

CS 348
Introduction to Database Management
Assignment 2

Due: 30 October 2012 9:00AM
Returned: 8 November 2012
Appeal deadline: One week after return
Lead TA: Jiewen Wu

Submission Instructions: By the indicated deadline you are to use the submit command to submit a file containing SQL queries that implement each of the requests for information given in Question 1. You are also to submit written or typed answers to Question 2 in the assignment boxes. For the online submission, put all your queries (in the order specified in this handout) in one file named `a2.sql`. We should be able to run the queries using the command `db2 -f a2.sql`. Assume a database connection already exists (i.e., you do not need `connect` or `disconnect` statements in your submitted file). To submit your assignment use `submit cs348 a2 .` (notice the dot “.” at the end indicating that you submit the entire directory).

Please submit the answers to second and third questions by placing hard copy solutions in the assignment boxes.

Question 1.

For the first question, you must use your Unix accounts and DB2 to compose and evaluate a number of SQL queries for a database that records information about courses. The schema for the database is illustrated in Figure 1 that depicts a relational database schema that includes an indication of primary and foreign key constraints in the manner discussed in class. Sample commands for defining the base tables for this schema can be downloaded from the course web site (go to the Assignments tab). Note that the schema stores information about both ongoing and past classes for a course. Also note that no marks are recorded for any enrollment of an ongoing class, and that, for a past class, a mark is recorded for each of its enrollments. Finally, you may assume that each class has at least one enrollment.

- (a) The number, name and department of each instructor who has taught in the past on Mondays.
- (b) The number of instructors who are currently teaching CS348.
- (c) The course number, course name and grade of each course completed by a student whose student number is 1234.
- (d) The number, name and year of each student who is not in his or her first year, who has a final grade of at most 70 in every course that she/he has completed

