CS254: Database Management Systems (DBMS)

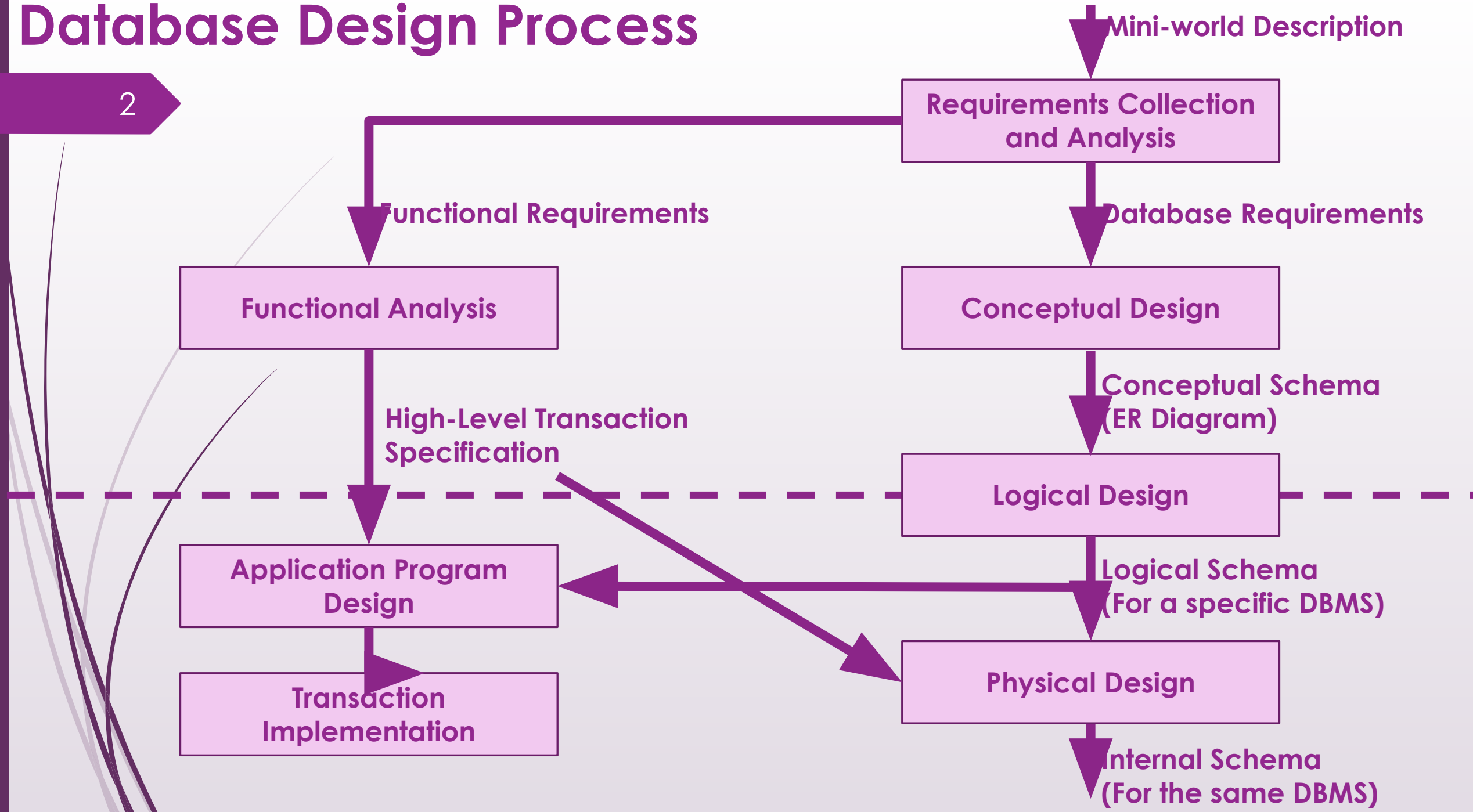
Lec04

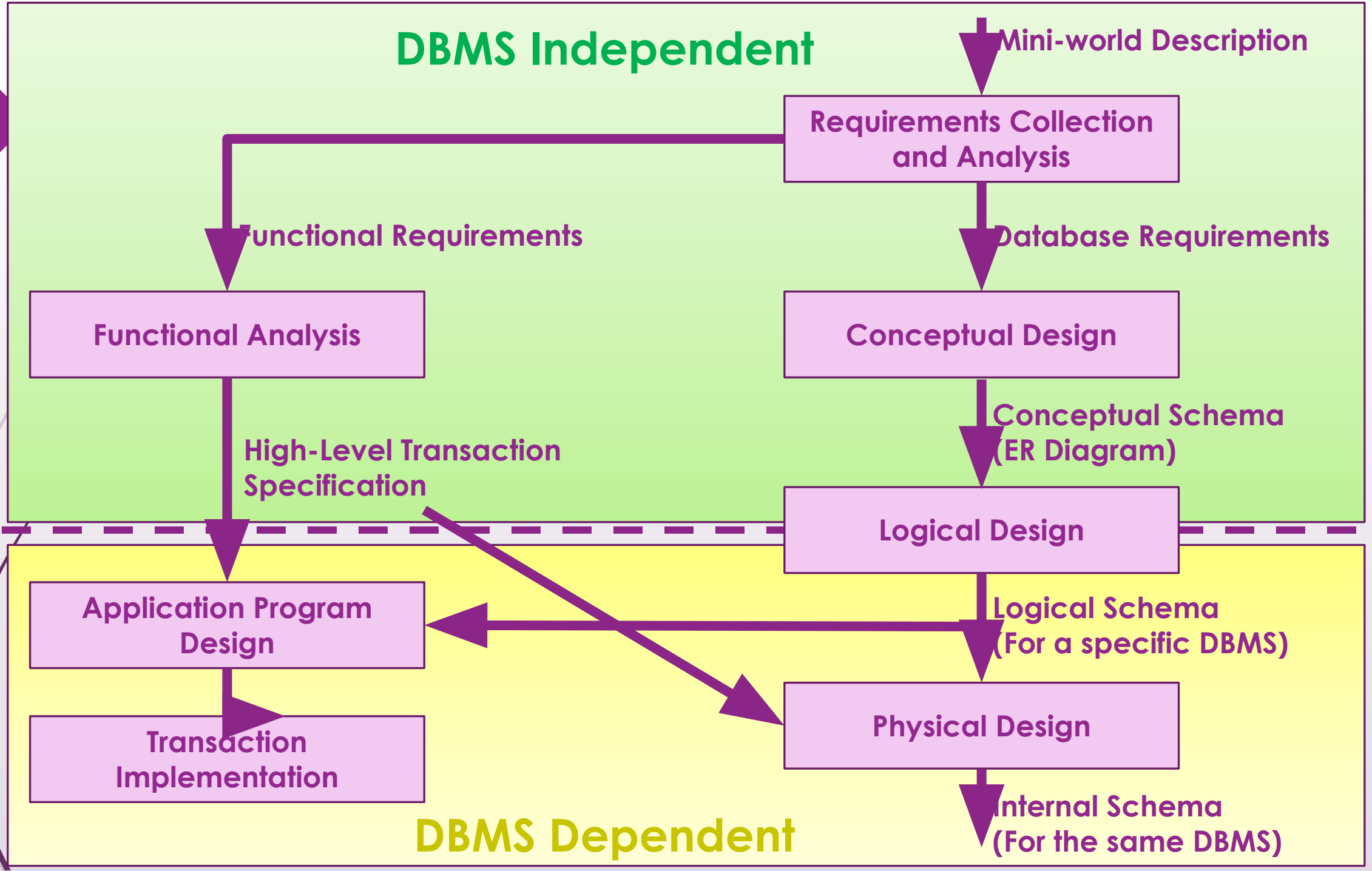
Suyash Kandele

suyashk@iitbhilai.ac.in

22 February 2019

Database Design Process




DBMS Independent
Mini-world Description
Requirements Collection and Analysis
Functional Requirements
Database Requirements
Functional Analysis
Conceptual Design
Conceptual Schema (ER Diagram)
High-Level Transaction Specification
Logical Design
Application Program Design
Logical Schema (For a specific DBMS)
Transaction Implementation
Physical Design
Internal Schema (For the same DBMS)
DBMS Dependent

Phases of Database Design

- Requirements collection and analysis
  - Database Requirements
  - Functional Requirements
- Conceptual Design & Functional Analysis
  - Create a conceptual schema
  - High level transaction specification corresponding to operations on database
- Logical Design
  - Map conceptual database schema to implementation database schema
- Physical Design
  - Internal storage structures and file organizations are specified
- Application Program Design & Transaction Implementation
  - In parallel to physical design

Example: Company Database

- The company is organized into Departments.
