Diseño conceptual: Entidad relación extendido

CI-0127 Bases de Datos, Universidad de Costa Rica

Sivana Hamer

Importante: Este documento recopila contenidos de diversos de sitios web especializados, académicos y documentos compartidos por universidades. Toda la información es utilizada con fines estrictamente académicos. Para más información, puede ver las referencias.

1 Extended entity relationship model

The *extended entity relationship* (EER) data model extends the traditional ER model to better express the data in certain situations. The model adds the concepts of subtype (Section 2), specialization & generalization (Section 3) with their constraints (Section 4).

Four our example, we shall continue with the student example described previously. The EER diagram is shown in Fig. 1. Another EER notation is shown in Fig. 2.

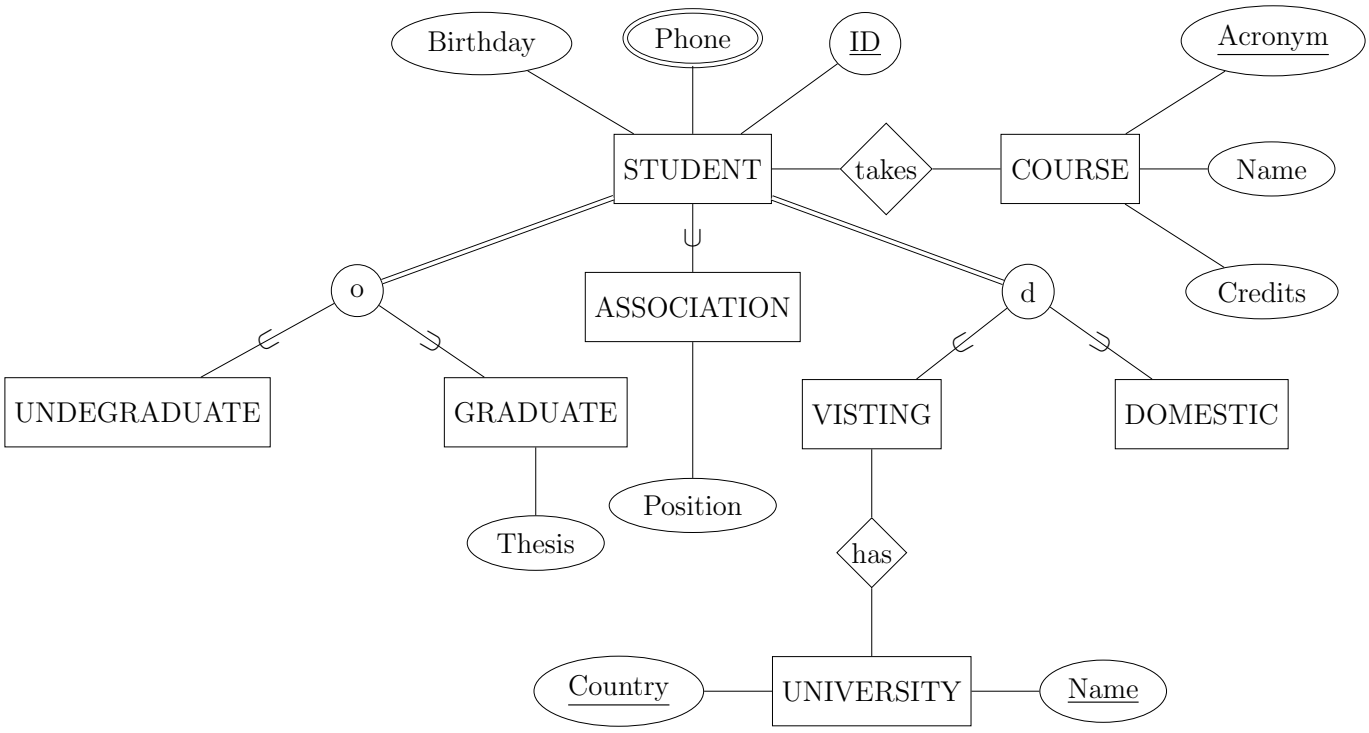

Figure 1: EER diagram for students

2 Subtype

A *subtype* or *subclass* of an entity are subgroups for an entity type. For example, the STUDENT entity can be divided as UNDERGRADUATE (associates or bachelors), GRADUATE (masters or doctorate), DOMESTIC, VISITING or as part of the student ASSOCIATION. The different groups for an entity are the subtype, while the *superclass* or *supertype* is the higher level entity. This relationship is also called ISA (IS A or IS AN) type like a UNDERGRADUATE is a STUDENT. We

