

# DAT475 / DIT930 Advanced databases

## 2023-2024, Study Period 4

### Assignment 1

#### Background

In the introductory databases course, a text description of a domain is provided and students are asked to draw an Entity-Relationship (E-R) diagram based on the text. The students then use “mechanical” rules to generate a good relational design from the E-R diagram. In the first assignment in this course, we’ll go in the opposite direction, starting with a relational design and constructing an E-R diagram that covers the information that is present in the relational design. However, the relational design is not a good one! So you should make some improvements as you construct the E-R diagram.

#### Task description

Draw an E-R diagram based on the following relational design.

SeniorTeachers(teacherId, name, dept, division)

TeachingAssistants(teacherId, name, dept, division).

# These can be currently registered students who are employed part-time as teaching assistants,  
# or they can be PhD students who are employed full-time.

Programmes(programmeCode, programmeName, dept, director)  
director → SeniorTeachers.teacherId

Courses(courseCode, courseName, credits, level, dept, division, ownedBy)  
ownedBy → Programmes.programmeCode  
# level is one of “first cycle” or “second cycle”.

programmeCourses(programme, academicYear, studyYear, course, courseType)  
programme → Programmes.programmeCode  
course → Courses.courseCode  
# studyYear is “1” or “2” for Master’s programmes and “1”, “2” or “3” for Bachelor’s programmes.  
# courseType is one of: “compulsory”, “compulsory-elective” or “elective”.  
# Programmes can include different courses in different academic years.  
# The same course could have a different courseType in different academic years.

CourseInstances(courseCode, studyPeriod, academicYear, instanceId, examiner)  
courseCode → courses.courseCode  
examiner → SeniorTeachers.teacherId

CoursePlanning(course, planningNumStudents, seniorHours, assistantHours)  
course → CourseInstances.instanceId  
# planningNumStudents is the number of students used for planning course resources  
# seniorHours is the estimated number of senior teachers’ hours that the course will need.  
# assistantHours is the estimated number of teaching assistant hours that the course will need.

AssignedHours(courseCode, studyPeriod, academicYear, teacherId, hours)  
(courseCode, studyPeriod, academicYear) →  
CourseInstances.(courseCode, studyPeriod, academicYear)

ReportedHours(course, teacherId, hours)  
course → CourseInstances.instanceId

Students(studentId, name, programme, year, graduated)  
(programme, year) → ProgrammeCourses.(programme, academicYear)  
# year is the academic year in which the student joined the programme.  
# graduated is a Boolean attribute.

Registrations(courseInstance, student, status, grade)  
course → CourseInstances.instanceId  
student → Students. studentId  
# status is one of: "not started", "registered", "completed", "withdrawn".  
# For students who have completed the course, the grade is one of "5", "4" or "3".

In addition to the E-R diagram, your report should include comments:

- Stating any additional assumptions that you have made. For example, keys for the relations are not stated in the task description, and some references (foreign keys) are missing.
- Explaining design decisions that you have made. For example, if you considered more than one way to model a particular part of the domain, explain your choice and why you preferred it to the alternative that you considered.

## Data files

The assignment page in Canvas includes links to .csv files with data corresponding to the eleven relations that are listed above. You will work with these data files in assignments 3 and 4. These files are not needed for assignment 1, but they are provided now in case looking at their contents can help clarify your understanding of the domain.

## "20 queries"

In the spirit of Jim Gray's "20 queries" heuristic, here are some questions that we will want to be able to answer with a data system that is based on the E-R model that is designed in this assignment:

- Find the name, director and department of all programmes.
- Find the names of all students who worked as teaching assistants in courses given by the D3-2 division in study period 2 in academic year 2023/2024.
- Find all teachers who are assigned more than 120 hours in course 1015 in study period 1 in academic year 2018/2019.
- Find all students registered for course instance I-910 that were **not** registered for course instance I-911.
- Find all programmes along with the total number of owned courses. List the results in descending order of number of owned courses.

- Find the number of senior teachers.
- Find the number of all people.
- Find the name, department and division of all teaching assistants.
- Find all students who are registered in course 1516 and course 1518 in study period 4 in academic year 2017/2018.
- Find all course instances of courses owned by the programmes P-52 or P-71 where the total assigned hours is larger than the total number of reported hours.
- Find all programmes that do **not** have a compulsory course in study period 3 of academic year 2013/2014.
- Find the programme that owns the most courses.
- Find the course given by the D1-4 division in academic year 2023/2024 that has the most registered students.

Note: you are not expected to answer these queries in this assignment!  
Rather, you can use this list of queries to help check your E-R model.

### Frequently asked questions

Q. Can an instance belong to more than one subclass?

A. Yes

Q. Can there be two or more relationships between the same pair of entity classes.

A. Yes

Q. In programmeCourses, what is the difference between academicYear and studyYear?

A. Consider an example. The [programme syllabus for study year 1 the MPDSC programme for academic year 2022/2023](#) lists courses that are compulsory, elective, or compulsory elective. If we look at the programme syllabus for MPDSC for other academic years, the list of courses might be different.

Q. How do divisions relate to departments?

A. A department can have several divisions; each division is part of one department. For example, the Department of Computer Science and Engineering (CSE) has four divisions: Computer and Network Systems (CNS), Computing Science (CS), Data Science and AI (DSAI), Interaction Design and Software Engineering (IDSE).

Q. Must the examiner for a course be employed at the division that arranges a course?

A. No, the examiner can be a senior teacher from another division or even from another department.

Q. Can we create entity classes in the E-R model for which there is no corresponding relation in the task description?

A. Yes (assuming that this improves the design and does not make it unnecessarily complicated).

## Drawing tools

The Dia Diagram Editor<sup>1</sup> can be used to draw Entity-Relationship diagrams. However, you are not required to use Dia in this assignment, and you are welcome to use another drawing program if you prefer.

If you want to use Dia, you can either install it on your own computer, or you can use the version of Dia that is installed centrally on the Linux system at Chalmers (either working directly with a Linux computer in a classroom, or by using Remote Desktop to connect to a Chalmers Linux server).

## Submitting work

You should submit:

- A PDF file that contains an E-R diagram and your comments on its design.
- (Optional) If the E-R diagram in the PDF file is not sufficiently clear, you can submit it as a separate image file in a standard image format (e.g., PNG).

In each file that you submit, give the names of the people submitting the work and your assignment group number (as given in Canvas).

(Please do not submit .zip files; these cannot be processed conveniently in Canvas.)

Deadline: Tuesday 9 April 2024 at 23:59.

---

<sup>1</sup> <http://dia-installer.de/>