- Each department has a name, number, and an employee who manages the department.
- We keep track of the start date if the department has a manager.
- A department may have several locations.
- Each department controls a number of Projects.
- Each project has a name, number, and is located at a single location.
- We store each Employees social security number, name, addresses, salary, gender, and birth date.
- Each employee works for one department but may work on several projects.
- We keep track of the number of hours per week that an employee currently works on each project.
- We also keep track of the direct supervisor of each employee.
- Each employee may have a number of Dependents.
- For each dependent, we keep their name, gender, birthdate, and relationship to the employee.

Entities and Attributes

Entities

- Entities are specific objects or things in the mini-world that are represented in the database; for example the Employee John Smith, the Research Department, the Database Project.

Attributes

- Attributes are properties used to describe an entity; for example, an Employee entity may have a Name, ID, Address, Gender, BirthDate.

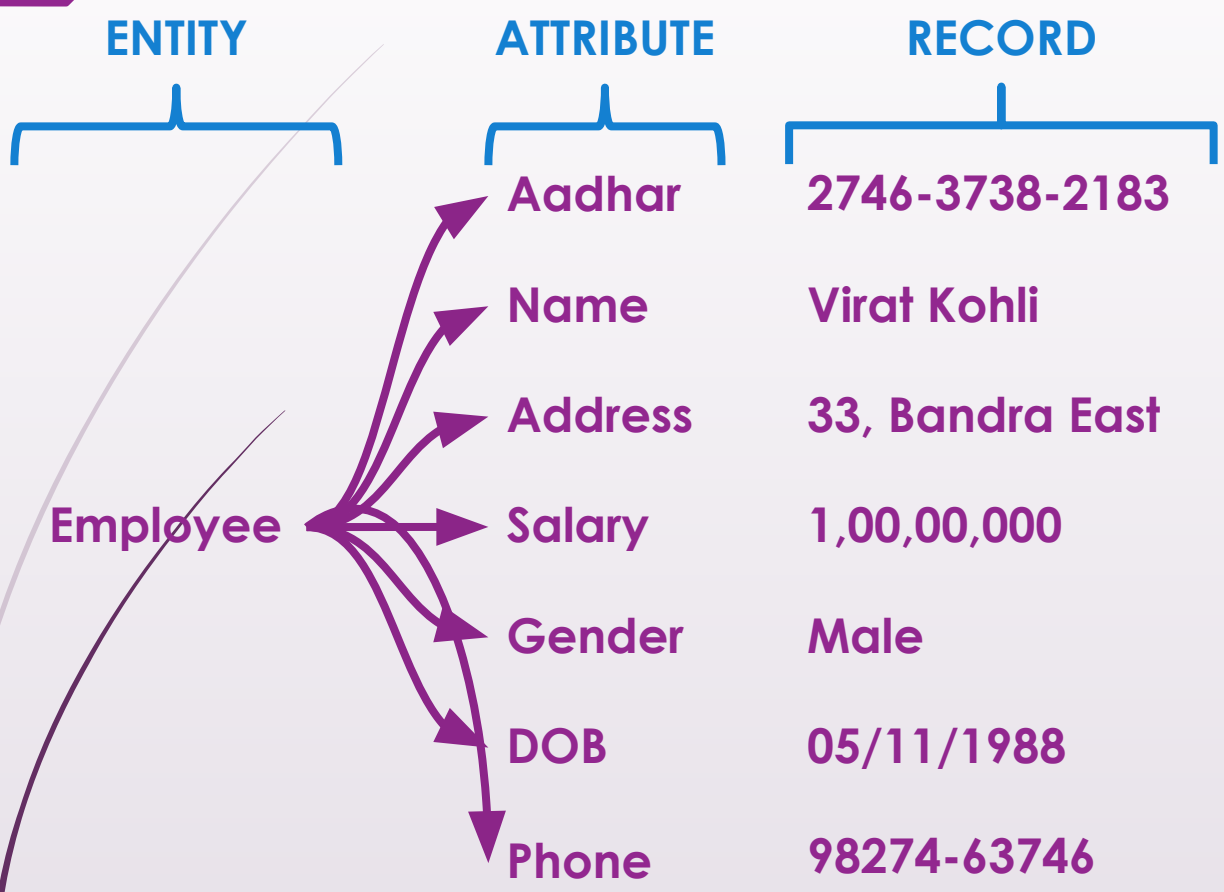
- A specific entity will have a value for each of its attributes.