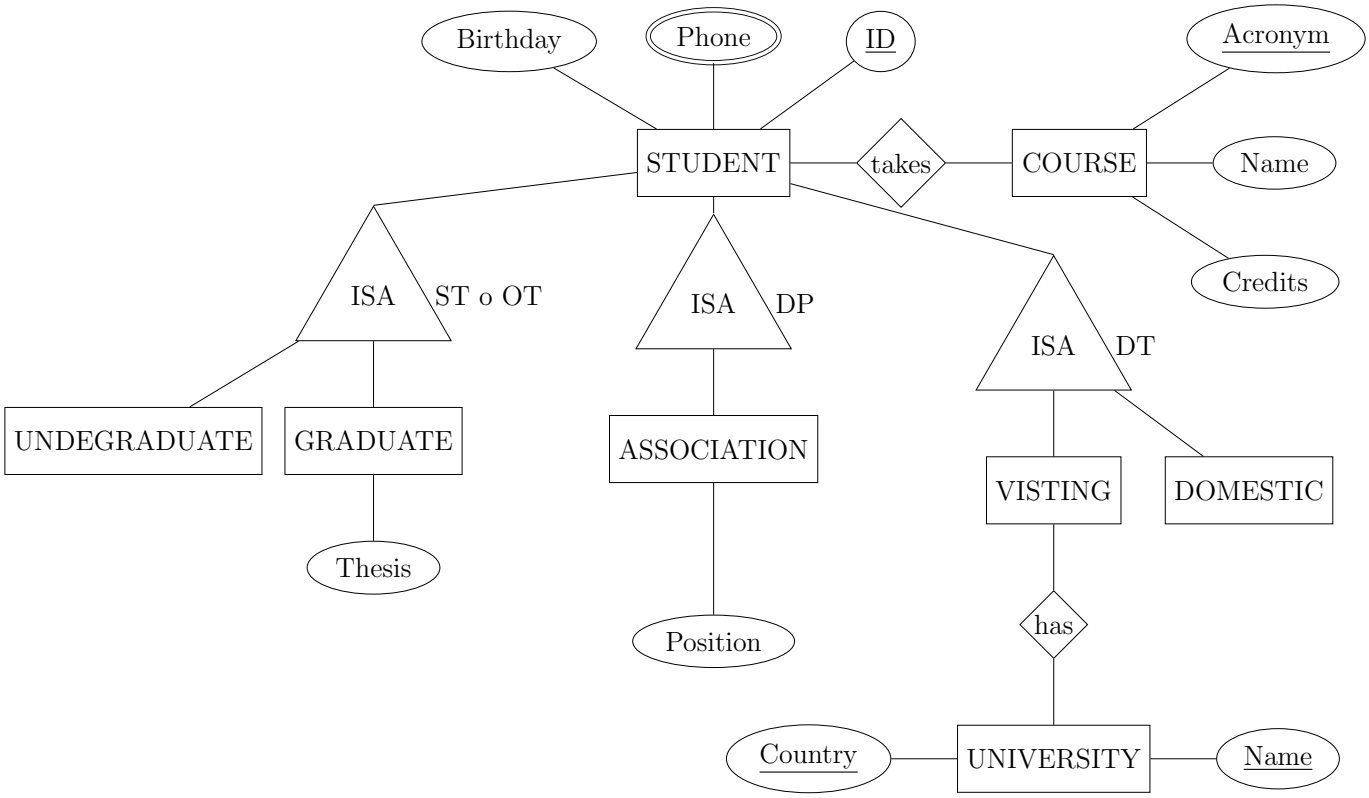

Figure 2: Alternative notation for EER diagram for students

can see that this is either represented by a line with a circle and another line defining the hierarchical structure between supertypes and subtypes, or with a triangle that has an ISA.

Each subtype may have different attributes, like GRADUATE students having a flag to indicate if they completed their dissertation or thesis. Subtypes may also have different relationships than the supertype, like VISITING student IS from another UNIVERSITY. These relationships may also have attributes. The attributes and relationships of the superclass are inherited by the subclass.

3 Specialization & Generalization

Specialization is the process where we define a *set of subclasses* for a *superclass*. For example, in our previous example we have three sets. First, we have the type of degree the student is getting, either being UNDERGRADUATE or GRADUATE. Second, the set where the students are getting the degree. They can be either DOMESTIC or VISITING from some other university. Finally, we have a set if the students are part of the ASSOCIATION of students. We can see that every set has a different circle or triangle with the ISA. We define these sets with these subclasses starting from the top (superclass) to the bottom (subclass).

Meanwhile, *generalization* is the opposite process where we define a superclass for a set of subclasses due to their similarity (in attributes and or relationships). For example, we could have defined UNDERGRADUATE and GRADUATE as entities and then have noticed that we could generalize certain attributes and relationships into one entity, STUDENT. We generalize the subclasses by analyzing the bottom (subclass) and determining their similarities at the top (superclass).

At the end of the day, defining which process was followed is somewhat subjective, so we shall not differentiate between both in our EER diagrams.

4 Constraints

There are different constraints we can define between the set of subclasses.

Disjointness constraint: *Disjointness* defines that members of the subclass can only be one the subclasses of the specialization. For example, a STUDENT may either be DOMESTIC or VISITING but cannot be both. We represent this using a *d* to denote disjointness. Meanwhile, *overlapping* entity types allow for an entity to belong to multiple types within the specialization. Such as someone can be both an UNDERGRADUATE and GRADUATE student at the same time. We represent this using a *o* to denote overlapping. Sometimes, in Spanish we also use *s* for *solapamiento*.

Completeness or totalness constraint. *Totalness* determines if all entities in the superclass must be part of a subclass. For example, all STUDENTs must detail if they are either UNDERGRADUATE or GRADUATE. This is denoted by a double line between the circle and the superclass, or a *t* for total besides the ISA. Meanwhile, if it is not mandatory for the entities in the superclass to have a subclass, we have a *partial specialization*. For example, a student may be part of the student ASSOCIATION. This would be partial specialization as not all students would be part of the ASSOCIATION (only a select few). We denote partialness with only one line between the circle and the superclass, or with a *p* for partial besides the ISA.

Therefore, we can have due to the constraints the following combinations:

- Disjoint + Total (DT)
- Disjoint + Partial (DP)
- Overlapping + Total (OT o ST)
- Overlapping + Partial (OP o SP)

References

[1] R. Elmasri and S. Navathe, *Fundamentals of database systems*, 7th ed. Pearson, 2016, chapters 3 and 4.

[2] A. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*, 7th ed. New York, NY: McGraw-Hill, 2020, chapter 6.

[3] C. Faloutsos and A. Pavlo. Lecture #2. [Online]. Available: https://15415.courses.cs.cmu.edu/fall2016/

