IS5102
Database Management Systems

Lecture 2: E-R Diagrams

Alexander Konovalov

alexander.konovalov@st-andrews.ac.uk

(with thanks to Susmit Sarkar)

2021

Physical Level

Describe how a data record (e.g. student) is stored

Logical Level

Describe the data and relationships between data

View Level

Describe selected aspects (**views**) of the data

Many views of same data possible

Also important to **hide** data in views (think security)

Analogous to types and values in programming languages

- Schema
  - The overall design of the database
  - **Physical** schema – database design at the physical
  - **Logical** schema – database design at the logical level level
  - Changes are infrequent
- Instance
  - Content of the database at a particular point in time
  - Changes may be frequent

Data Models

A collection of conceptual tools for describing

- Data
- Data relationships
- Data semantics
- Data constraints

Types of Data Models

Physical data model – geared towards implementation

Logical/conceptual data model – more abstract

Examples:

- **Entity-Relationship (E-R)** data model (mainly for database design)
- **Relational** model (lower level, later)
- Object-based data models (Object-oriented and Object-relational)
- Semistructured data model (XML)
- Other older models:
  - Network model
  - Hierarchical model

A **database** can be modeled as:

- a collection of **entities**,
- **relationship** among entities.

Entities and Entity Sets

- An **entity** is an object that exists and is distinguishable from other objects.
  - Example: specific person, company, event, plant
- Entities have **attributes**
  - Example: people have names and addresses
- An **entity set** is a set of entities of the same type that share the same properties.
  - Example: set of all persons, companies, trees, holidays

- An entity is represented by a set of **attributes**, that is descriptive properties possessed by all members of an entity set.

  Example:

  ```
  instructor = (ID, name, street, city, salary)
  course = (course_id, title, credits)
  ```

- **Domain** – the set of permitted values for each attribute

- Attribute types:
  - **Simple** and **composite** attributes
    - Example of a composite attribute: `address`
  - **Single-valued** and **multivalued** attributes
    - Example of a multivalued attribute: `phone_numbers`
  - **Derived** attributes
    - Can be computed from other attributes
    - Example: `age`, if given `date_of_birth`

- A **relationship** is an association among several entities
  Example:
  44553 (Student X) advisor 22222 (Instructor Y)
  student entity $\rightarrow$ relationship set $\rightarrow$ instructor entity

- A **relationship set** is a mathematical relation among $n \geq 2$ entities, each taken from corresponding entity sets

$$\{ (e_1, e_2, \ldots, e_n) \mid e_1 \in E_1, e_2 \in E_2, ..., e_n \in E_n \}$$

where $(e_1, e_2, ..., e_n)$ is a relationship
Example: $(44553, 22222) \in$ `advisor`

Attributes in Relationship Sets

- An attribute can also be property of a relationship set
- For instance, the `advisor` relationship set may have the attribute `date` which tracks when the student started being associated with the advisor

Degree of a Relationship Set

- **Binary** relationship
  - involves two entity sets (i.e. has **degree** two).
  - most relationship sets in a database system are binary.

- Relationships between more than two entity sets are less common, but also occur
  - Example: students work on research projects under the guidance of an instructor
    relationship `proj_guide` is a ternary relationship between instructor, student, and project

Participation Constraints

determined by the **minimum** number of times entity participates in relationship

– if zero, then **partial** participation

– if more than zero, then **total** participation

Cardinality Constraints

maximum number of times entity participates in relationship

Mapping Cardinality Constraints

- Express the **number** of entities to which another entity can be **associated** via a relationship set
- Most useful in describing binary relationship sets

- For a binary relationship set the mapping cardinality must be one of the following types:
  - **One to one**
  - **One to many**
  - **Many to one**
  - **Many to many**

Keys

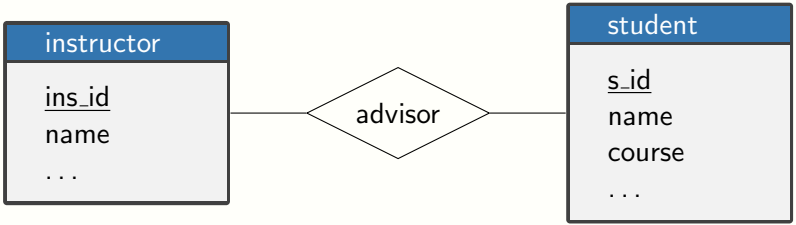
- A **super key** of an entity set is a set of one or more attributes whose values uniquely determine each entity.
- A **candidate key** of an entity set is a minimal super key
  - `ID` is candidate key of instructor
  - `course_id` is candidate key of course
- Although several candidate keys may exist, one of the candidate keys is selected to be the **primary key**.

Keys for Relationship Sets

- The **combination of primary keys** of the participating entity sets forms a super key of a relationship set.

