

Advanced databases - Assignment 2

José Cutileiro

April 23, 2024

1 E-R diagram

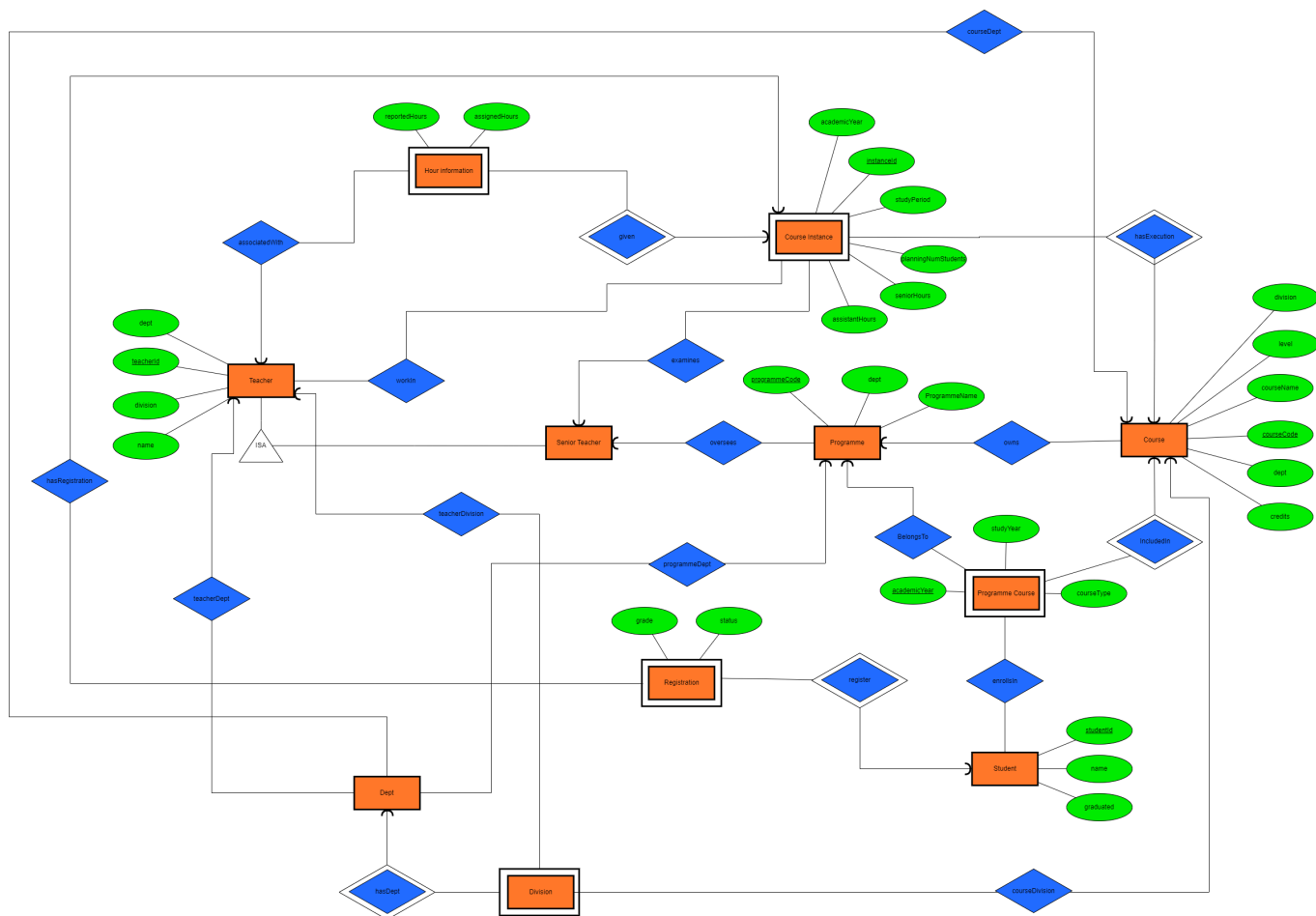
The technical changes involving alterations to the scheme were made in the second delivery of the project as well as the changes related to the report. For this second assignment, I simply added color to make it easier to visualize when we're framing it in Protege.

The colors are related to the Protege UI:

Orange - Classes

Blue - Object Properties

Green - Data properties



Note: I will once again include the diagram in the submission to avoid the risk of poor readability

2 Resubmission of the project

After talking to the professor in the lab session, I corrected some things that were wrong.

1. The organization of the classes in Protege goes with the E-R diagram without the addition of additional concepts.
2. The concept of division and department have earned the right to have a class given their presence in many different classes.

3 OWL 2 (turtle format)

3.1 Introduction

I aimed to make the transition from the diagram to OWL 2 as smooth as possible, avoiding major disparities between the two. However, some semantic and practical nuances lead to slight differences. Therefore, let's now analyze some of these distinctions.

3.2 Classes

3.2.1 The concept of person was included:

In my original E-R diagram, I omitted the 'person' class. However, I consider it crucial to define the concept of person, especially since one of the objectives is to determine the number of people present. Therefore, I chose to include it in my solution for the second task. Both students and teachers are considered people.

3.2.2 Information associated with a person:

It can be observed that both the hour information associated with teachers and the registration information associated with students are relevant to individuals. Therefore, I chose to incorporate the concept of information about people, aiming to provide a greater semantic context to these relationships.

3.2.3 Information associated with a course:

The same reasoning was followed for information about courses, including information about course programs and even instances of the courses.

3.3 Object Properties

Firstly, I would like to clarify the organization of the relationships between the different classes into three main categories:

1. Actions by people: Encompasses actions performed by individuals. For example, 'examines' (A senior professor examines an instance of a course).
2. Course description: Includes descriptive aspects related to a course or program. For example, 'owns' (A programme owns a course).
3. Associated with someone: Refers to relationships involving information about individuals. For example, 'associatedWith' (Schedule information is associated with a specific professor).

Furthermore, when defining the domain and range of each object property, I used the property names themselves as a guide. For instance, with 'oversees', we understand that a professor supervises a specific program

`<Senior_Teacher> <oversees> <Programme>`

Drawing from the guidance provided in class, I assumed that the class on the left belongs to the domain and the class on the right belongs to the range.

3.4 Data properties

In the distribution of attributes, I chose to categorize them into three distinct classes:

1. **Course attributes:** Encompass characteristics associated with a specific course or program.
2. **Person attributes:** Include characteristics exclusively related to the person themselves, without information about the courses they attend or teach.
3. **Attributes associated with personal information:** In this category, we find attributes linked to the courses that a specific person attends or teaches.

Note that the attributes 'division' and 'department' do not fall into any of the three main classes. This is because they are attributes shared by both people and courses/programs, and therefore do not align with any specific main family.

Furthermore, I remained quite faithful to the original diagram, essentially performing a direct translation from the diagram to OWL 2. Regarding the types of attributes, I guided myself by the CSV files provided in the previous task, aiming to be as accurate as possible. For example, intuitively, we might consider a student identifier to be an integer, but in the CSV files, we find formats like '19970308-0000', where the presence of the '-' character indicates it is a string. At least in the current phase, I chose not to make changes in that regard, although I acknowledge that this could be analyzed later on.