DATABASE SYSTEMS

CS – 355/ CE – 373

Instructor: Maria N. Samad

September 3rd, 2024

DATABASE DESIGN

- The task of creating a database application is a complex one, involving:
  - Design of the database schema
  - Design of the programs that access and update the data
  - Design of a security scheme to control access to data.

- The needs of the users play a central role in the design process.

DATABASE DESIGN PHASES

Several phases are required for a complete database design:

- User Requirements Specification: Involves interaction with the users.
- Conceptual Design: User requirements are translated into a conceptual schema of the database.
- Specification of functional requirements: Users describe the kinds of operations (or transactions) that will be performed on the data.

The output of these phases is a fully specified conceptual design of the database.

DATABASE DESIGN PHASES

- In the logical-design phase, the designer maps the high-level conceptual schema onto the implementation data model of the database system that will be used.
- The implementation data model is typically the relational data model.
- This step typically consists of mapping the conceptual schema defined using the entity-relationship model into a relation schema.

DATABASE DESIGN PHASES

- Finally, the designer uses the resulting system-specific database schema in the subsequent physical-design phase, in which the physical features of the database are specified.

- These features include the form of file organization and choice of index structures.

DATABASE DESIGN – PITFALLS

In designing a database schema, we must ensure that we avoid two major pitfalls:

- **Redundancy**: a bad design may result in repeat information.
- Redundant representation of information may lead to data inconsistency among the various copies of information
- **Incompleteness**: a bad design may make certain aspects of the enterprise difficult or impossible to model.

Avoiding bad designs is not enough. There may be a large number of good designs from which we must choose.

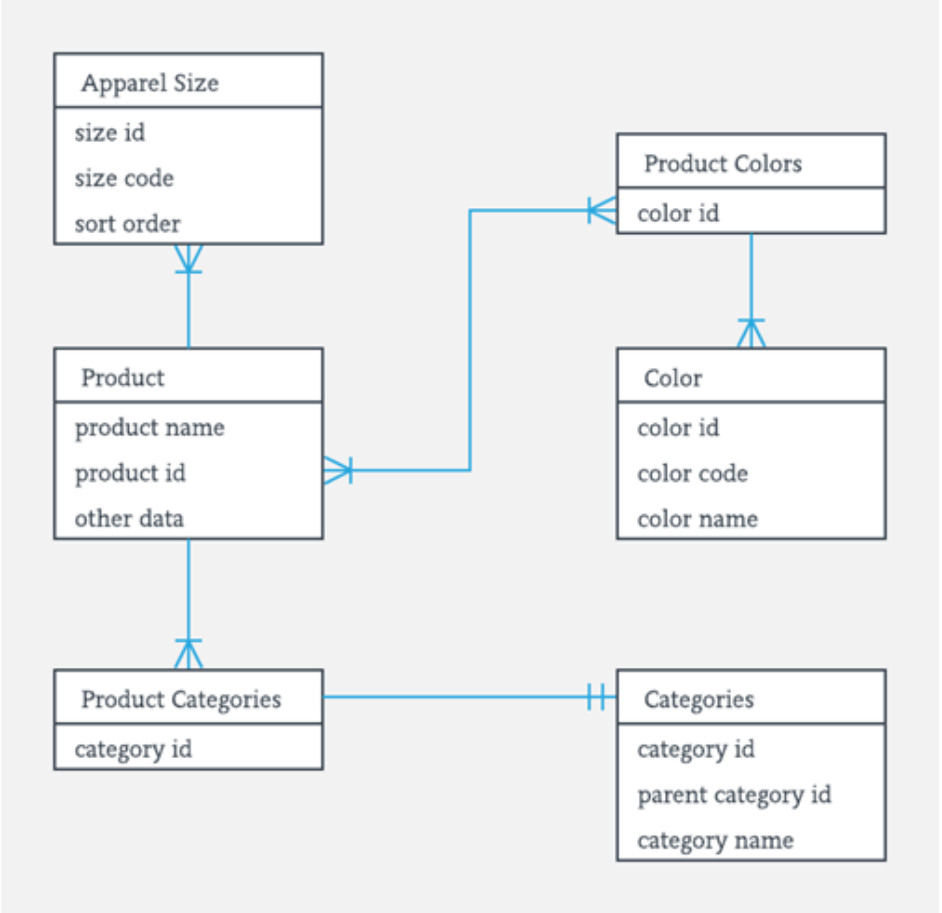
** Content from Dr. Faisal Alvi's slides **

DATA MODEL

- A collection of conceptual tools to describe:
  - ***Data***
  - ***Data Relationships***
  - ***Data Semantics/Rules***
  - ***Consistency Constraints***
- Provides a way to describe the overall design of a database at all the levels of abstraction
- The 4 main categories of data models are:
  - Relational Model
  - Entity-Relationship Model
  - Object-Based Data Model
  - Semistructured Data Model
- This chapter focuses on Design of Entity-Relationship Model

DATA MODEL – ENTITY-RELATIONSHIP MODEL

- Uses a collection of basic objects called ***entities*** and their ***relationships*** among each other
- An ***entity*** can represent any thing or object of the real world
- Depicted by E-R Diagrams



ENTITY-RELATIONSHIP MODEL

- This data model is used to represent the overall logical structure of a database
- Useful in mapping the meanings and interactions of real-world enterprises onto a conceptual schema
  - Typically includes only the main concepts and the main relationships of the different relations in a database
  - It gives a first-cut model, with insufficient detail to build an actual database
  - It hides the internal details of physical storage and targets on describing entities, datatypes, relationships and constraints.

ENTITY-RELATIONSHIP MODEL

- The E-R data model employs three basic concepts:
  - ***Entity Sets***
  - ***Relationship Sets***
  - ***Attributes***
- Many database design tools draw on concepts from the E-R model
- The associated diagrammatic representation, commonly used by database designers is the E-R Diagram
  - This is used to depict the E-R Model

ENTITY SETS

- An ***entity*** is a thing or any object of the real world, that is distinguishable from all other objects
  - Example:
    - A **person** in university – **instructor** and/or **student**
    - A **course** being offered in a semester
    - A **book** in a library
    - A **flight booking** for an upcoming journey
- Each entity has one or more properties, i.e. ***attributes***, and values of these attributes are used to uniquely identify a single entity
  - Example:
    - A **person** has a CNIC number that uniquely identifies him/her from others
    - A **book** has a catalog number associated with it in a library that differentiates it from others

ENTITY SETS

- An entity does not only contain these uniquely identified properties but others as well to describe the objects and their transactions completely
- Hence, it is represented by a ***set of attributes***, which are descriptive properties possessed by all members of an entity set
  - Example:
    - Instructor = {ID, Name, Department, Salary}
    - Course = {CourseID, Title, CreditHours, Department}
    - Book = {BookID, Title, Author, Genre}
- The subset of these attributes that uniquely identify each entity form the set of ***primary keys***

ENTITY SETS

- An ***entity set*** is a set of entities of the same type that share the same properties, i.e. have the same set of attributes
  - Example:
    - Set of all people in university, i.e. ***Instructor***, ***Student***
    - All the courses offered by a university, i.e. ***Course***
    - All the books in library, i.e. ***Book***
    - All the holidays of the current year, i.e. ***Holidays***
- Similar to relations or tables in the Relational Model

ENTITY SETS

- An ***entity set*** is represented in an E-R diagram by a rectangle, which is divided into two parts:
  - The first part contains the name of the entity set.
  - The second part contains the names of all the attributes of the entity set.
- Similar to Schema Diagrams of Relational Model

ENTITY SETS

- The E-R diagram below shows two entity sets ***instructor*** and ***student***.

