

COS221

L07 - DB Design using UML

(Chapter 10 in Edition 6 and Chapter 3.8 and 4.6 in Edition 7)

Linda Marshall

6 March 2023

Multiple methodologies exist to represent database design. These include:

- ▶ Chen's - extended to include generalisations and specialisations. Good educational notation.
- ▶ Information Engineering (IE) - most popular and has the most products that can be used to model and translate the model to SQL
- ▶ Barker's - used by Oracle and when using Oracle products
- ▶ UML - Only notation (used for OO modelling) that adheres to a standard.

Introduction to UML

- ▶ *Class*, comprises of *attributes* and *operations*. Attribute domains (data types) are optional. Operations are not specified on ER diagrams. In UML operations can be viewed as *stored procedures*.
- ▶ Modelling complex attributes - *Composite* attributes modelled as structured domains, e.g. Name. *Multivalued* attributes are modelled as a separate class, e.g. LOCATION.
- ▶ *Relationships* are modelled using associations.
 - ▶ Associations have *multiplicity* and *roles*.
 - ▶ There are two variants of association, aggregation and composition.
 - ▶ Associations can be unidirectional (arrowed) or bidirectional.
 - ▶ *Reflexive* associations link to self.

UML notation used in data modelling

- ▶ Adding an association name to the association (dashed line on association) provides *ordering* to the relationship based on the attributes specified in the box.
- ▶ Weak entities are modelled using UML *qualified associations/aggregations*. A *discriminator* is used to model the partial key.

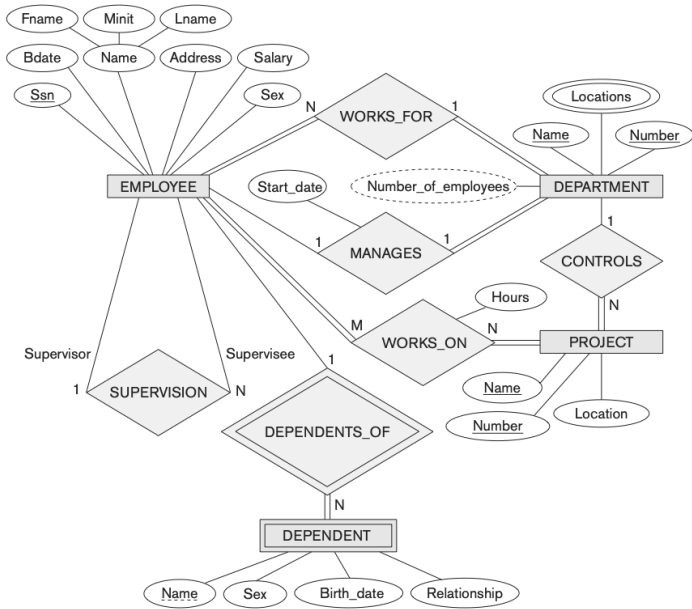


Figure 3.2

An ER schema diagram for the COMPANY database. The diagrammatic notation is introduced gradually throughout this chapter and is summarized in Figure 3.14.

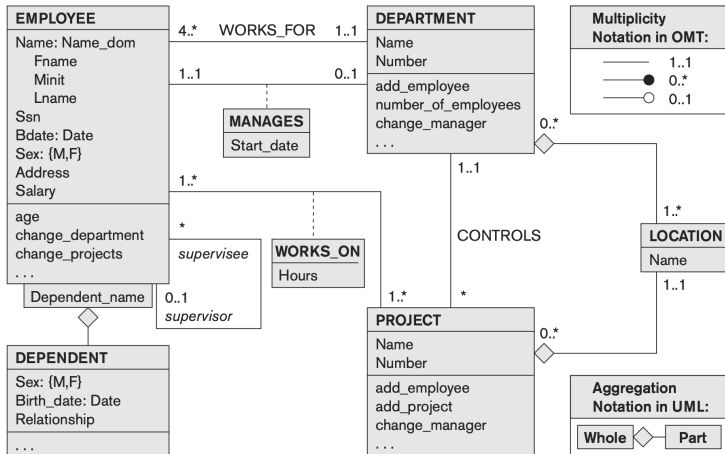


Figure 3.16

The COMPANY conceptual schema in UML class diagram notation.