  `(s_id, i_id)` is the super key of `advisor`

- Must consider the **mapping cardinality** of the relationship set when deciding what are the candidate keys



- Rectangles represent **entity sets**.
- Diamonds represent **relationship sets**.
- **Attributes** listed inside entity rectangle
- Underline indicates **primary key attributes**

We express cardinality constraints by drawing:

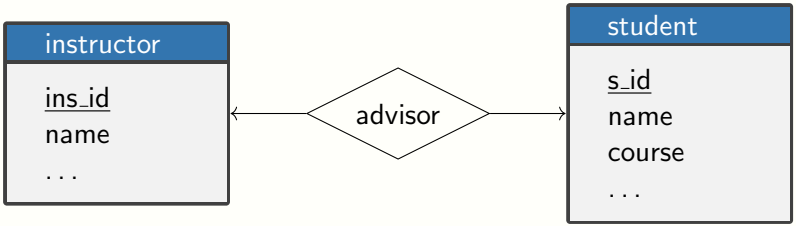
either a directed line ($\longrightarrow$), signifying **"one"**

or an undirected line ($\_\_$), signifying **"many"**

between the relationship set and the entity set

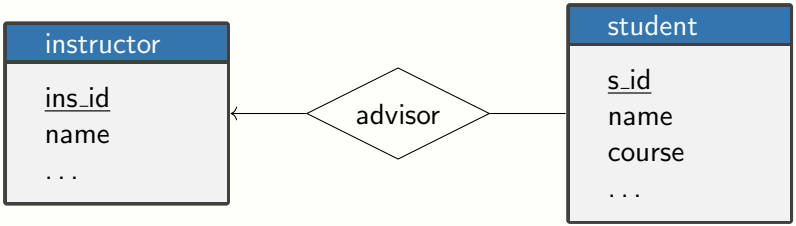
One-to-one relationship between an instructor and a student

- an instructor is associated with at most one student via advisor
- and a student is associated with at most one instructor via advisor



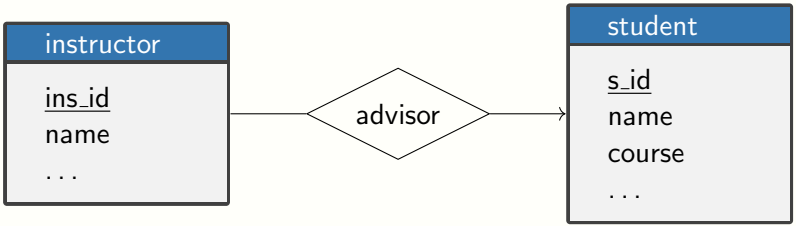
One-to-many relationship between an instructor and a student

- an instructor is associated with several (including 0) students via advisor
- a student is associated with at most one instructor via advisor



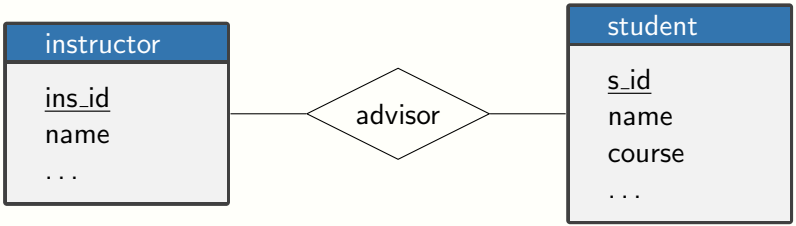
Many-to-one relationship between an instructor and a student

- an instructor is associated with at most one student via advisor,
- and a student is associated with several (including 0) instructors via advisor



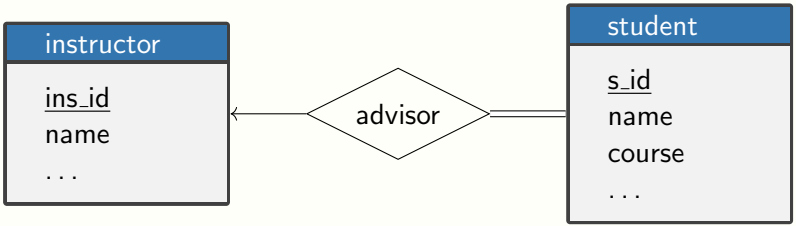
Many-to-many relationship between an instructor and a student

- An instructor is associated with several (possibly 0) students via advisor
- A student is associated with several (possibly 0) instructors via advisor



Participation of an Entity Set in a Relationship Set

- **Total participation** (indicated by double line): every entity in the entity set participates in at least one relationship in the relationship set
  Example: participation of student in advisor is total
  every student must have an associated advisor
- **Partial participation**: some entities may not participate in any relationship in the relationship set
  Example: participation of instructor in advisor is partial



- Data models:
  - Chapters 4-5, Database Design
  - Chapter 1, Database System Concepts
- E-R models:
  - Chapter 8, Database Design
  - Chapter 7, Database System Concepts