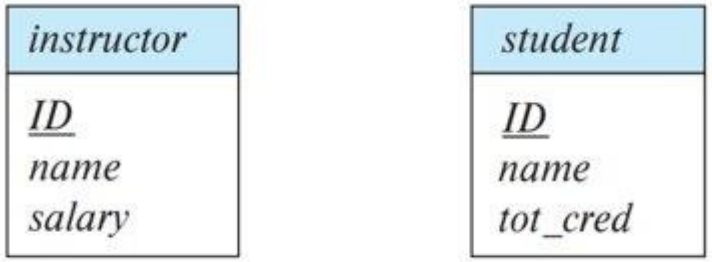


Figure 6.1 E-R diagram showing entity sets *instructor* and *student*.

- The attributes associated with ***instructor*** are *ID*, *name*, and *salary*.
- The attributes associated with ***student*** are *ID*, *name*, and *tot_cred*.
- Attributes that are part of the ***primary key*** are underlined

ENTITY SETS – EXAMPLES

- Activity Sheet
  - Attempt **Question 1**

ENTITY SETS – EXAMPLES

- Activity Sheet **Question 1** Solution:
  - ER Model Q1 Solution

RELATIONSHIP SETS

- A ***relationship*** is an association among several entities.
  - Example:
    - A relationship of an ***instructor*** with his/her ***student*** – **advising**
    - A relationship of ***customer*** with a ***bank*** – **account_holder**
    - Relationship(s) between ***persons*** and ***film*** – **acting**, **directing**, etc
    - A relationship between ***student*** and ***section*** – **takes** that course
- A ***relationship set*** will then represent a set of relationships of the same type.
  - Example:
    - All ***instructors*** who are **advising**, and their ***students***
    - All the **account holders** of a specific branch of a ***bank***
    - All the **actors** of a ***film***

RELATIONSHIP SETS

- The following diagram shows a relationship between an ***instructor*** and a ***student***

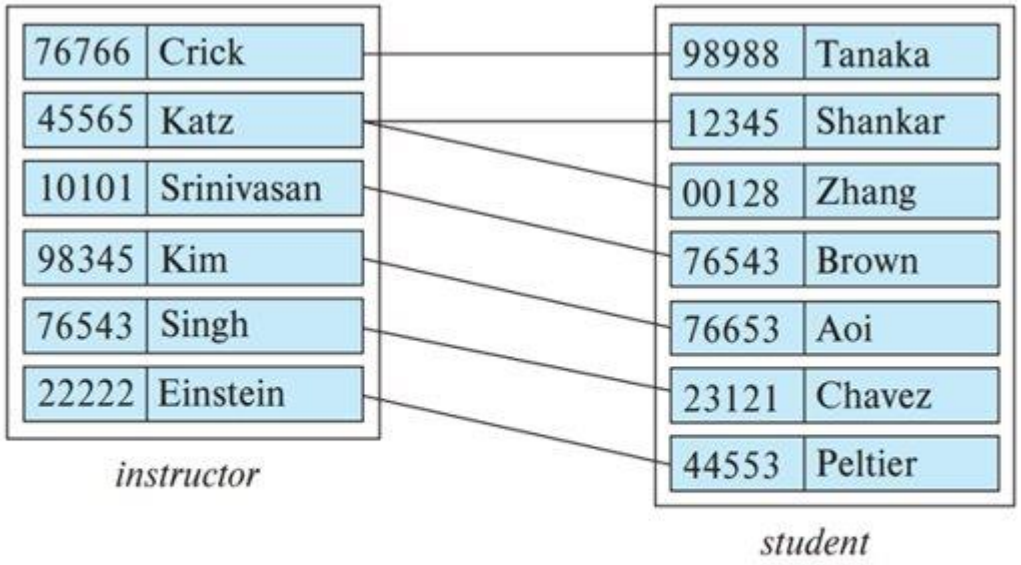


Figure 6.2 Relationship set *advisor* (only some attributes of *instructor* and *student* are shown).

RELATIONSHIP SETS

- A relationship instance in an E-R schema represents an ***association/relationship*** between the named ***entities*** in the real-world enterprise that is being modeled.
- A relationship set is represented in an E-R diagram by a diamond, which is linked via lines to a number of different entity sets (rectangles).

Figure 6.3 E-R diagram showing relationship set *advisor*.

RELATIONSHIP SETS – EXAMPLES

- Activity Sheet
  - Attempt **Question 2**

RELATIONSHIP SETS – EXAMPLES

- Activity Sheet **Question 2** Solution:
  - ER Model Q2 Solution

ROLES

- The function that an **entity** plays in a **relationship** is called that entity's ***role***.
- Roles are implicit and are not usually specified.
- However, they are useful when the meaning of a relationship needs clarification.

ROLES

- For example, in the ***instructor*** and ***student*** relationship:

- Roles are implicit, i.e. we can tell that either using Instructor ID or name, and Student's ID or name, we can define the relationship
- Let's say we have a relationship such that an instructor named **Katz** is an **advisor** to student named **Shankar**
- It is understandable whether we don't specify roles (as in above diagram) or explicitly write them (as in the diagram below)

ROLES

- However, when the entity sets of a relationship set are not distinct; i.e., the same entity set participates in a relationship set, in different roles, the roles play an important role
- In this type of relationship set, sometimes called a ***recursive relationship*** set, explicit role names are necessary to specify how an entity participates in a relationship instance.

ROLES

- Example:
  - A ***course*** records information about all the courses offered in the university
  - Each course is a prerequisite for another course, which means the same entity, i.e. ***course*** is related to another ***course***
  - This is then considered a recursive relationship

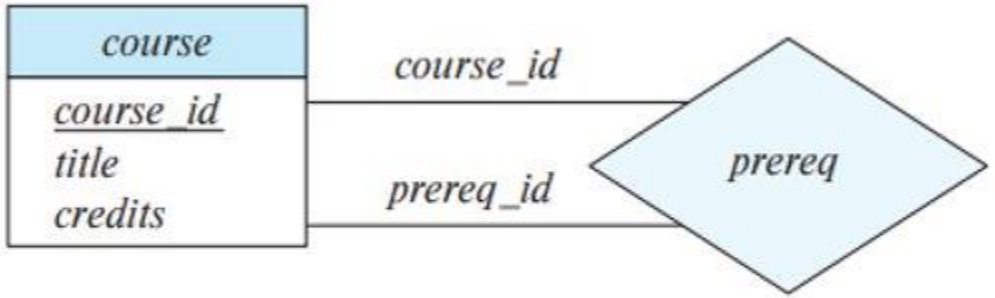


Figure 6.4 E-R diagram with role indicators.