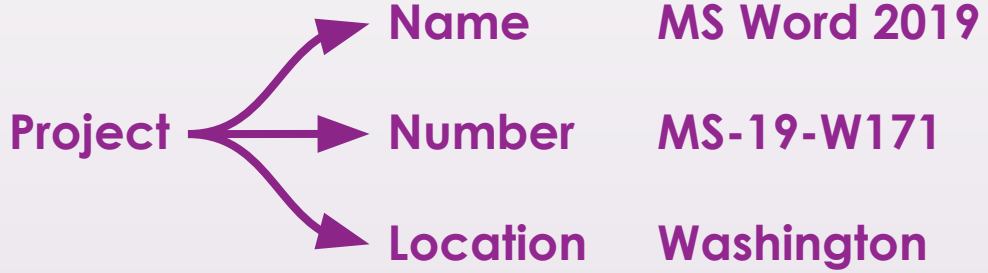
Exercise: IDENTIFY ALL ENTITIES

Requirements for the Company Database

- The company is organized into Departments.
- Each department has a name, number, and an employee who manages the department.
- We keep track of the start date if the department has a manager.
- A department may have several locations.
- Each department controls a number of Projects.
- Each project has a name, number, and is located at a single location.
- We store each Employees Aadhar number, name, addresses, salary, gender, and birth date.
- Each employee works for one department but may work on several projects.
- We keep track of the number of hours per week that an employee currently works on each project.
- We also keep track of the direct supervisor of each employee.
- Each employee may have a number of Dependents.
- For each dependent, we keep their name, gender, birthdate, and relationship to the employee.

Entities and Attributes





Entity Types

- Entities with the same basic attributes are grouped or typed into an entity type.
  - For example,
    - ***Company*** entity type.
    - ***Department*** entity type.
    - ***Employee*** entity type
    - ***Project*** entity type.

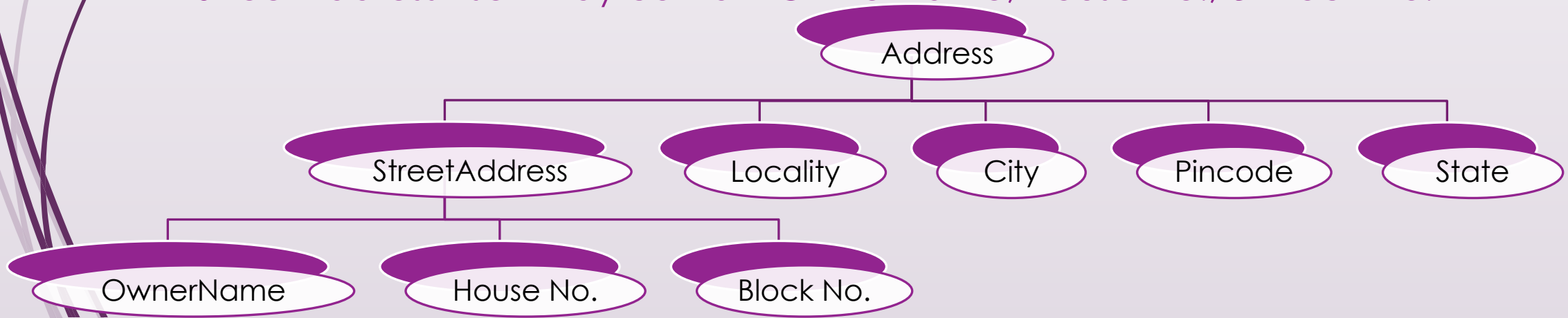
Types of Attributes (1)

Simple

- Each entity has a single atomic value for the attribute.
  - For example ID, CourseNo.

Composite

- The attribute may be composed of several components.
  - For example Address contains StreetAddress, Locality, City, Pincode and State.
  - StreetAddress itself may contain OwnerName, House No., & Block No.



Types of Attributes (2)

Single-valued

- An Entity has only one value for that attribute.
  - For example AadharID, Name, etc.

Multi-valued

- An Entity may have multiple values for that attribute.
  - For example ColorOfCar, CourseOfStudent, BestFriend.