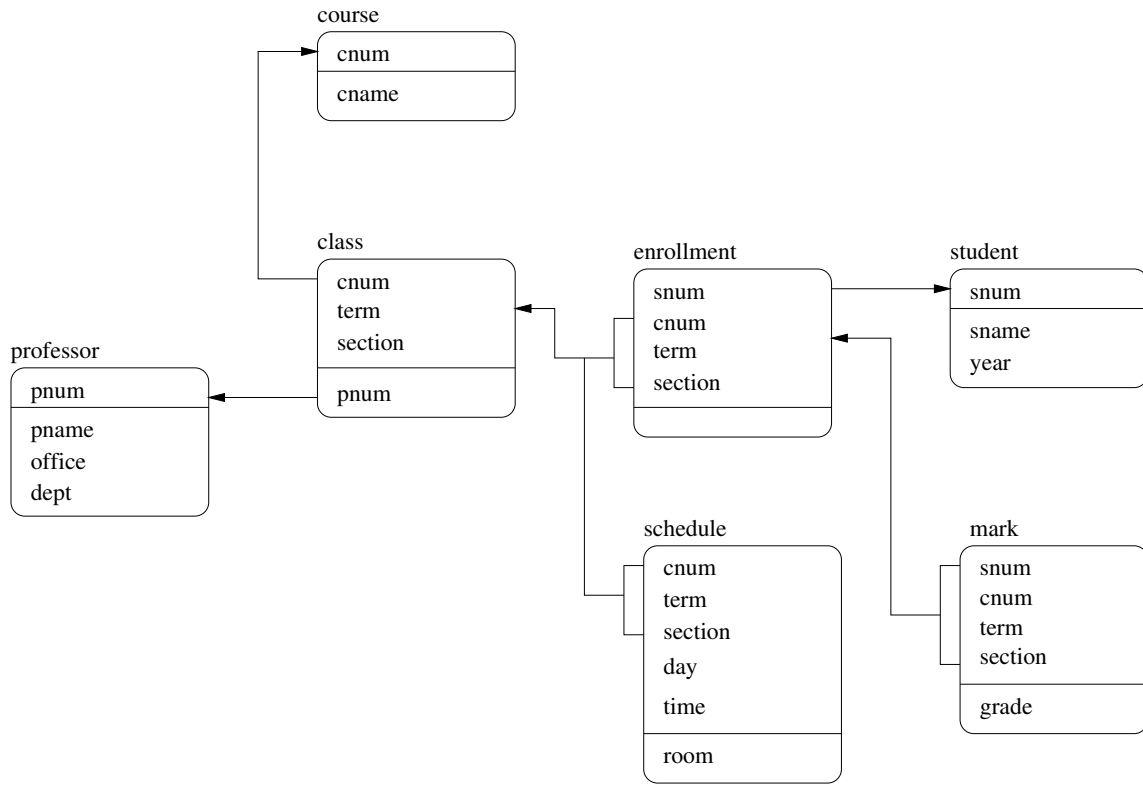


Figure 1: Schema for Question 1

- and who was not taught by a professor in the philosophy department (including current courses). Note: use department name “Philosophy”.
- (e) The number, name and department of each professor who has no current teaching commitments on either Mondays or Fridays, sorted by department and then by the professor’s name. Note: use full day names, e.g., “Monday”.
 - (f) The course number, term and section number of each class taught in the past by some professor, together with the number and name of that professor.
 - (g) The course numbers and total enrollment of courses with total enrollment counts among the three lowest. (Note that one possible result could be: [‘CS448’, 120], [‘CS446’, 120], [‘CS341’, 105], [‘CS246’, 110]. Also note that all classes, past and ongoing, need to be considered.)
 - (h) The number of different third year students in each section of each course taught by each professor in the current term. The result should include the professor number, professor name, course number and section, and should also be sorted first by the name of the professor, then by the professor number, third by the course number, and finally by section. (Note that a section is identified by a term and a section number. Also assume that sorting by section means sorting by term and then by section number. The result will therefore have a total of six columns.)
 - (i) The minimum and maximum final grade for each class that was taught in the past by a professor in the statistics department. The result should include the number and name of the professor, and the course number, course name, term and section of the class. Note: use department name “Statistics”.
 - (j) The percentage of departments with professors who have never taught more than one class in the same term.

Question 2.

Write queries in the relational algebra for each of the above specifications for which this is possible. (Hint: there are five of them.)

Question 3.

Consider the following schema given in Figure 2 where the underlined attributes are primary keys and the arrows between relations depict the foreign key-primary key relationships. Also consider the following information about relations and give the appropriate DDL statements to create each relation.

Note: A straightforward solution would not allow you to insert tuples into EMPLOYEE and/or DEPARTMENT after creation. So, you need to be careful how you create these tables and modify them after an insertion. If you can figure out how to do this, there is a bonus of 5 points.

- In EMPLOYEE, Fname and Lname can vary in length with a maximum of 15 characters; Minit is a single character; SIN consist of 9 characters; Address is variable length up to 30 characters; Sex is identified as 'M' or 'F'; Salary is a decimal number with 10 digits and 2 digits after the decimal point; Super_Id has the same specification as SIN; Dno is an integer; Bdate is of type date. For each