ROLES – EXAMPLES

- Activity Sheet
  - Attempt **Question 3**

ROLES – EXAMPLES

- Activity Sheet **Question 3** Solution:
  - ER Model Q3 Solution

DESCRIPTIVE ATTRIBUTES

- A relationship may also have attributes called ***descriptive attributes***.
- For example, consider the relationship set **takes** which relates entity sets ***student*** and ***section***
- We may store a ***descriptive attribute***, **grade** with the relationship to record the grade that a ***student*** received in that ***section*** of course offering
- This depicts that a ***student*** gets a certain **grade** in the ***section*** of the course he was **taking**

DESCRIPTIVE ATTRIBUTES

- An attribute of a relationship set is represented by an undivided rectangle.
- This rectangle is linked with a dashed line to the diamond representing that relationship set.
- So for the ***student-section*** relationship example from the previous slide, the E-R diagram would be depicted as follows:

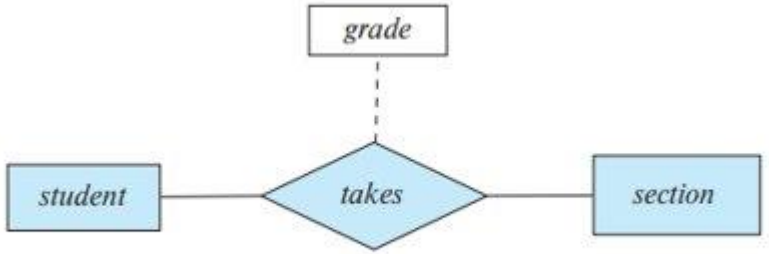


Figure 6.5 E-R diagram with an attribute attached to a relationship set.

- A relationship set may have multiple descriptive attributes

DESCRIPTIVE ATTRIBUTES – EXAMPLES

- Activity Sheet
  - Attempt **Question 4**

DESCRIPTIVE ATTRIBUTES – EXAMPLES

- Activity Sheet **Question 4** Solution:
  - ER Model Q4 Solution

DEGREE OF RELATIONSHIPS

- The relationship sets **advisor** and **takes** provide examples of a binary relationship set—that is, one that involves two entity sets.
- Most of the relationship sets in a database system are ***binary***.
- Occasionally, however, relationship sets involve more than two entity sets.
- The number of entity sets that participate in a relationship set is the ***degree of the relationship set***.
- A ***binary*** relationship set is of degree 2; a ***ternary*** relationship set is of degree 3.

DEGREE OF RELATIONSHIPS

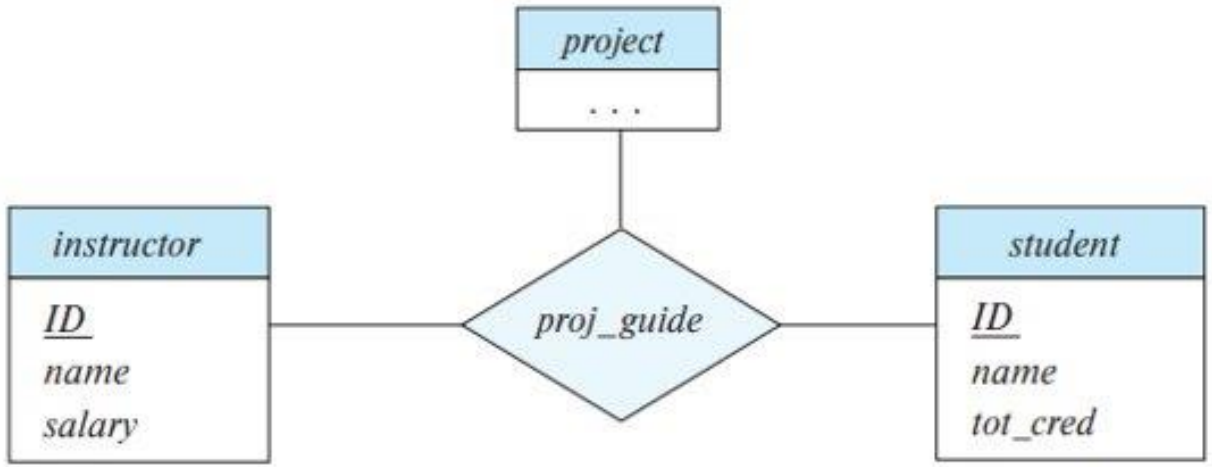


Figure 6.6 E-R diagram with a ternary relationship *proj_guide*.

ATTRIBUTES

- For each attribute, there is a set of permitted values, called the ***domain***, or ***value set***, of that attribute
- Each attribute instance must lie within its ***domain*** for the database to be correctly defined and become functional
- For example, the ***domain*** of attribute, **course_id** might be the set of all text strings of a certain length.
- Or, the ***value set*** of attribute, **semester** might be strings from the set {Fall, Winter, Spring, Summer}

ATTRIBUTES

- As each ***entity set*** contains several attributes, then each ***entity*** can be described by a set of **(attribute, data value)** pairs, with one pair for each attribute in the ***entity***
- For example, a particular ***instructor entity*** may be described by the set:
  - {(**ID**, 76766), (**name**, Crick), (**dept_name**, Biology), (**salary**, 72000)}
- This is where the integration of the conceptual schema with the actual enterprise starts being modeled

ATTRIBUTES

- An ***attribute***, as used in the E-R model, can be characterized by the following attribute types:
  - ***Simple*** and ***Composite*** attributes
  - ***Single-valued*** and ***Multivalued*** attributes
  - ***Derived*** attributes

SIMPLE AND COMPOSITE ATTRIBUTES

- ***Simple*** attributes are **atomic** – they cannot be divided into subparts.
  - Example:
    - Instructor ID is a 5-digit integer. It cannot be divided further, hence it's atomic, and thus considered a ***simple*** attribute
    - Department Name is a string, which is also representing a singular value, thus defined as a ***simple*** attribute
- ***Composite*** attributes can be divided into subparts.
  - Example:
    - Instructor Name could be structured as ***composite*** attribute, if the name consisted of first name, middle name and last name all together
    - Phone number can also be defined as a ***composite*** attribute, by specifying country code, city code, and/or extension separately
  - Specified as indented values in the E-R diagram

COMPOSITE ATTRIBUTES

- A ***composite*** attribute may also appear as a hierarchy.
- The further attributes of each ***composite*** attribute are called the ***component attributes***
- Example: An address can be a ***composite*** attribute, containing street, city, state, and zip code, in which street can become the ***component*** attribute that can be further divided into street number, street name and apartment number, forming a hierarchical structure