Types of Attributes (3)

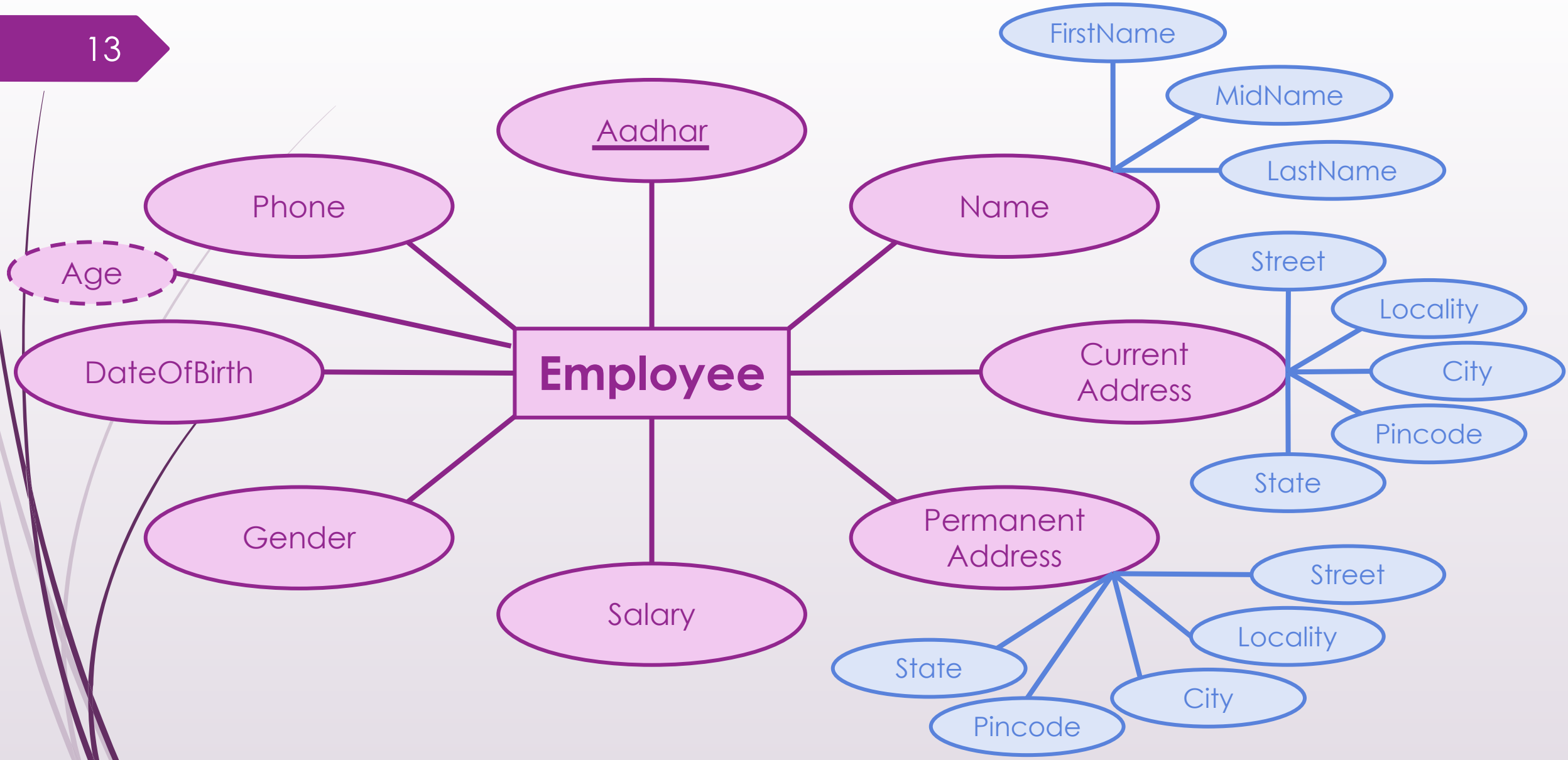
Stored

- Basic attribute that needs to be stored.
  - For example: BirthDate, CourseGrades, etc.

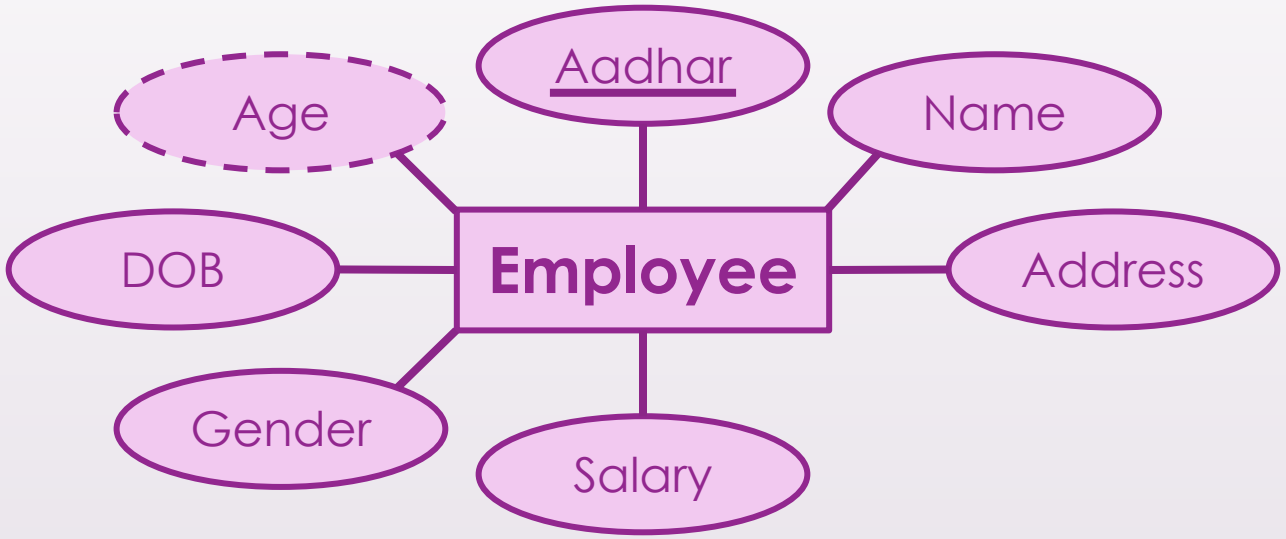
Derived

- The domain value of attribute can be determined from one or more other attributes.
  - For example, Age from BirthDate, SGA from CourseGrades.

