Software Engineering

Database
Data Modeling

Dr. Young-Woo Kwon

Outline

- Overview of Database Design Process
- ER Model Concepts
  - Entities and Attributes
  - Entity Types, Value Sets, and Key Attributes
  - Relationships and Relationship Types
  - Weak Entity Types
  - Roles and Attributes in Relationship Types
- ER Diagrams - Notation
- ER Diagram for COMPANY Schema
- Alternative Notations – UML class diagrams, others

Overview of Database Design Process

- Two main activities:
  - Database design
    - Focuses on conceptual database design
    - To design the conceptual schema for a database application
  - Applications design
    - focuses on the programs and interfaces that access the database
    - Generally considered part of software engineering

Overview of Database Design Process

- Step 1: requirements collection and analysis
  - Functional / Data requirements
- Step 2: conceptual design
  - Conceptual schema
- Step 3:logical design (or data model mapping)
  - Logical schema
- Step 4: physical design
  - Internal schema: internal storage structures, file organizations, indexes, access paths, and parameters for database files

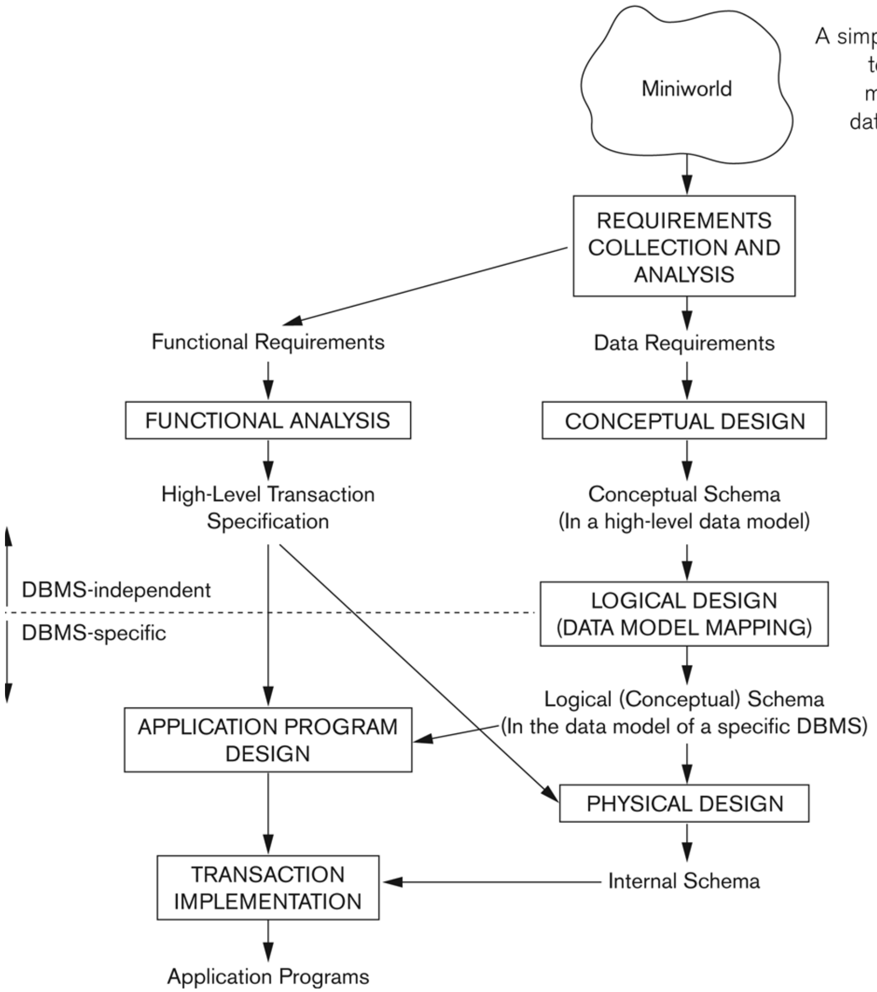

Miniworld
REQUIREMENTS COLLECTION AND ANALYSIS
Functional Requirements
Data Requirements
FUNCTIONAL ANALYSIS
CONCEPTUAL DESIGN
High-Level Transaction Specification
Conceptual Schema (In a high-level data model)
DBMS-independent
DBMS-specific
LOGICAL DESIGN (DATA MODEL MAPPING)
Logical (Conceptual) Schema (In the data model of a specific DBMS)
APPLICATION PROGRAM DESIGN
PHYSICAL DESIGN
TRANSACTION IMPLEMENTATION
Internal Schema
Application Programs