Modelling Specialisation and Generalisation

- ▶ Make use of *inheritance*
 - ▶ Extend the UML inheritance notation to make provision for *overlapping* constraints (filled-in triangle). *Disjoint* constraints makes use of the standard inheritance triangle.
 - ▶ In UML classes can be modelled as Abstract, Concrete and Template. For database design the models typically make use of Concrete classes for entities.

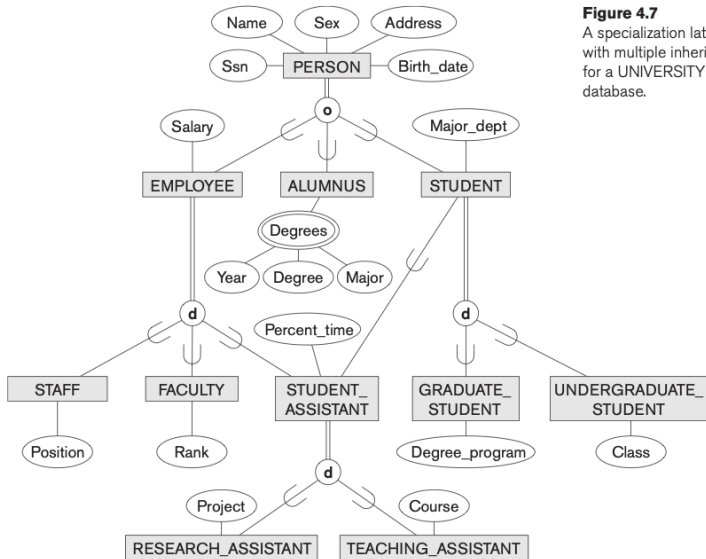


Figure 4.7

A specialization lattice with multiple inheritance for a UNIVERSITY database.

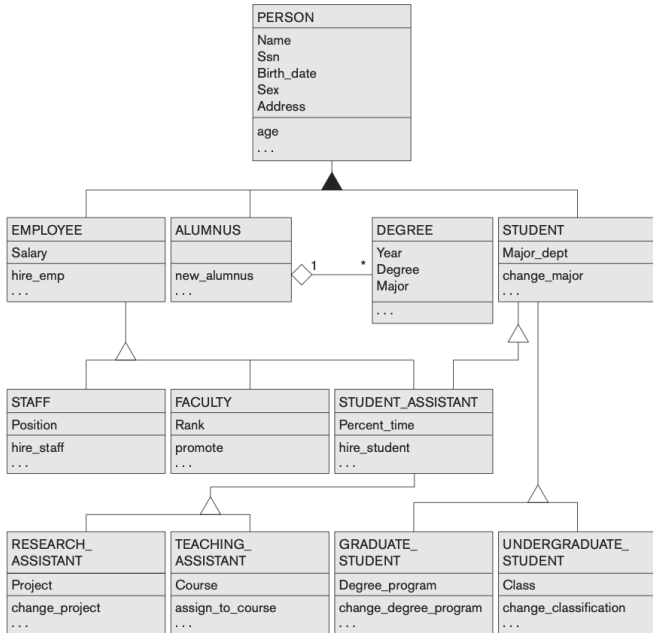


Figure 4.10

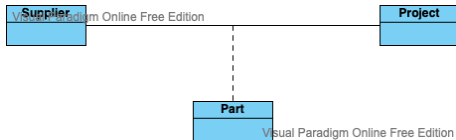
A UML class diagram corresponding to the EER diagram in Figure 4.7, illustrating UML notation for specialization/generalization.

N-ary Relationships

- N-ary relationships are modelled in a similar manner to Chen. A diamond links the associations between the classes.



- To transform the relationship to a binary relationship, make use of the dashed line on the association in UML to model the ternary relationship - association relationship.



UML Extensions

- ▶ Some tool manufacturers have developed UML extensions to model concepts such as primary keys in the UML database models. The depiction is similar to the Barker notation for primary keys.