Visualizing Entity Types & Attributes



Visualizing Entity Types & Attributes



Relationships and Relationship Types

Relationship

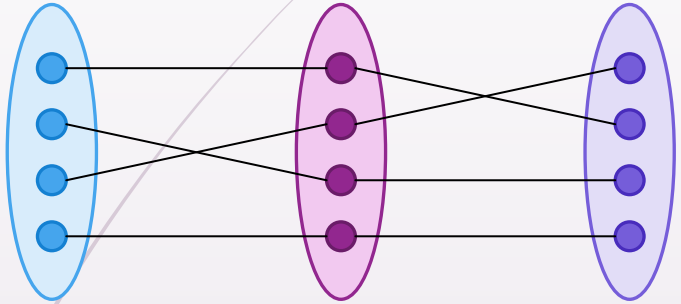
- Relates two or more distinct entities with a specific meaning.
  - For example, Employee John Smith works on the Database Projec
  - For example, Employee Jim Wong manages the Research Department.

Relationship Types

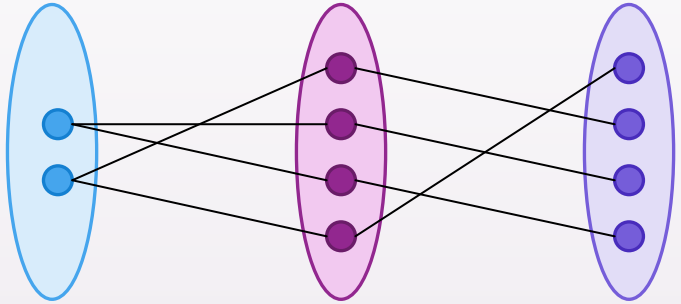
- Relationships of the same type are grouped or typed into a relationship type.
  - For example, the Works_On relationship type in which Employee's and Project's participate.

Relationship Types (1)