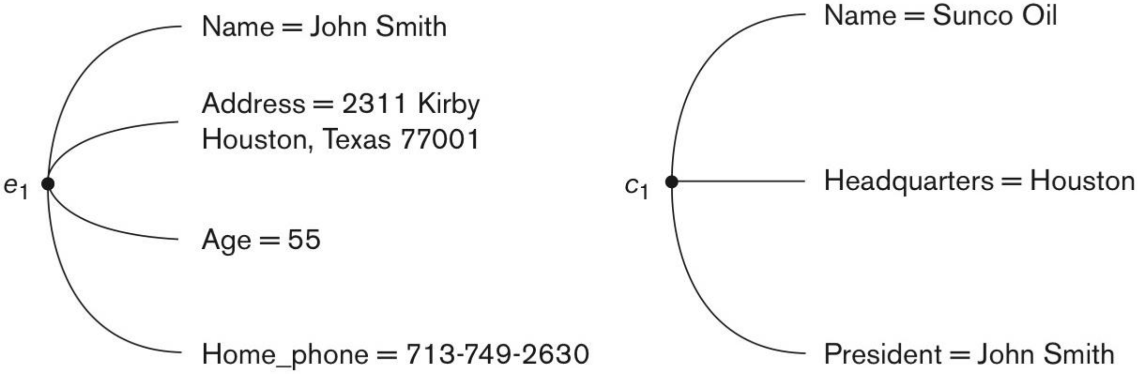
Figure 3.1
A simplified diagram to illustrate the main phases of database design.

Methodologies for Conceptual Design

- Entity Relationship (ER) Diagrams
- Enhanced Entity Relationship (EER) Diagrams
- Use of Design Tools in industry for designing and documenting large scale designs
- The UML (Unified Modeling Language) Class Diagrams are popular in industry to document conceptual database designs

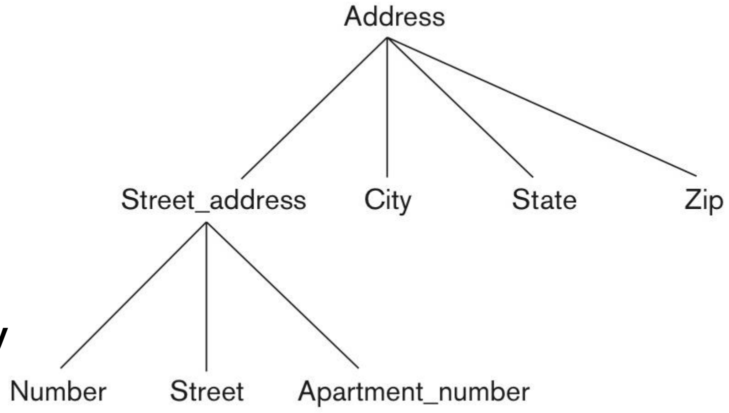
ER Model Concepts

- Entities
  - Entity is a basic concept for the ER model.
  - Entities are specific things or objects in the mini-world that are represented in the database.
- Attributes
  - Attributes are properties used to describe an entity.
  - A specific entity will have a value for each of its attributes.
  - Each attribute has a value set (or data type) associated with it – e.g. integer, string, date, enumerated type, …



Types of Attributes (1)

- Simple
  - Each entity has a single atomic value for the attribute. For example, SSN or Sex.
- Composite
  - The attribute may be composed of several components. For example:
    - Address(Apt#, House#, Street, City, State, ZipCode, Country), or
    - Name(FirstName, MiddleName, LastName).
    - Composition may form a hierarchy where some components are themselves composite.



- Multi-valued
  - An entity may have multiple values for that attribute.
    - For example, Color of a CAR or PreviousDegrees of a STUDENT. Denoted as {Color} or {PreviousDegrees}.

Types of Attributes (2)

- Composite and multi-valued attributes may be nested arbitrarily to any number of levels (very rare)
  - PreviousDegrees of a STUDENT is a composite multi-valued attribute denoted by {PreviousDegrees (College, Year, Degree, Field)}
  - Multiple PreviousDegrees values can exist
  - Each has four subcomponent attributes:
    - College, Year, Degree, Field

Example COMPANY Database

- We need to create a database schema design based on the following (simplified) **requirements** of 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 of the department manager. A department may have several locations.
  - Each department *controls* a number of PROJECTs. Each project has a unique name, unique number and is located at a single location.