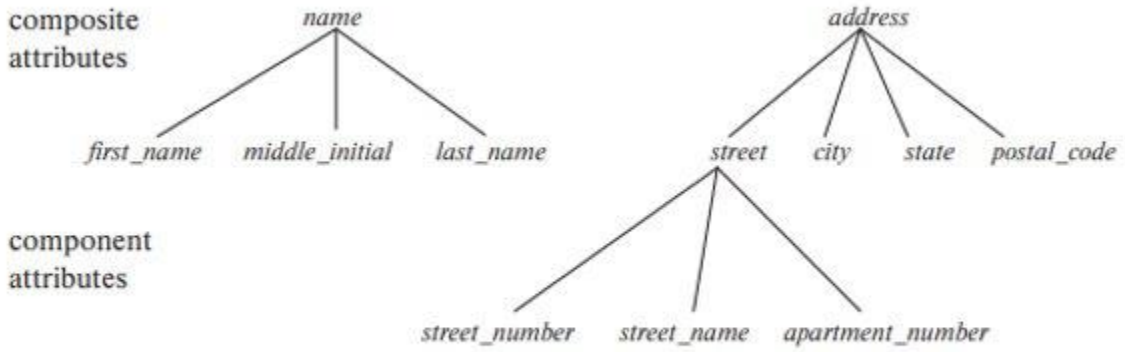
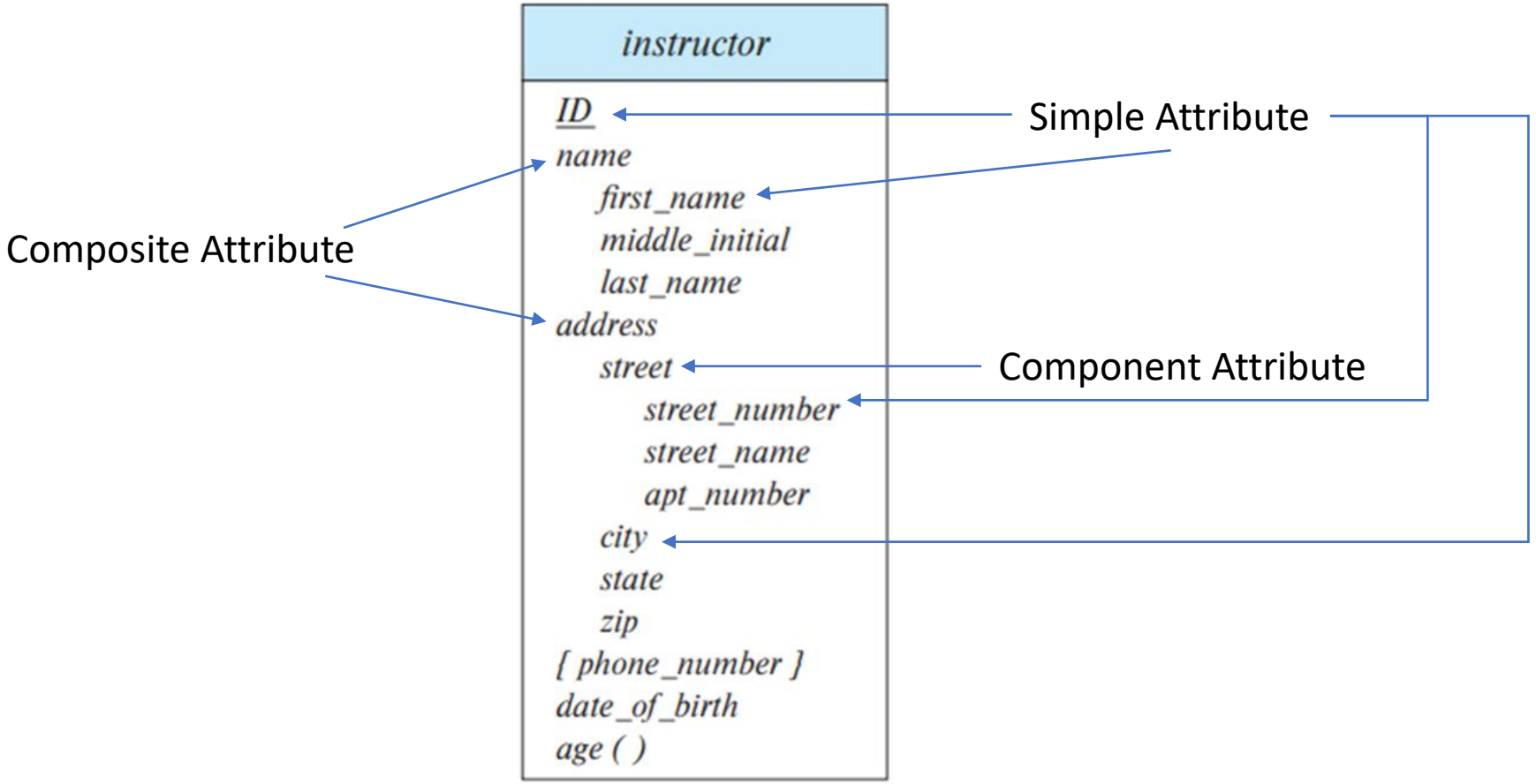


Figure 6.7 Composite attributes instructor *name* and *address*.

SIMPLE & COMPOSITE ATTRIBUTES IN E-R DIAGRAMS



E-R diagram with composite, component, and simple attributes

SIMPLE & COMPOSITE ATTRIBUTES – EXAMPLES

- Activity Sheet:
  - Attempt **Question 5**

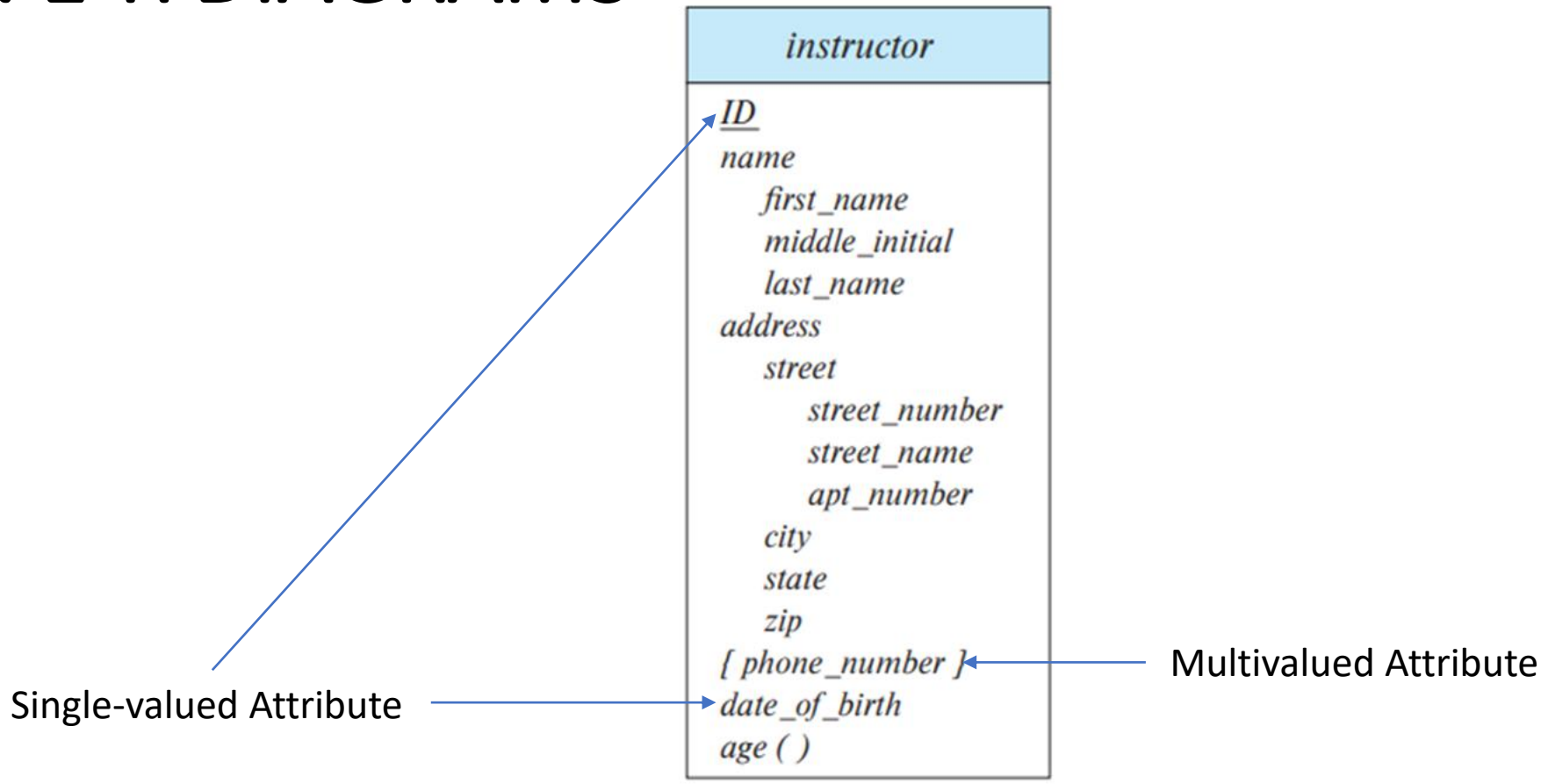
SIMPLE & COMPOSITE ATTRIBUTES – EXAMPLES