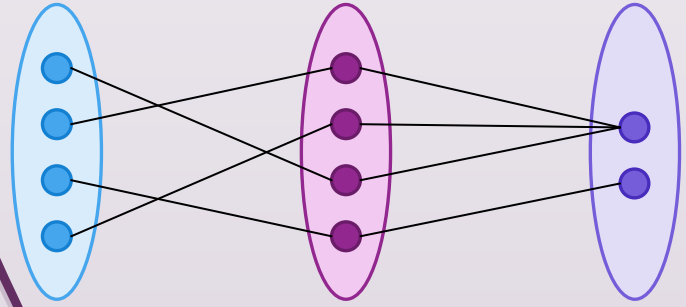
ONE-TO-ONE



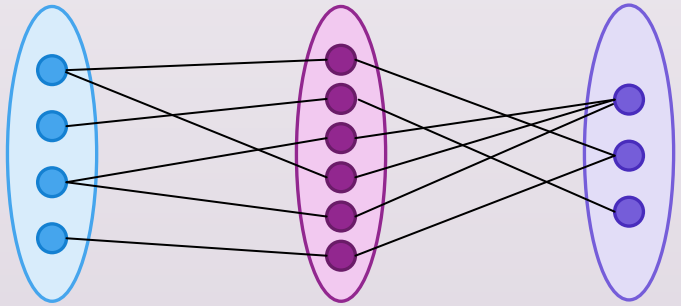
ONE-TO-MANY



MANY-TO-ONE

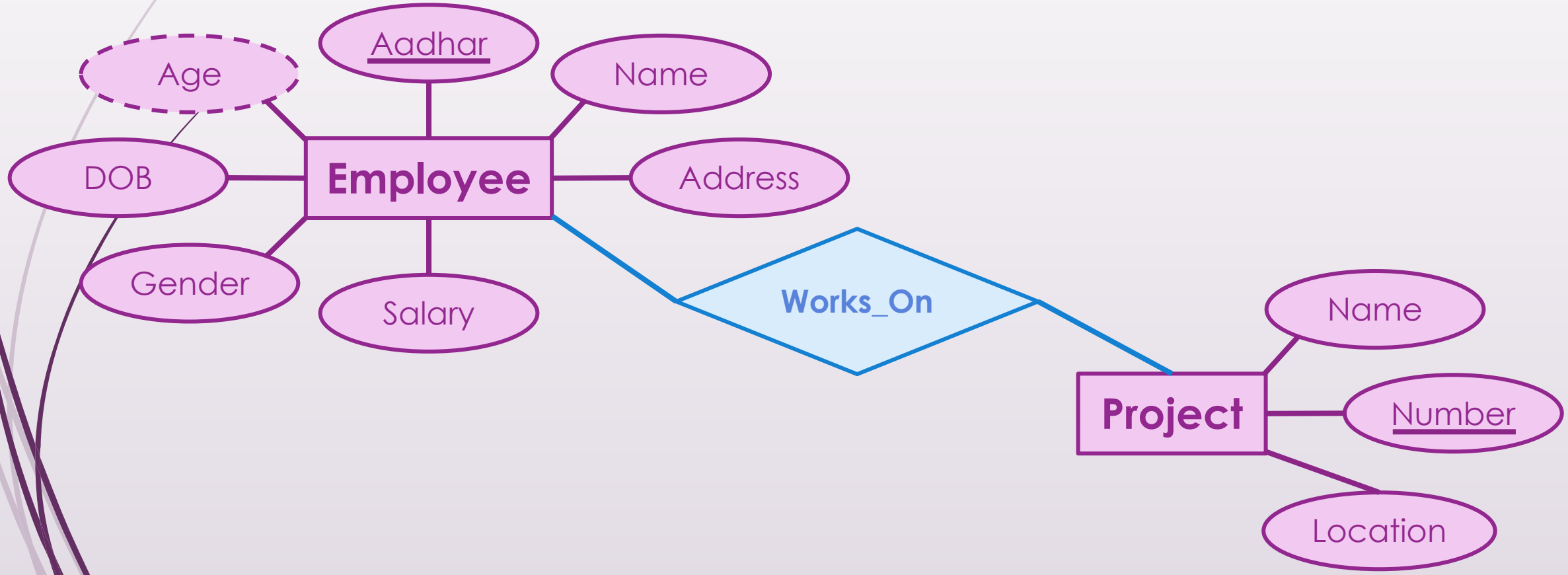


MANY-TO-MANY



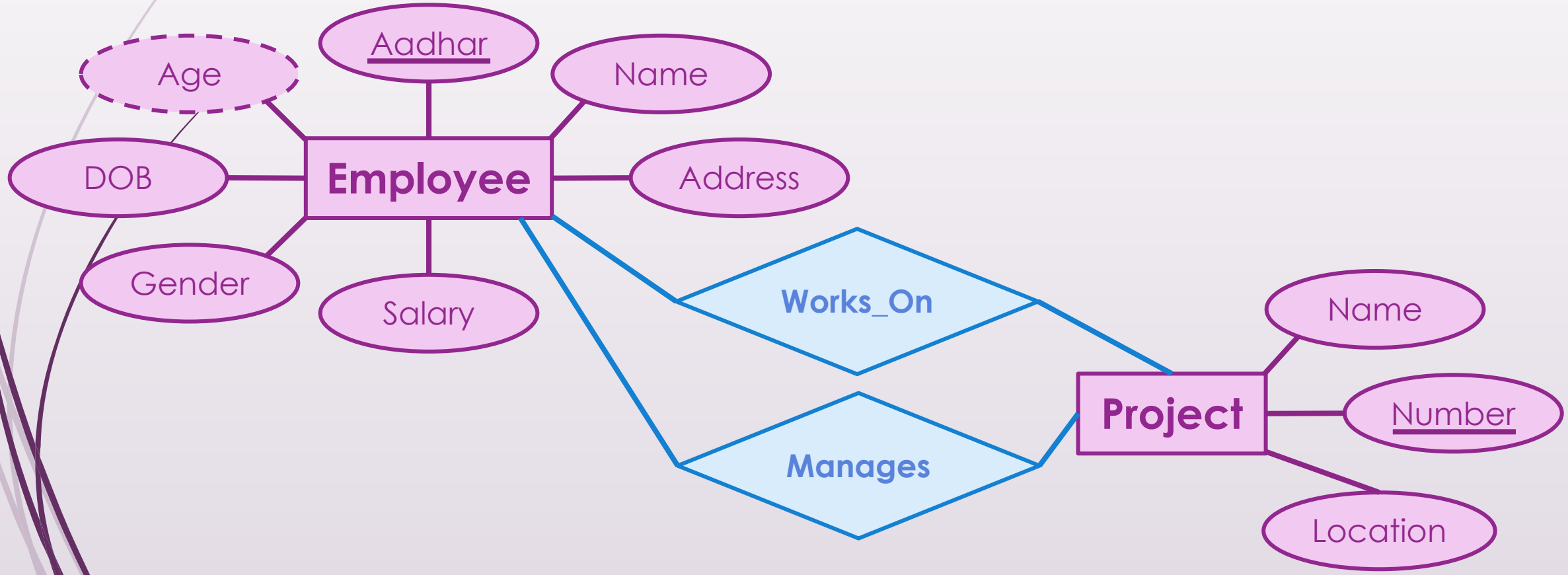
Relationship Types (2)

- The degree of a relationship type is the number of participating entity types.
  - For example, Works_On is a binary relationship.