Example COMPANY Database

- The database will store each EMPLOYEE's social security number, address, salary, sex, and birthdate.
    - Each employee *works for* one department but may *work on* several projects.
    - The DB will keep track of the number of hours per week that an employee currently works on each project.
    - It is required to keep track of the *direct supervisor* of each employee.
- Each employee may *have* a number of DEPENDENTs.
    - For each dependent, the DB keeps a record of name, sex, birthdate, and relationship to the employee.

Initial Conceptual Design of Entity Types for the COMPANY Database Schema

- Based on the requirements, we can identify four initial entity types in the COMPANY database:
  - DEPARTMENT
  - PROJECT
  - EMPLOYEE
  - DEPENDENT
- Their initial conceptual design is shown on the following slide
- The initial attributes shown are derived from the requirements description

Initial Design of Entity Types:

EMPLOYEE, DEPARTMENT, PROJECT, DEPENDENT

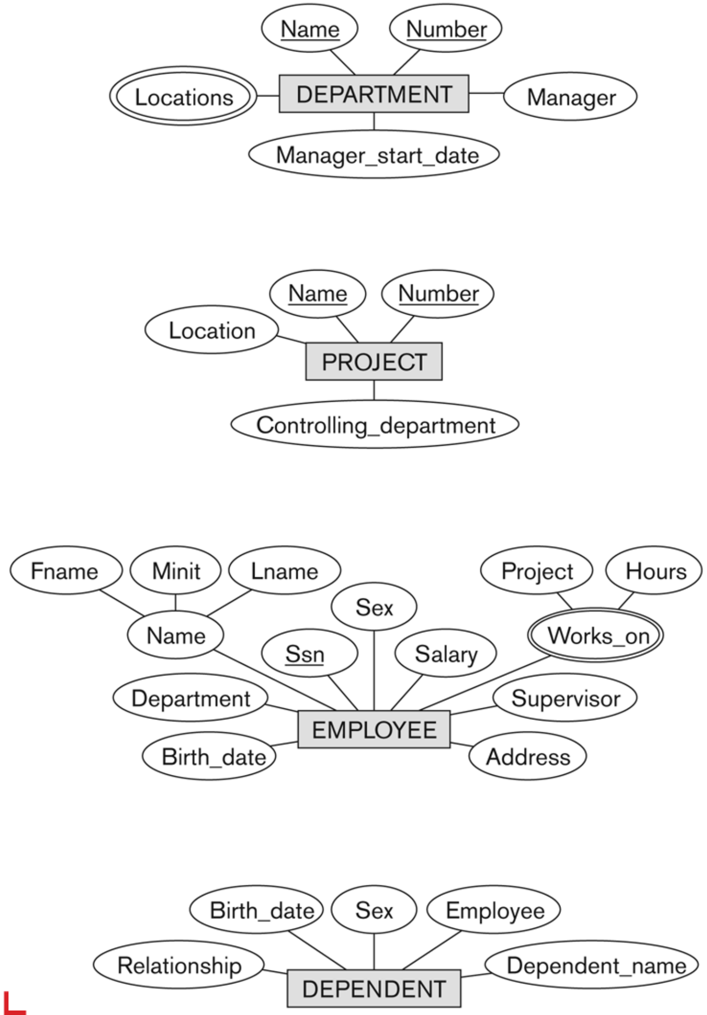
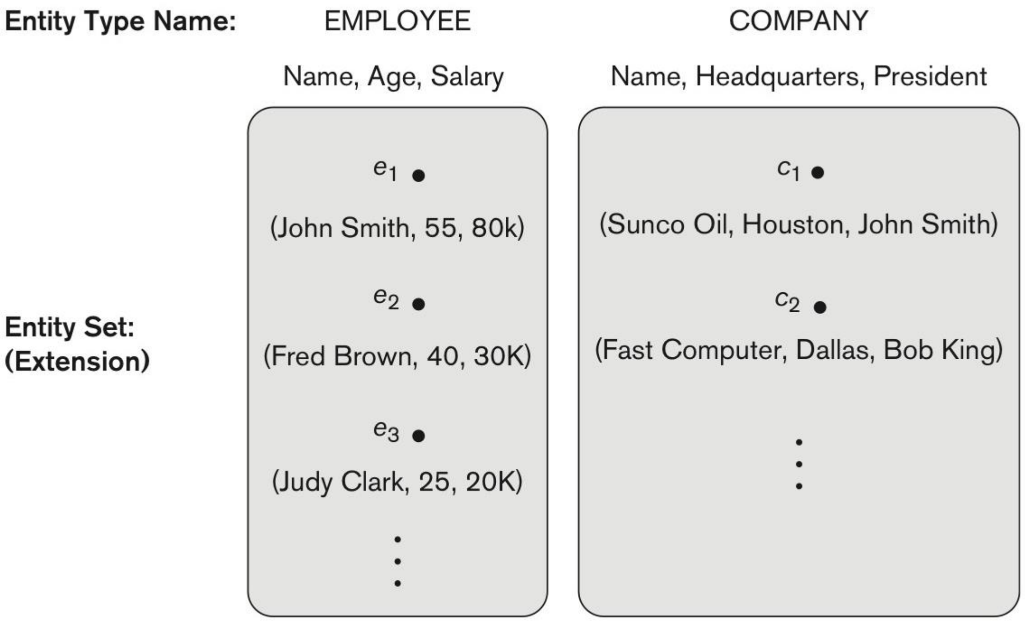


