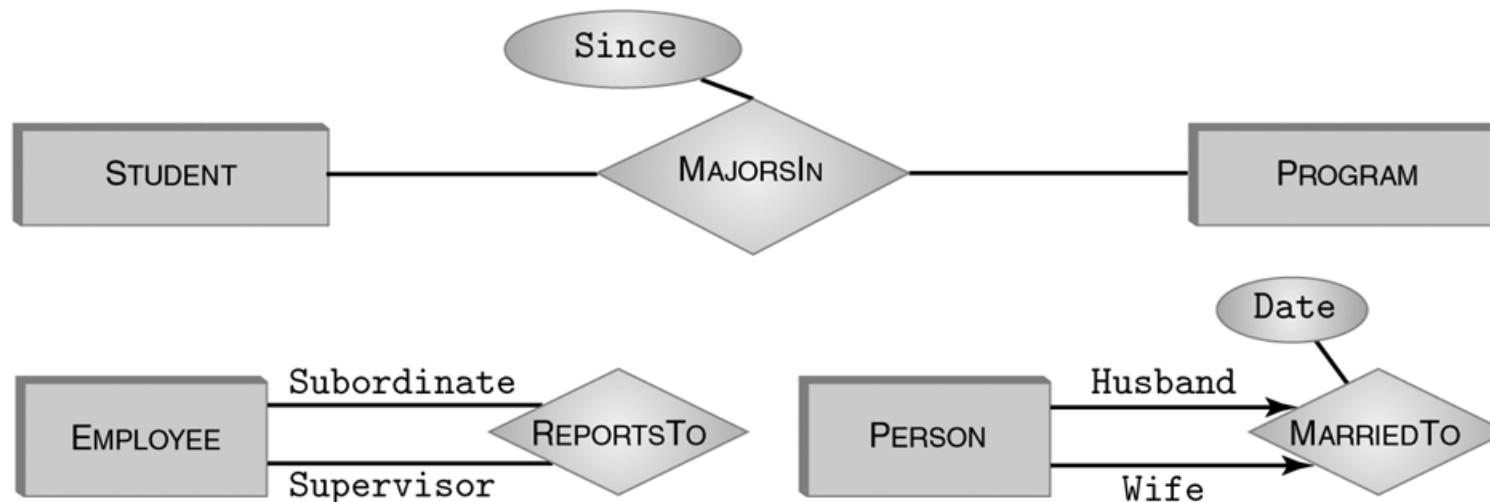
- One entity type participates **0 or 1** times in a relationship.
- Then a relation is not needed to represent the relationship.

Examples:

STUDENT can store Since and primary key of PROGRAM

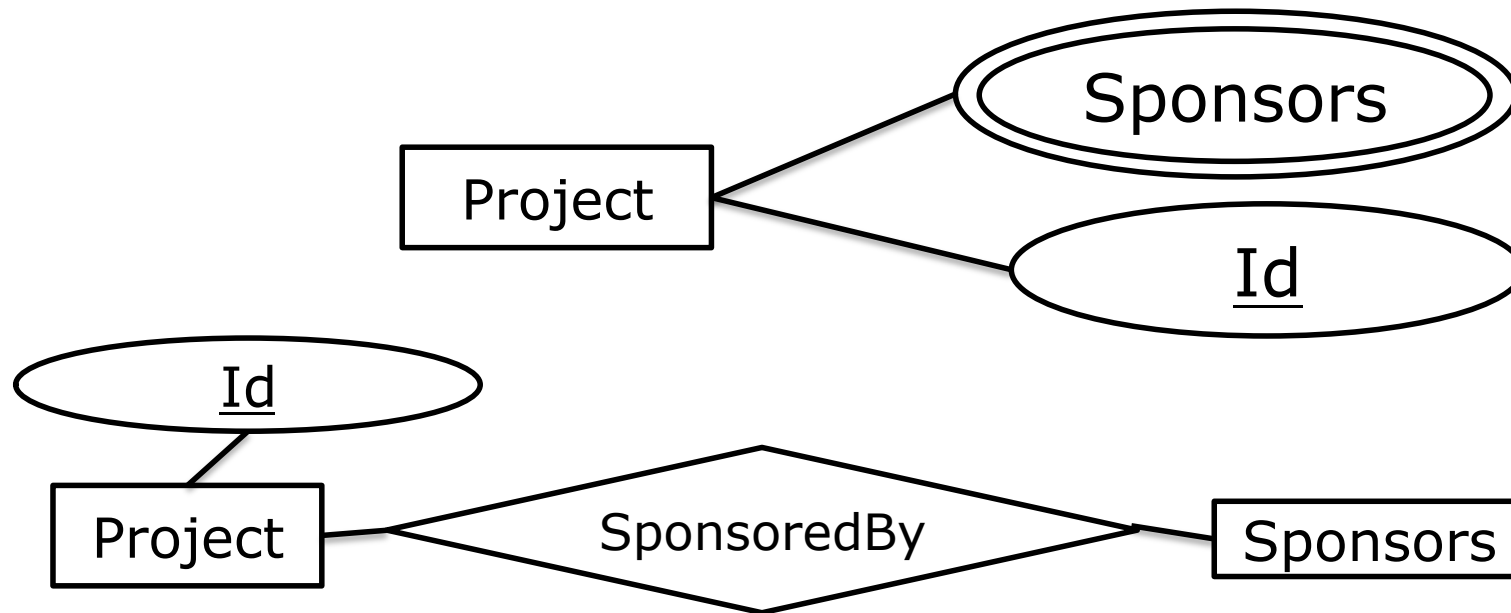
EMPLOYEE can store primary key for the Supervisor