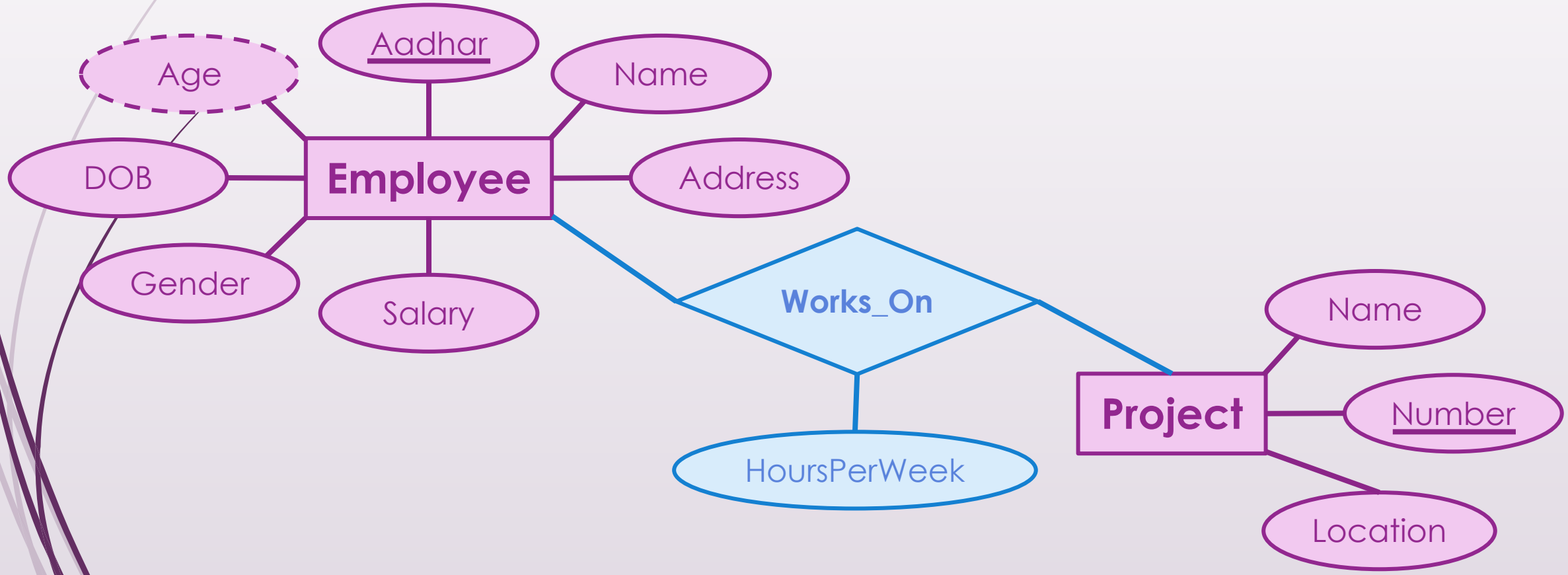
Relationship Types (3)

- More than one relationship type can exist with same participating entity types.
  - For e.g., Manages & Works_For.



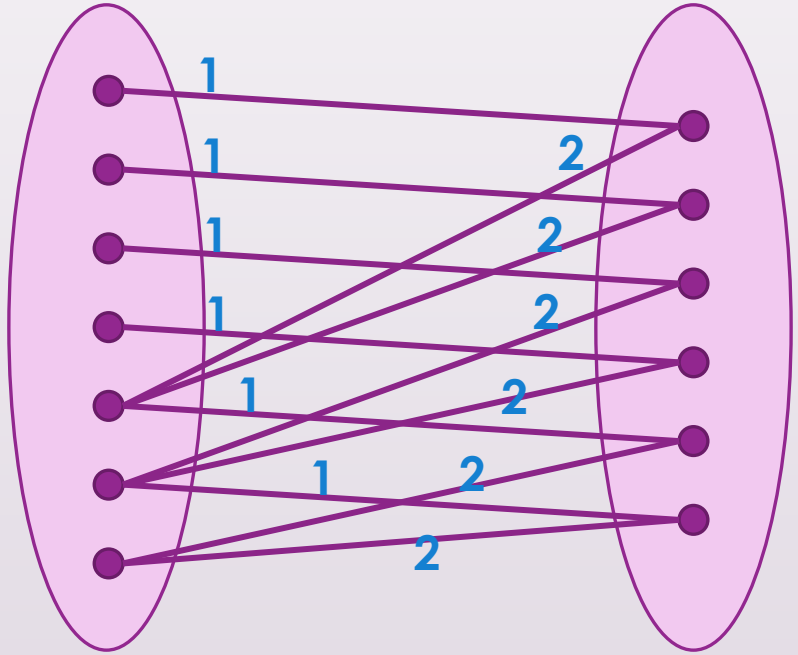
Relationship Types (4)

- A relationship type can have attributes; for example: HoursPerWeek of Works_On; its value for each relationship instance describes the number of hours per week that an Employee works on a Project