Figure 3.8
Preliminary design of entity types for the COMPANY database. Some of the shown attributes will be refined into relationships.

Entity Types and Key Attributes (1)

- Entities with the same basic attributes are grouped or typed into an entity type.
  - For example, the entity type EMPLOYEE and PROJECT.
- An attribute of an entity type for which each entity must have a unique value is called a key attribute of the entity type.
  - For example, SSN of EMPLOYEE.



Entity Type Name: EMPLOYEE COMPANY

Name, Age, Salary

Name, Headquarters, President

Entity Set: (Extension)

e_1 ●

(John Smith, 55, 80k)

e_2 ●

(Fred Brown, 40, 30K)

e_3 ●

(Judy Clark, 25, 20K)

⋮

c_1 ●

(Sunco Oil, Houston, John Smith)

c_2 ●

(Fast Computer, Dallas, Bob King)

⋮

Entity Types and Key Attributes (2)

- A key attribute may be composite.
  - VehicleTagNumber is a key of the CAR entity type with components (Number, State).
- An entity type may have more than one key.
  - The CAR entity type may have two keys:
    - VehicleIdentificationNumber (popularly called VIN)
    - VehicleTagNumber (Number, State), aka license plate number.
- Each key is underlined (Note: this is different from the relational schema where only one “primary key is underlined).

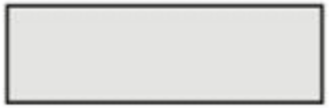
Entity Set

- Each entity type will have a collection of entities stored in the database
  - Called the **entity set** or sometimes **entity collection**
- Previous slide shows three CAR entity instances in the entity set for CAR
- Same name (CAR) used to refer to both the entity type and the entity set
- However, entity type and entity set may be given different names
- Entity set is the current *state* of the entities of that type that are stored in the database

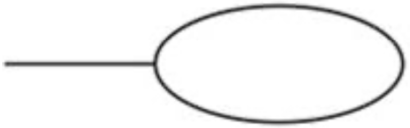
Value Sets (Domains) of Attributes

- Each simple attribute is associated with a value set
  - E.g., Lastname has a value which is a character string of up to 15 characters, say
  - Date has a value consisting of MM-DD-YYYY where each letter is an integer
- A **value set** specifies the set of values associated with an attribute
- Value sets are similar to the basic data types available in most programming languages: e.g. integer, string, double, ...