- Activity Sheet **Question 5** Solution:
  - **ER Model Q5 Solution**

SINGLE-VALUED AND MULTIVALUED ATTRIBUTES

- ***Single-valued*** attributes have a single value for a particular entity.
  - Example:
    - Student ID for a specific ***student*** entity refers to only one Student
    - Instructor's Joining Date is also single-valued because there can only be one day when they started, and will remain unchanged
- ***Multivalued*** attributes, on the other hand, may have zero, one or several values.
  - Example:
    - Phone number attribute for a person – they can have one or more phone numbers, or they can change their phone numbers
    - An employee can have multiple dependents for their health insurance policies, including their spouses, children, etc
  - Represented in braces in E-R diagrams. For example, {phone_number}, {dependents}, etc

SINGLE-VALUED AND MULTIVALUED ATTRIBUTES IN E-R DIAGRAMS



E-R diagram with multivalued attributes.

SINGLE-VALUED & MULTIVALUED ATTRIBUTES – EXAMPLES

- Activity Sheet:
  - Attempt **Question 6**

SINGLE-VALUED & MULTIVALUED ATTRIBUTES – EXAMPLES

- Activity Sheet **Question 6** Solution:
  - ER Model Q6 Solution

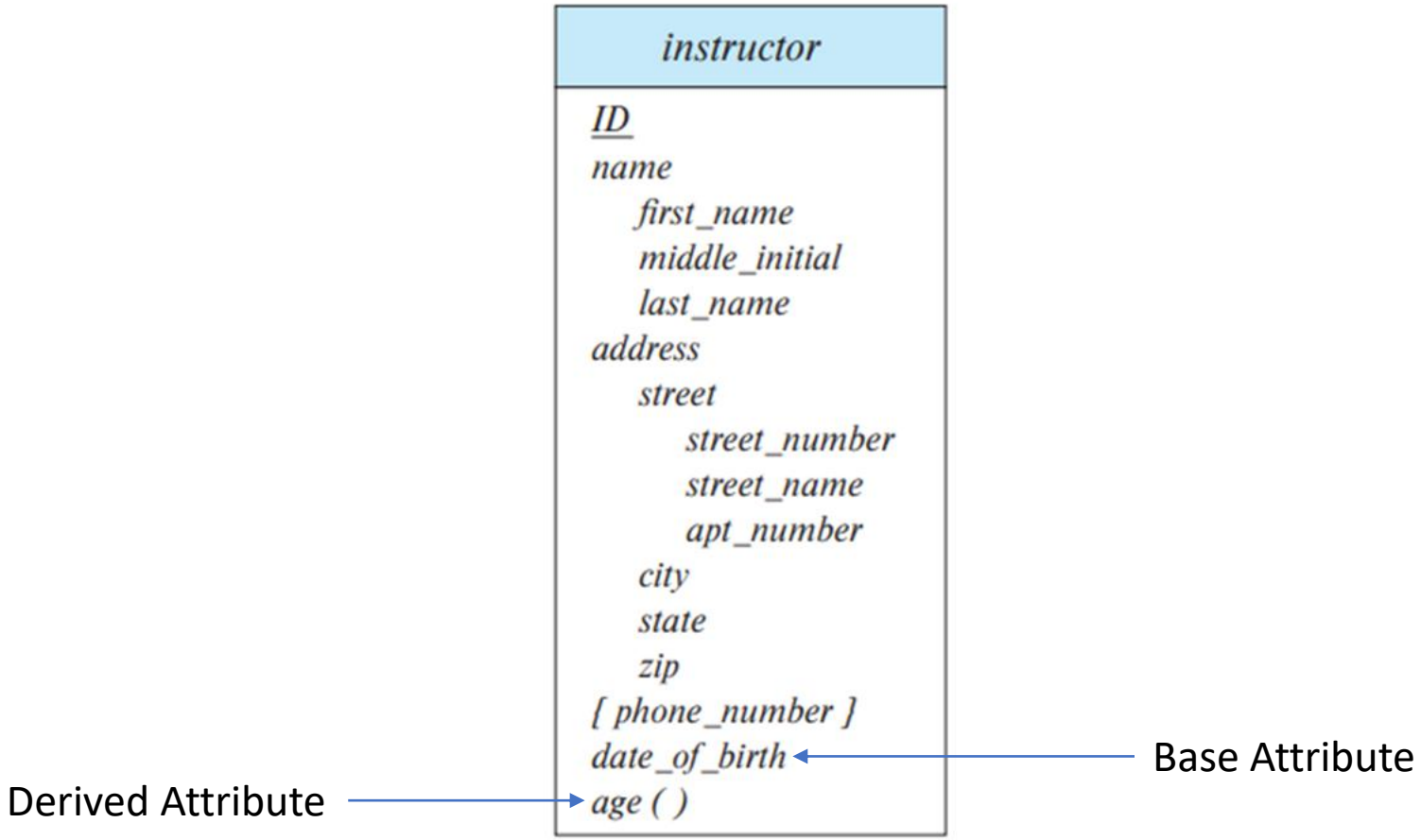
DERIVED ATTRIBUTES

- An attribute that does not exist in the entity set
- The value, however, can be derived from the other related attributes.
- Such attributes are defined as the ***derived*** attributes
- The attribute used for such a computation may be called a ***base attribute*** or a ***stored attribute***.
- The ***derived attribute*** may not be stored but computed when needed.
- Similar to functions in a program

DERIVED ATTRIBUTES

- Example:
  - Each ***instructor entity*** has an attribute which represents the student being advised. We can derive an attribute called **students_advised**, which represents the number of students advised by the instructor. This can be done by counting the number of ***student entities*** associated with that particular ***instructor***
  - If a student has an attribute, **DOB**, a derived attribute, i.e. **age** can be calculated by comparing that student's date of birth with the current date

