Corso di laurea in Informatica - Dipartimento di Informatica Università di Pisa

Progetto VIII – appello marzo 2021

BD: Consegna entro il 24 marzo 2021 ore 23:59

Progetto di Basi di dati

General project delivery rules

The project must be uploaded to the MOODLE page of the course by the date set for the exam session.
Out of fairness towards group mates, whoever submits also sends an email containing the project to the teacher and to all group members. The email must have subject: ProgettoBD. All emails (sender and recipients) must be institutional (@ unipi.it / @ students.unipi.it).

In addition to the information published on MOODLE:

- Diagrams (including attributes and association names) can be drawn digitally or drawn on paper and photographed, as long as they are clearly legible and as long as they are legible without the need to rotate the screen.
- A single file in PDF format or an archive containing the requested material must be delivered. In any case, the title of the course, the name, surname and serial number of the members of the group, the title and the delivery date of the project must be present in the title page of the document.
- The name of the file uploaded to Moodle must have as a prefix the concatenation of the surnames (with a capital letter) of the project members.

The Project

A professional firm wants to keep track of ongoing practices (or “cases”). Each customer can be a person or an organization. Each customer has a denomination, an address, and a telephone number. For people, we are also interested in their name, surname and tax code. For organizations, the VAT number is of interest. In addition, for each organization, it is interesting to know which of the customers who are persons plays a role in that organization, and what role they play; the same person can have different roles in multiple organizations at the same time. For each case, it is important to keep track of a customer who is responsible for the case (person or organization). When the responsible is an organization, it is also interesting to know a group (in general) of persons who follow the procedure on behalf of the organization; these persons are in general a subset of the persons who have a role in the organization. The organization issues invoices, each sent to a customer and each related to a specific case. Customers make payments, which can be linked to invoices in a complex way (payments may only partially cover an invoice, a payment can be linked to several invoices…) and payments can be performed in any way (cash, bank transfer, credit card…). Invoices and payments must be recorded by the system.

This description is intentionally incomplete and open to different interpretations. It is up to the student to choose a reasonable interpretation and specify the specifications consistently. Therefore, starting from the general description given previously and integrated above, the student is asked to produce a text structured as follows:

1. Description of the domain

Inspired by the description provided, the student gives a precise description of the domain, on the model of those used in the homework of past years (http://pages.di.unipi.it/ghelli/bd1/2019.04.03.BD.compitino1 .solutions.pdf). Nothing too long, a period is enough for each class identified, and a number of classes between six and ten is reasonable. The description provided must make it possible to understand what are the associations between the classes and the relative cardinalities. The student must ensure the internal consistency of the description produced. Provide at least one instance of subclass hierarchy.

2. Conceptual schema

The student produces an object-oriented conceptual scheme in graphic format with attributes (and association names) that corresponds to the domain of point 1, indicating in a textual way the constraints not captured graphically. Specify both interrelational and intrarelational constraints. Provide at least one instance of subclass hierarchy.

The student is asked to use the graphic notation seen during the course.

3. Relational logical schema

The student produces a relational logic scheme, in graphic and textual format with the notation R (IdR, ..., A *) which corresponds to the scheme of point 2. Specify if there are functional dependencies and if these dependencies respect the normal form of Boyce Codd .

The student is asked to use the graphic notation seen during the course.

4. SQL queries

The student defines a list of at least 6 queries (describing them in plain text and then writing the respective queries in SQL) relating to the domain specified in point 1, at his choice, which have the following characteristics respectively:

a. use of projection, join and restriction;
b. use of group by with having, where and sort;
c. use of join, group by with having and where;
d. use of nested select with existential quantification;
is. use of nested select with universal quantification;
f. use of a quantified comparison against a subquery, such as for example:
attribute =ANY / > ALL (SELECT .. FROM .. WHERE)

5. Access plans

I. Write a logical access plan of queries a), b), c);

II. Write an efficient physical access plan for the three logical access plans in point I that does not use indexes, and does not use NestedLoop (MergeLoop is ok) and (optional) check if sorting before Group By can be avoided.

III. Write an efficient physical access plan for the three logical access plans in point I that make use of two indexes (or in any case the maximum number of possible indexes), and (optional) check if the sort before the Group By can be avoided.

The omission of one of the above points in the response can make the entire project insufficient.