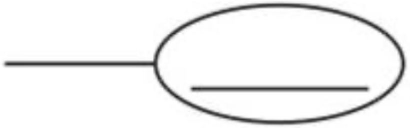
NOTATION for ER diagrams



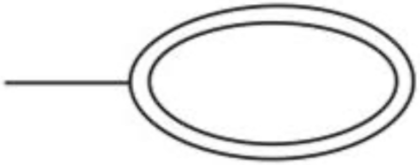
Entity



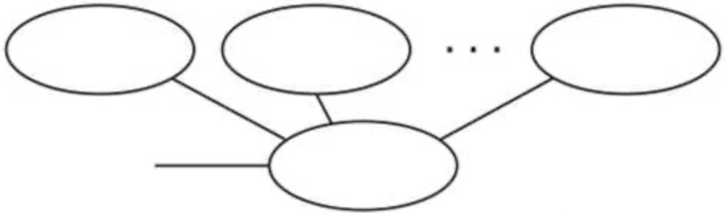
Attribute



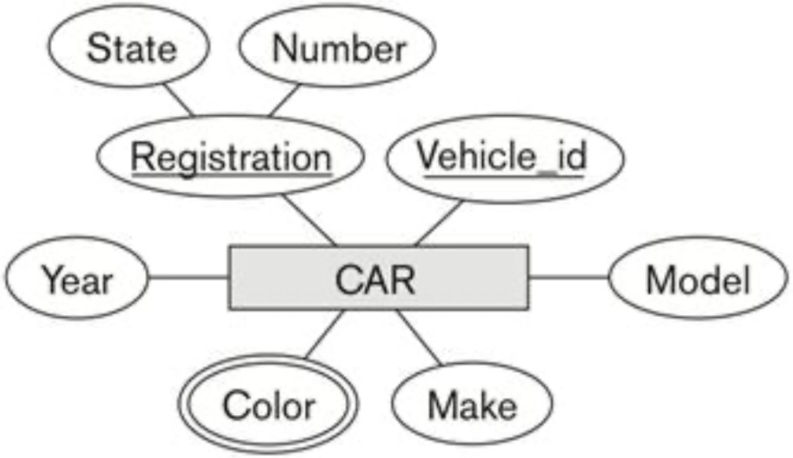
Key attribute



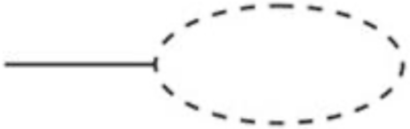
Multivalued Attribute



Composite attribute



[ER diagram notation of the CAR entity type]



Derived attribute

Initial Design of Entity Types:

EMPLOYEE, DEPARTMENT, PROJECT, DEPENDENT

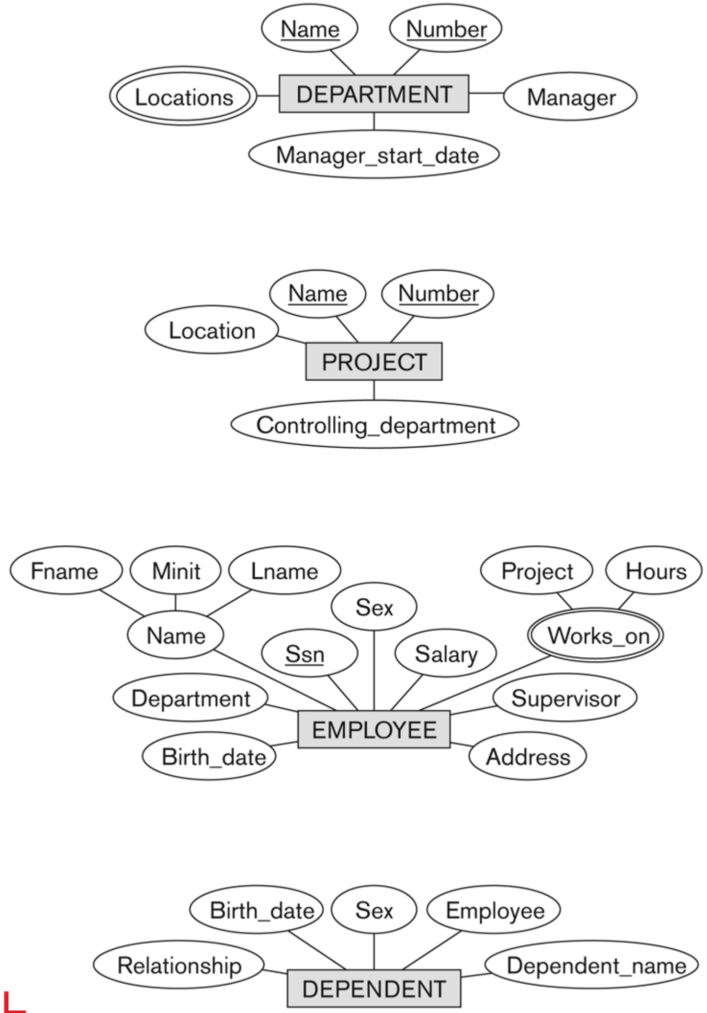


Figure 3.8
Preliminary design of entity types for the COMPANY database. Some of the shown attributes will be refined into relationships.