William O’Malley

CS 348

HW 3

**Part A**

Consider the following relational schema:

Student (student-id, student-name, department)

Course (course-id, semester, location, enrolled\_number, instructor, department)

Faculty (faculty-id, faculty-name, department)

Prerequisite (course-id, prereq-course-id)

CourseScore (student-id, course-id, score)

Assume that:

All students and faculty members must major/work in one department; One course can have more than one prerequisites; Scores are numbers in [0,100]; Instructor in Course table is a subset of faculty-id in Faculty table.

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. (10 pts) Print all names of students who have taken both CS348 and CS252.

|  |  |  |
| --- | --- | --- |
| student-id | student-name | department |
|  | P.ALL | \_x |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| course-id | semester | location | enrolled\_number | instructor | department |
| \_y |  |  |  |  | \_x |

|  |
| --- |
| conditions |
| \_y = CS348 AND CS252 |

1. (10 pts) Print all students with their ids and names who have more than the average score of all students of course id ‘CS348’.

|  |  |  |
| --- | --- | --- |
| student-id | student-name | department |
| P. | P. | \_x |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| course-id | semester | location | enrolled\_number | instructor | department |
| CS348 |  |  |  |  | \_x |

|  |  |  |
| --- | --- | --- |
| student-id | course-id | score |
|  | CS348 | \_y |

|  |
| --- |
| conditions |
| \_y > AVG.ALL |

1. (10 pts) Print the course-ids of which have a prerequisite course of other departments.
2. (10 pts) Print the name of the faculty member that teach most unique courses.
3. (10 pts) Delete the courses that have average(over all semesters) enrolled number less than 5.

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.

6. (10 pts) Find all direct and indirect prerequisites of course CS348, i.e. if CS2XX is prerequisite of CS348, and CS1XX is prerequisite of CS2XX, then CS1XX is an indirect prerequisite of CS348.

7. (10 pts) Find the students who have taken at least one course in the CS Department but have never taken any courses in the Statistics Department.

8. (10 pts) For each student, find the courses in which the student has received his/her highest scores. That is, if a student has taken 3 courses and got 60, 70, 80 correspondingly, then the course in which he/she got 80 is the course with the highest personal score.

**Part B**

(20 pts) A Datalog rule is safe if every variable appears in some positive literal. What are possible outcomes if a rule is unsafe? Please illustrate at least two different scenarios that a rule is unsafe by designing unsafe rules for the relational schema in Part A and explain the answers of these rules.