CS430/630 – Homework 3

Released Apr 05, Due Apr 16

50 points (5/100 of final grade)

Instructions: The homework is due on **Fri Apr 16th, 23:59:59**. Submissions must be **TYPESET**; submission must be in a single file called **HW3.pdf**.

The timestamp will be considered according to the UMB email server. Submissions received (i.e., tagged) later than 23:59:59 EST on Apr 16 **WILL NOT BE GRADED**. You are only allowed a **SINGLE** submission (i.e., you cannot send multiple emails – only the **first received** will be graded).

Email submissions must be sent to Nicholas J Pankewytch at <N.Pankewytch001@umb.edu>

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 OR b=20`
(b) `a=10 AND b=20`
(c) `a<10 OR a>=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:

(a) 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`.

(b) Query the view above to retrieve the set of distinct salaries of managers who manage a department called "Sales".

(c) 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.

.