BASE & DERIVED ATTRIBUTES IN E-R DIAGRAMS



E-R diagram with derived attribute

BASE & DERIVED ATTRIBUTES – EXAMPLES

- Activity Sheet:
  - Attempt **Question 7**

BASE & DERIVED ATTRIBUTES – EXAMPLES

- Activity Sheet **Question 7** Solution:
  - ER Model Q7 Solution

NULL VALUES

- An attribute may have a ***null value***.
- A null value indicates that either the value is:
  - Not applicable
  - It is missing.
- An example of a “not applicable” value is a missing middle name. Sometimes a person may not have a middle name.
- An example of a “missing” value is a null value of ID.
- A primary key must never have a null value

CARDINALITIES

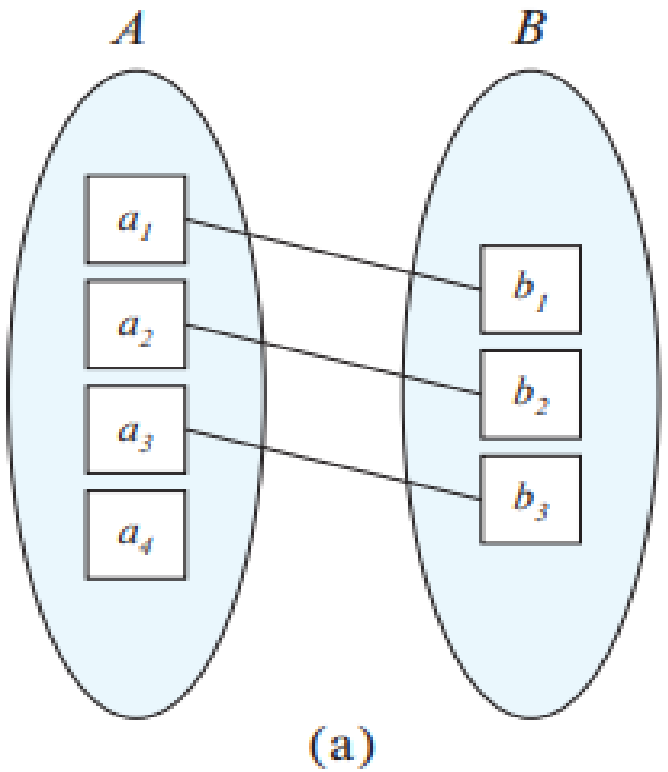
- ***Mapping cardinalities***, or ***cardinality ratios***, express the number of entities to which another entity can be associated via a relationship set.
- Mapping cardinalities are most useful in describing binary relationship sets
- They can also contribute to the description of relationship sets that involve more than two entity sets.

MAPPING CARDINALITIES

- For a binary relationship ***R*** between entity sets ***A*** and ***B***, the mapping cardinality can be one of the following:
  - One-to-one
  - One-to-many
  - Many-to-one
  - Many-to-many

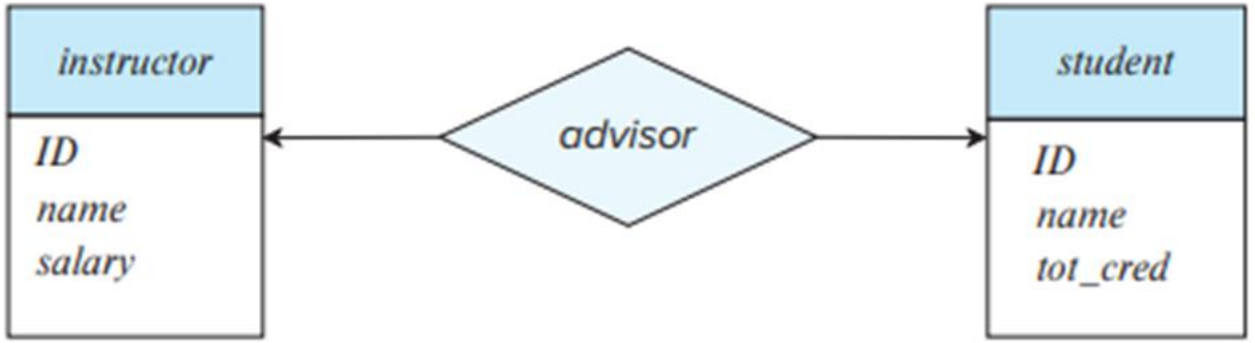
MAPPING CARDINALITIES – ONE-TO-ONE

- An entity in ***A*** is associated with <u>at most one</u> entity in ***B***, and an entity in ***B*** is associated with <u>at most one</u> entity in ***A***.



MAPPING CARDINALITIES – ONE-TO-ONE

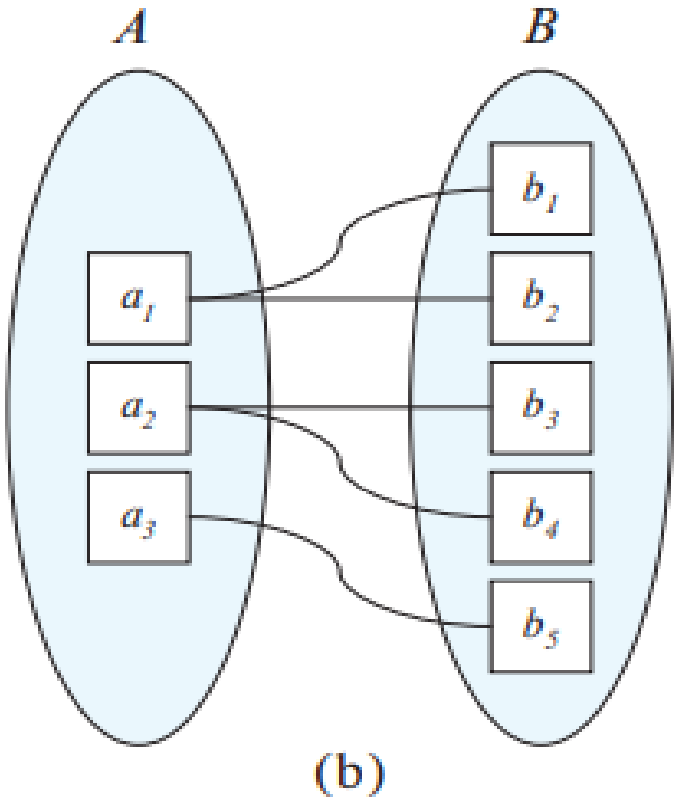
- Examples:
  - John Doe has a British passport
  - Every employee is entitled for one company car
  - Alex Johnson's locker number is 42
- ERD Representation of One-to-One cardinality:
  - For example, Ms. Maria Samad is the faculty advisor of Abdul Ahad



