

CS430/630 – Homework 3

Released Mar 06, Due Mar 27

50 points (5/100 of final grade)

Instructions: The homework is due BEFORE CLASS on Tue Mar 27th. Please hand in paper copies (preferably typeset, although hand-written copies will be accepted as well) for Questions 1 and 2. For Question 3, you have to create scripts for the queries, and place them in your directory for the course. Create a folder “HW3” under your main folder for the course, and put all queries in-order in a single file named “Q3.sql”. Ensure that the file is not readable by others (comments in the SQL file to distinguish between queries are expected, as this is a matter of good coding style).

Question 1 (20 points)

A university database contains information about professors (identified by social security number *SSN*) and courses (identified by *courseid*). Professors also have a name, an address and a phone number. Courses have a name and a number of credits. Professors teach courses. For each of the following situations, draw an ER diagram that describes it (assuming no further constraints hold).

- (a) Every professor must teach some course.
- (b) Every professor teaches exactly one course (no more, no less).
- (c) Every professor teaches exactly one course (no more, no less), and every course must be taught by some professor.
- (d) [630 students only] Modify the diagram from (a) such that a professor can have a set of addresses (which are street-city-state triples) and a set of phones. Recall that in the E/R model there can be only primitive data types (no sets).
- (e) [630 students only] Modify the diagram from (d) such that professors can have a set of addresses, and at each address there is a set of phones.

Question 2 (15 points)

Let *a* and *b* be integer-valued attributes that may be NULL in some tuples. For each of the following conditions that may appear in a WHERE clause, describe exactly the set of (*a*, *b*) tuples that satisfy the condition, including the case where *a* and/or *b* is NULL.

- (a) $a=10 \text{ OR } b=20$
- (b) $a=10 \text{ AND } b=20$
- (c) $a<10 \text{ OR } a\geq 10$
- (d) [630 students only] $a=b$

Question 3 (15 points)

Consider a database schema with three relations:

```
Employee (eid:integer, ename:string, age:integer, salary:real)
Works (eid:integer, did:integer, pct_time:integer)
Department(did:integer, dname:string, budget:real, managerid:integer)
```

The keys are underlined in each relation. Relation `Employee` stores employee information such as unique identifier `eid`, employee name `ename`, age and salary. Relation `Department` stores the department unique identifier `did`, department name `dname`, the department budget and `managerid` which is the `eid` of the employee who is managing the department. The `managerid` value must always be found in the `eid` field of a record of the `Employee` relation. The `Works` relation tracks which employee works in which department, and what percentage of the time s/he allocates to that department. Note that, an employee can work in several departments.

Provide SQL statements for the following:

- Create a view `ManagerSummary` that lists for every department the department name, manager ID and manager name, manager salary and the number of employees in that department. The view will have five columns with headings: `DeptName`, `MgrID`, `MgrName`, `MgrSalary` and `EmpCount`.
- Query the view above to retrieve the set of distinct salaries of managers who manage a department called "Sales".
- Query the view above to find the name of the manager who manages most employees. If the same employee works in several departments, that employee is counted once in each of the departments. The manager is included in the count the same as all other employees, i.e., based on his or her records in the `Works` table.