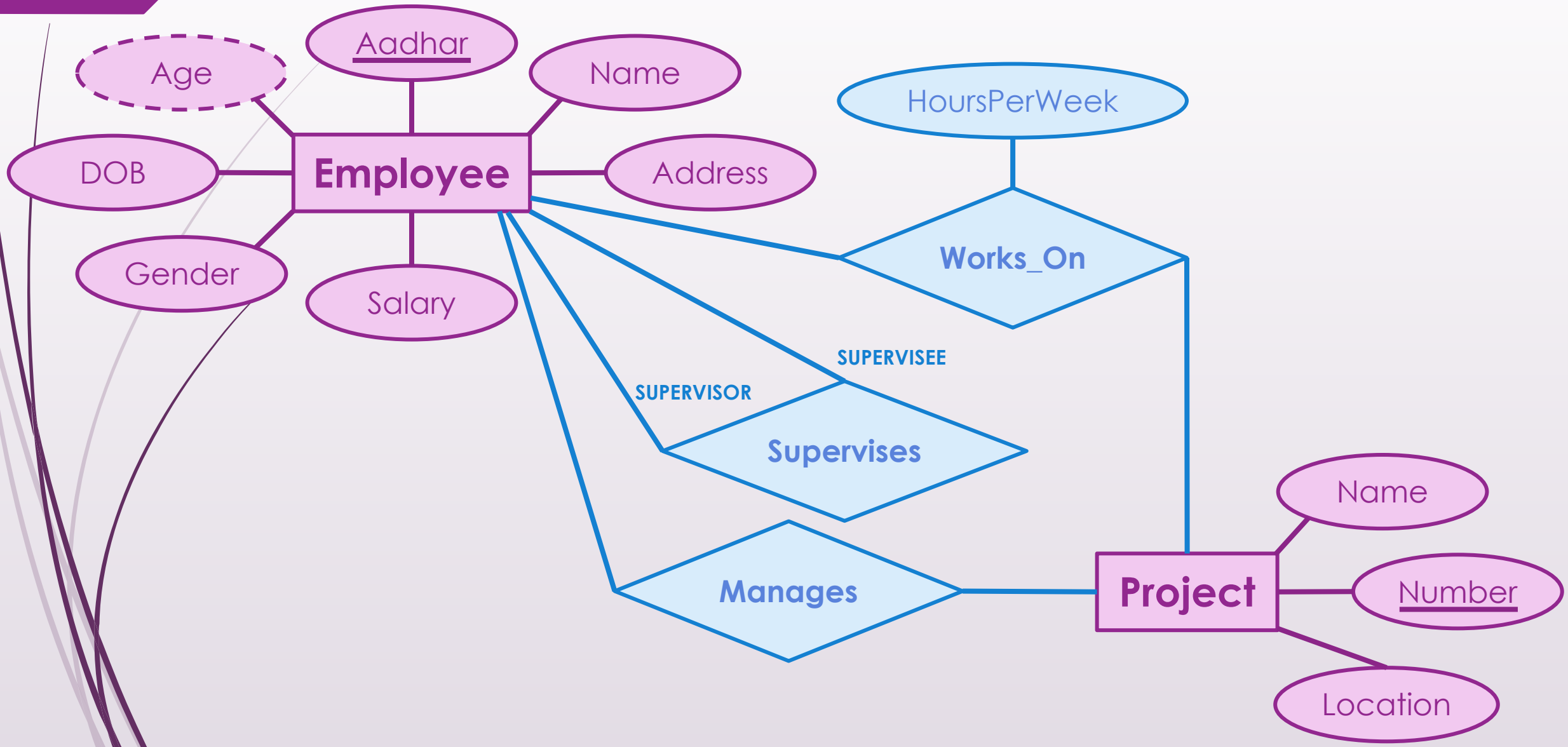
Relationship Types – Roles

- Roles of distinct entity types in a relationship are not necessary. But are required if the same entity type participates in the relationship. For example, a Supervision relationship type relates one Employee (in the role of Subordinate) to another Employee (in the role of supervisor). This is called a **recursive relationship type**.



1: Subordinate Role
2: Supervisor Role

Visualizing Relation Types



Exercise-1: Draw ER Diagram

- The university keeps track of each student's name, student ID, Aadhar number, current address and phone number, permanent address and phone number, date of birth, gender, class (freshman, sophomore, ..., graduate), major department, minor department (if any), and degree program (B.A., B.S., ..., Ph.D.). Some user applications need to refer to the city, state, and ZIP Code of the student's permanent address and to the student's last name. Both Aadhar number and student ID have unique values for each student.
- Each department is described by a name, department code, office number, office phone number, and college. Both name and code have unique values for each department.
- Each course has a course name, description, course number, number of semester hours, level, and offering department. The value of the course number is unique for each course.
- Each section has an instructor, semester, year, course, and section number. The section number distinguishes sections of the same course that are taught during the same semester/year; its values are 1, 2, 3, ..., up to the number of sections taught during each semester.
- A grade report has a student, section, letter grade, and numeric grade (0,1, 2, 3, or 4).

Exercise-2: Draw ER Diagram

- Design an ER schema for keeping track of information about votes taken in the U.S.House of Representatives during the current two-year congressional session.
- The database needs to keep track of each U.S. STATE's Name (e.g., 'Texas', 'New York', 'California') and include the Region of the state (whosedomain is {'Northeast', 'Midwest', 'Southeast', 'Southwest', 'West'}).
- Each CONGRESS_PERSON in the House of Representatives is described by his or her Name, plus the District represented, the Start_date when the congressperson was first elected, and the political Party to which he or she belongs (whose domain is {'Republican', 'Democrat', 'Independent', 'Other'}).
- The database keeps track of each BILL (i.e., proposed law), including the Bill_name, the Date_of_vote on the bill, whether the bill Passed_or_failed (whose domain is {'Yes', 'No'}), and the Sponsor (the congressperson(s) who sponsored—that is, proposed—the bill).
- The database also keeps track of how each congressperson voted on each bill (domain of Vote attribute is {'Yes', 'No', 'Abstain', 'Absent'}).

Exercise-3: Draw ER Diagram

Consider a MAIL_ORDER database in which employees take orders for parts from customers. The data requirements are summarized as follows:

- The mail order company has employees, each identified by a unique employee number, first and last name, and Zip Code.
- Each customer of the company is identified by a unique customer number, first and last name, and Zip Code.
- Each part sold by the company is identified by a unique part number, a part name, price, and quantity in stock.
- Each order placed by a customer is taken by an employee and is given a unique order number. Each order contains specified quantities of one or more parts. Each order has a date of receipt as well as an expected ship date. The actual ship date is also recorded.

Exercise-4: Draw ER Diagram

Consider a CONFERENCE_REVIEW database in which researchers submit their research papers for consideration. Reviews by reviewers are recorded for use in the paper selection process. The database system caters primarily to reviewers who record answers to evaluation questions for each paper they review and make recommendations regarding whether to accept or reject the paper. The data requirements are summarized as follows:

- Authors of papers are uniquely identified by e-mail id. First and last names are also recorded.
- Each paper is assigned a unique identifier by the system and is described by a title, abstract, and the name of the electronic file containing the paper.
- A paper may have multiple authors, but one of the authors is designated as the contact author.
- Reviewers of papers are uniquely identified by e-mail address. Each reviewer's first name, last name, phone number, affiliation, and topics of interest are also recorded.
- Each paper is assigned between 2 and 4 reviewers. A reviewer rates each paper assigned to him or her on a scale of 1 to 10 in four categories: technical merit, readability, originality, and relevance to the conference. Finally, each reviewer provides an overall recommendation regarding each paper.
- Each review contains two types of written comments: one to be seen by the review committee only and the other as feedback to the author(s).