(a) One-to-one

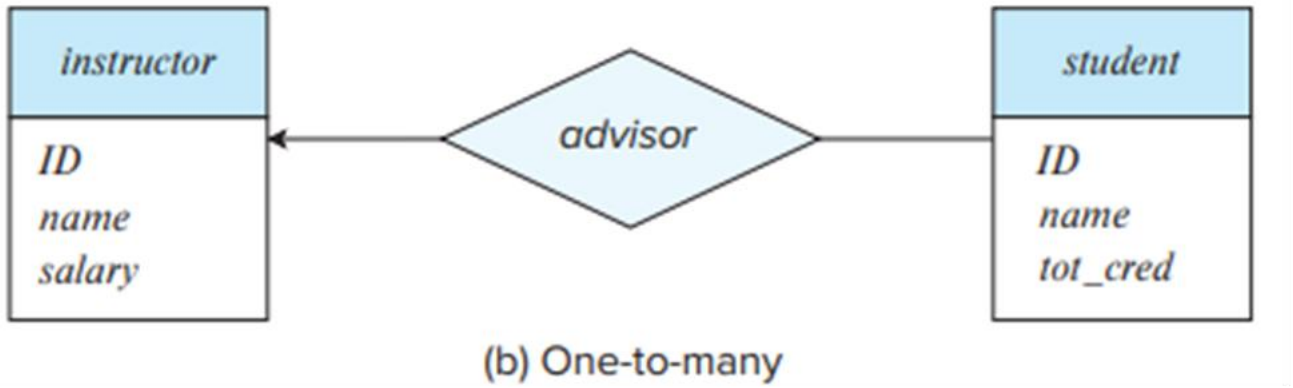
MAPPING CARDINALITIES – ONE-TO-MANY

- An entity in ***A*** is associated with <u>any number (zero or more)</u> of entities in ***B***. An entity in ***B***, however, can be associated with <u>at most one</u> entity in ***A***.



MAPPING CARDINALITIES – ONE-TO-MANY

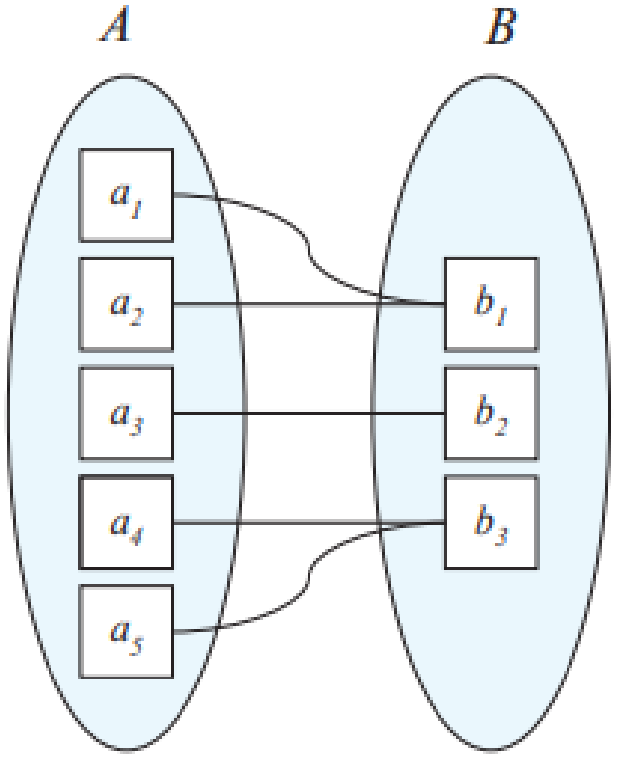
- Examples:
  - JK Rowling is the author of Harry Potter series
  - Jane Smith placed 3 orders on Amazon within a week
  - Marketing department has 10 employees
- ERD Representation of One-to-Many cardinality:
  - For example, Ms. Maria Samad is assigned 19 advisees this year



(b) One-to-many

MAPPING CARDINALITIES – MANY-TO-ONE

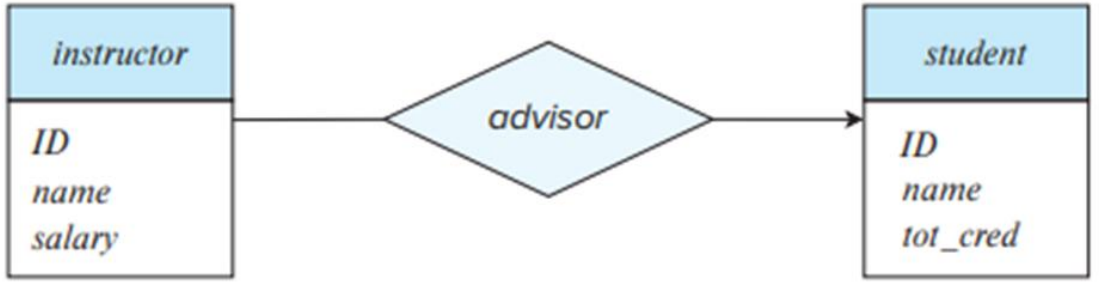
- An entity in ***A*** is associated with <u>at most one</u> entity in ***B***. An entity in ***B***, however, can be associated with <u>any number (zero or more)</u> of entities in ***A***.



(a)

MAPPING CARDINALITIES – MANY-TO-ONE

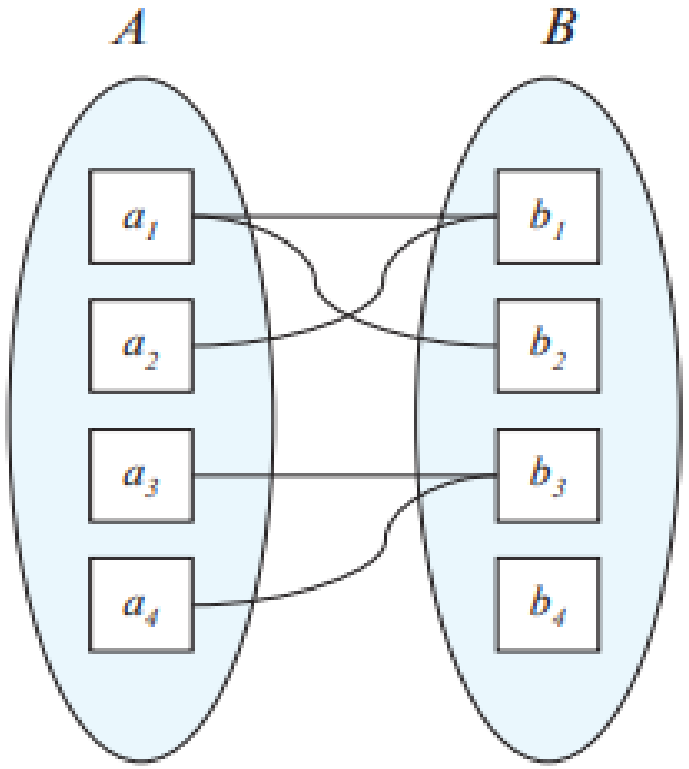
- Examples:
  - Alice, Bob and Charlie are enrolled in Lincoln High School
  - Tesla Powerwall, Tesla Model S, Tesla Cybertruck and Tesla Model X are produced by Tesla
  - There are 5 copies of DBS textbook available in HU library
- ERD Representation of Many-to-One cardinality:
  - For example, According to university policy, each instructor is supposed to advise only one student at a time, however, every student gets at least 2 faculty advisors



