Jason Rapsinski
CS 248 HW 3

Consider the following relational schema:

Department (deptid, dname, location)

Student (snum, sname, deptid, age)

Faculty (fid, fname, deptid)

Class (cname, deptid, meets_at, room, faculty)

Enrolled (snum, cname, score, grade)

Assume that:

- All students and faculty members must major/work in one department;

- Faculty teach courses only in the department they belong to;

- The names of the courses are unique for each department;

- Scores are numbers in [0,100];

- Faculty in Table Class is a subset of fid in Table Faculty.

• Write each of the following queries using Query-by-Example (QBE). If you believe any one of the following is not expressible in QBE, informally explain why it cannot be expressed.

1. (15 pts) Print the names of the department(s) that have more courses than the CS department (deptid-101).

Department	Deptid	Dname	Location
	_x	P.UNQ	

Class	Cname	Deptid	Meets_at	Room	Faculty
_a _b		_x _y			

Condition
_y = 101 CNT.ALL._a > CNT.ALL._b

2. (10 pts) Print the ids and names of the students along with their scores and grades in each course they have enrolled.

Student	Snum	Sname	Deptid	Age
	P._x	P.		

Enrolled	Snum	Cname	Score	Grade
	_x		P.	P.

3. (15 pts) Print the names of the faculty members who instruct at least two courses.

Faculty	Fid	Fname	Deptid
	_x	P.UNQ	

Class	Cname	Deptid	Meets_at	Room	Faculty
_y					_x

Condition
CNT.ALL._y >= 2

4. (10 pts) Print the names of the students who are enrolled in the greatest number of courses. This search cannot be performed by QBE because a subquery would be needed to find the number of the most classes any student is taking, and QBE does not provide subquery functionality.

5. (10 pts) Print the student ids with the maximum average score in the courses they have enrolled in.

Student	Snum	Sname	Deptid	Age
	G.P.UNQ._x			

Enrolled	Snum	Cname	Score	Grade
	_x		P.MAX.ALL	

6. (5 pts) Delete the departments that do not have any faculty or courses at all.

Department	Deptid	Dname	Location
D.	_x		

Faculty	Fid	Fname	Deptid
	_y		_x

Class	Cname	Deptid	Meets_at	Room	Faculty
	_z	_x			

Condition
CNT.ALL._y < 1 AND CNT.ALL._z <1

• Write each of the following queries using Datalog. If you believe any one of the following is not expressible in Datalog, informally explain why it cannot be expressed.

7. (10 pts) Find the faculty who have instructed at least one course in the CS department but have never instructed any course in the Mathematics department.
   V1(F) :- Class(Deptid “CS”, Faculty F), NOT teaches_math(Faculty F)
   teaches_math (Faculty F) :- Class(Deptid “Math”, Faculty F)

8. (10 pts) Find the names of courses each student gets the lowest score in, i.e., if a student takes 2 courses A and B and scores 50 and 60, respectively, then A is the name of the course in which that student (can be different for different students depending on their score) has scored the lowest.
   This query cannot be expressed because it would require an unsafe loop to find the lowest score each student has, which would lead to an infinite loop.

9. (15 pts) What are the possible outcomes if a Datalog rule is unsafe? Please illustrate two different scenarios for this case by designing unsafe rules for the relational schema mentioned above, and explain the answers to these rules.

   not_in_cs(name) :- NOT Student(name, “CS”)
   This rule is unsafe because the variable “name” only appears in a negative literal in the body of the rule

   highest score(snum, score s1, score s2) :- s1 > s2
   This rule is unsafe because the variables “s1” and “s2” are never used in a non-arithmetic positive literal in the body of the rule

   In both instances, the datalog rules will produce an infinite number of answers, with the first rule producing an infinite loop of all students who aren’t in the cs department, and the second rule producing an infinite loop of the highest score