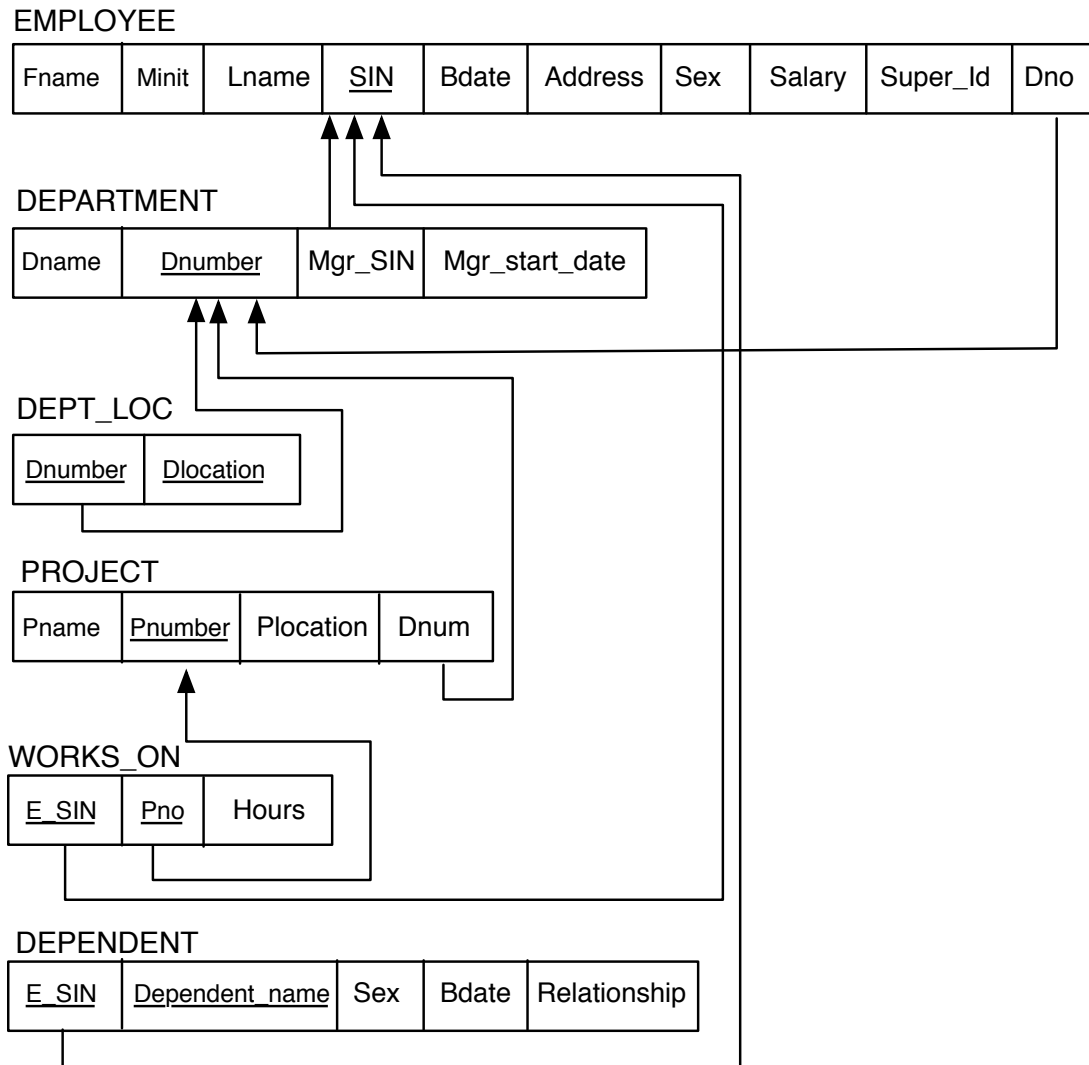


Figure 2: Schema for Question 3

tuple, Fname, Lname, SIN, have to be specified, and each employee has to be assigned to a department.

- In DEPARTMENT, Dname is variable length character string with a maximum length of 15; Mgr_SIN is a fixed length string of 9; Dnumber is an integer; and Mgr_start_date is a date. Dname and Mgr_SIN have to be specified, and every department has to have a manager.
- In DEPT_LOC, Dnumber is the same as what it is in DEPARTMENT, and Dlocation is a variable length character string of at most 15 characters. Both of these have to be specified for any tuple.
- In PROJECT, Pname and Plocation are maximum 15 character strings; Pnumber and Dnum are integers. Pname, Pnumber and Dnum have to be specified for each tuple.
- In WORKS_ON, E_SIN is a 9 character string; Pno is an integer; and Hours is a 3 digit decimal number with one digit after the decimal point. All of these have to be specified for each tuple.
- In DEPENDENT, E_SIN is the same as it is in WORKS_ON; Dependent_name is a character string of up to 15 characters; Sex is identified as 'M' or 'F'; Bdate is a date; and Relationship is a character string of up to 8 characters.