(c) Many-to-one

MAPPING CARDINALITIES – MANY-TO-MANY

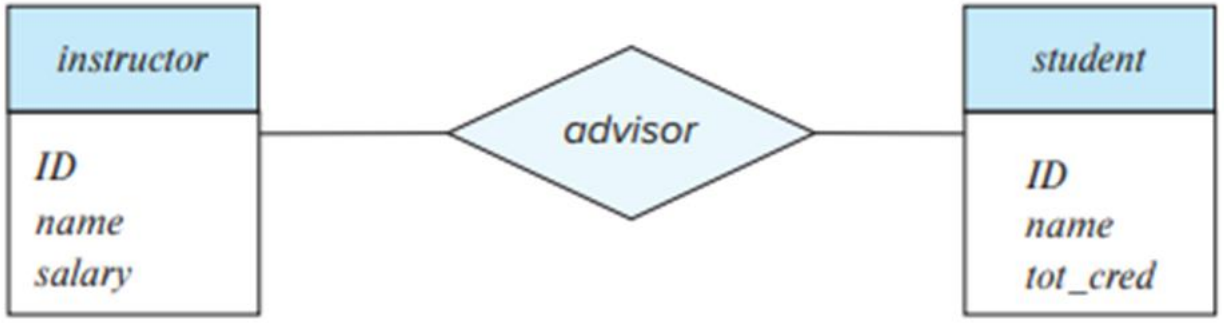
- An entity in ***A*** is associated with <u>any number (zero or more)</u> of entities in ***B***, and an entity in ***B*** is associated with <u>any number (zero or more)</u> of entities in ***A***.



(b)

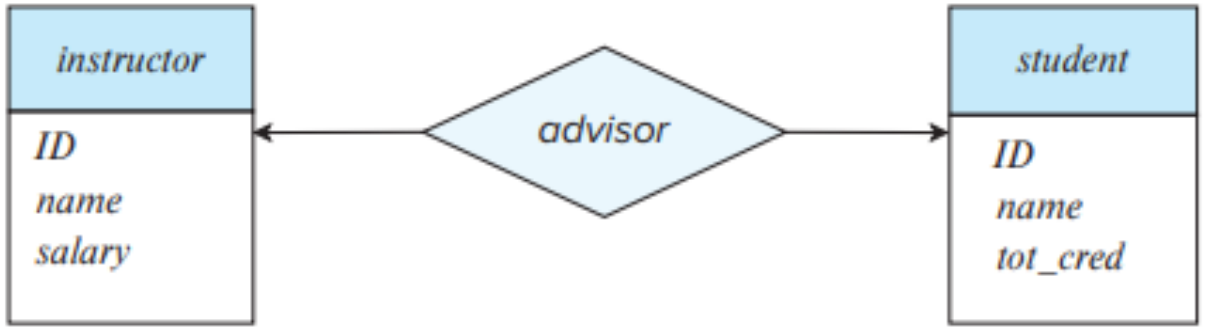
MAPPING CARDINALITIES – MANY-TO-MANY

- Examples:
  - Each CS Sophomore student is enrolled in Database Systems, Object-Oriented Programming and Digital Logic Design courses
  - Dr. Smith treats Jane and Bob
  - Best Buy sells different models of laptops
- ERD Representation of Many-to-Many cardinality:
  - For example, The CS department instructors are always guiding the students

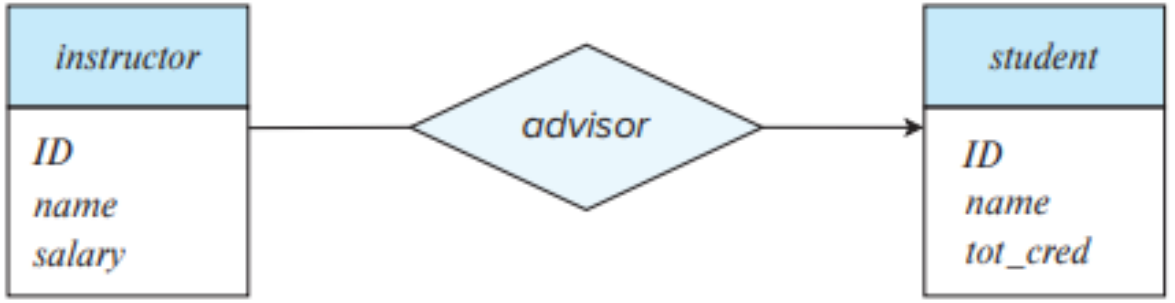


(d) Many-to-many

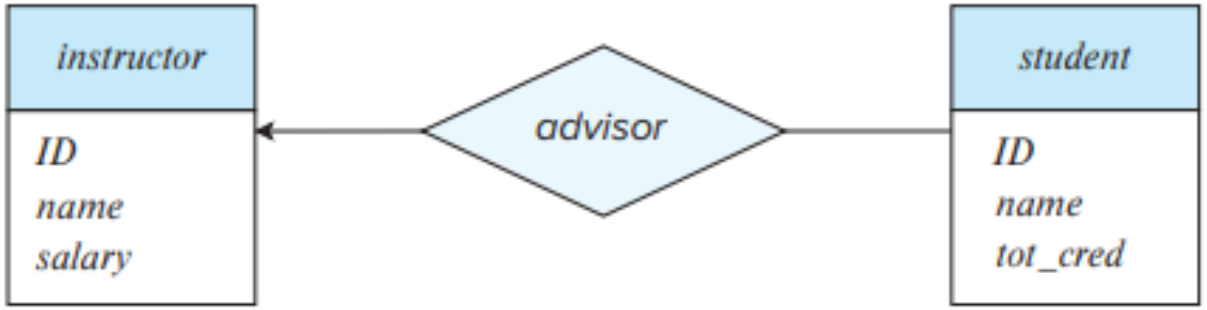
ERD REPRESENTATION OF ALL CARDINALITIES



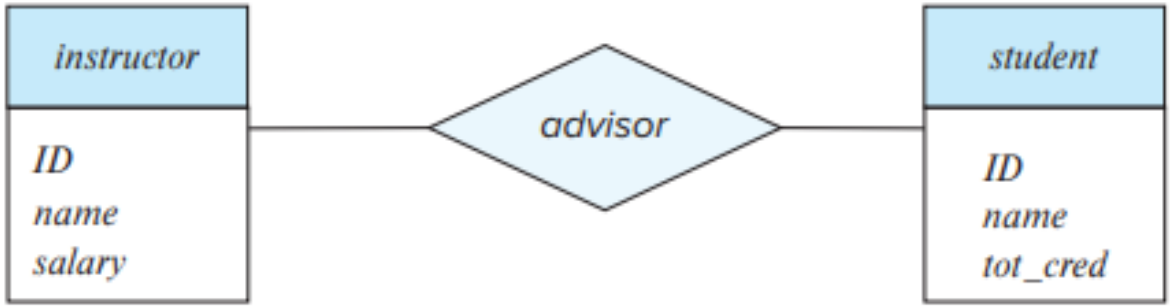
(a) One-to-one



(c) Many-to-one



(b) One-to-many



(d) Many-to-many

MAPPING CARDINALITIES – EXAMPLES

- Activity Sheet:
  - Attempt **Question 8**

MAPPING CARDINALITIES – EXAMPLES

- Activity Sheet **Question 8** Solution:
  - ER Model Q8 Solution