PERSON can store Date and SSN for Husband or Wife, resp.



Set valued attributes

- A set valued attribute can be expressed as a relationship instead:



Each set-valued attribute becomes a relation.



IsA

Can be done in different ways. Examples:

- 1.** Treat like any other relationship.
- 2.** All relations use the *same* primary key
 - the key of the top entity type
- 3.** A single relation for the whole hierarchy with attribute(s) defining the entity type(s)
 - Binary (indicate all memberships), or
 - Identifier for the most specific entity type.



Translating weak entities

- The primary key contains
 - key attributes for the entity type
 - primary key attributes of the master entity types.
- Otherwise as other entity types.



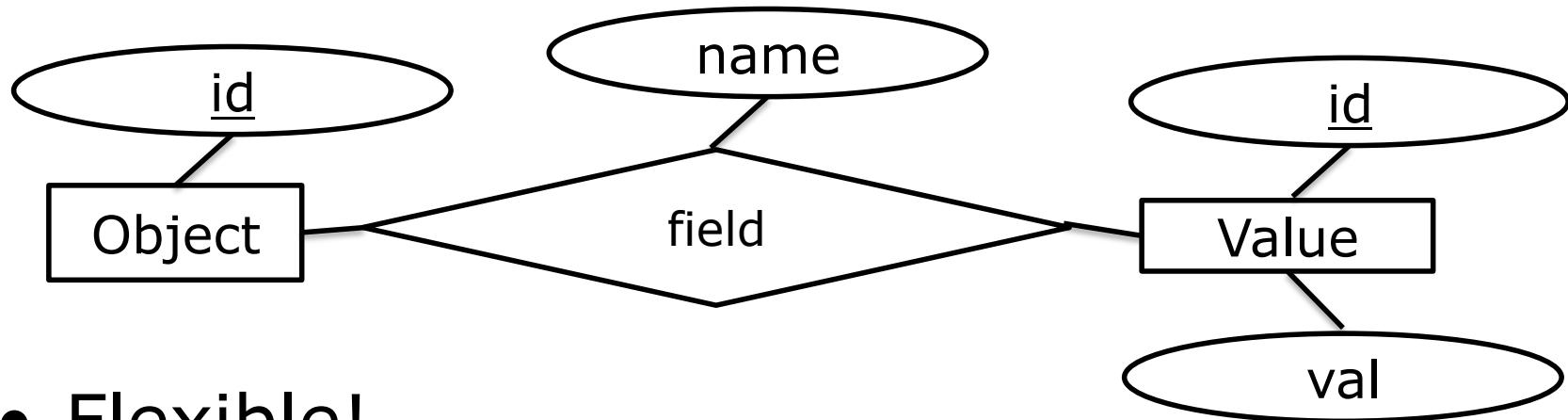
Participation constraints

- Translation to SQL constraints:
Exercise this afternoon!
- (Cannot be fully expressed using standard SQL...)



One E-R diagram to rule them all?

- Encode any relational information using:



- Flexible!
- Big downside: DBMS **lost** when it comes to type checking, indexing, fast query evaluation,...



Tools

- Tools such as MySQL Workbench modeling using E-R notation, and automatic translation to relation schemas.
- Ideally, should also support “reverse engineering” E-R model from relation schema.
 - Problem: This mapping is not unique.



Hand-in, part 1

- Available on course web page.
- Due in two weeks, September 20.
- We have covered the material to do most already.
 - The rest (normalization) next week.
- Whether you are lacking 0, 1, 2, or 3 group members, please write to Ninh, e-mail ndap, with your partial group information.
 - We start forming groups this afternoon.

