Database Systems, CSCI 4380-01
Homework # 2
Due Thursday September 20, 2018 at 11:59:59 PM

Homework Statement. This homework is worth 5% of your total grade. If you choose to skip it, Midterm #1 will be worth 5% more. Remember, practice is extremely important to do well in this class. I recommend that not only you solve this homework, but also work on homeworks from past semesters. Link to those is provided in the Piazza resources page.

This homework aims to get you practice your relational algebra skills but most importantly work on normalization theory.

Database Description. Suppose you are given the following database for keeping track of grades in this course. The data model from Homework #1 is significantly simplified where all gradable items (hw,quiz, exams) are combined into a single relation. Similarly all grades are also combined into a single relation).

```
students(rin, fname, lname, email, optin_date, optout_date)
gradables(gid, gtype, label, given_date, due_date, maxgrade, points, nextg_id)
grades(rin, gid, submission_date, grade)
```

Each `student` may have an opt-in date, if there is no date (i.e. the value is NULL) then the homeworks are optional for this student. If there is an opt-in date and no opt-out date, then homeworks are required. If there are dates for both opt-in and opt-out, only the homeworks that have a due date within within the given dates are required.

All gradable assignments are stored in `gradables`. The `gtype` is one of `'quiz'`, `'hw'`, `'exam'`, `'finalexam'`. The label is the name given to the gradable as a string such as `'Hw 1'` or `'Exam 2'`. If the gradable is a homework, we store the `gid` of the exam that it is directly before in `nextg_id`.

All grades are stored in `grades` which stores the grade for each student. All date fields are formatted as `mon-day-year`, e.g. `01-31-2016`.

Question 1. Write the following queries using relational algebra using any operator that you wish:

(a) Return the RIN of all students who missed a homework that was due during their opt-in period. Return the `gid` of the corresponding missed homeworks. Remember if there is no opt-out date, all homeworks after opt-in date are required.

(b) Find the RIN, first and last name of all students who had the highest grades for an exam (i.e. gtype `'exam'` or `'finalexam'`). Also return the `gid` and `label` of the exams they got the highest grades in.

Question 2. For each of the following new relations:

(1) list all the relevant functional dependencies based on the explanations below,
(2) find all keys based on your functional dependencies,
(3) discuss whether the relation is in BCNF (Boyce-Codd Normal Form) or not, explain why or why not.
(4) discuss whether the relation is in 3NF (Boyce-Codd Normal Form) or not, explain why or why not.

(a) The system keeps track of multiple submissions for the same homework gradable like submitty in a relation called `submissions`:

`submissions(gid, rin, filename, attemptno, submission_datetime, isactive, totalruntime)`

Each student, gradable and specific attempt corresponds to a specific filename. Each filename corresponds to a specific student, gradable and attempt. For each filename, there is a specific submission_datetime, isactive value and totalruntime value.

(b) Homeworks, quizzes and exams have individual questions. We will store the details of grades of each part separately using a relation called `grade_details`:

`grade_details(rin, gid, partno, topic, maxpoints, pointsearned)`

For each gradable (gid) and part, there is a maxpoints value. For each gradable, part and student, there is pointsearned. Each gradable part may have multiple topics.

Question 3. Given the following relation, functional dependencies and decomposition, answer the following questions:

Relation $R(A, B, C, D, E, F)$ with $\mathcal{F} = \{AB \rightarrow F, BD \rightarrow C, CE \rightarrow F, F \rightarrow D\}$

Decomposition: $R1(A, B, D)$, $R2(A, B, C, E)$, $R3(B, D, E, F)$

(a) Is this decomposition lossless? Show yes or no using Chase algorithm.

(b) Is this decomposition dependency preserving? Show your work.

Note: two show that two sets of functional dependencies, F_1 and F_2 are equivalent, it is sufficient to show that (1) all functional dependencies in F_1 are implied by F_2, and (2) all all functional dependencies in F_2 are implied by F_1.

Question 4. Given the following relation, use BCNF decomposition to convert it to relations in BCNF.

$R(A, B, C, D, E)$ $\mathcal{F} = \{AB \rightarrow C, C \rightarrow E\}$

Question 5. Given the following relation, use 3NF decomposition to convert it to relations in 3NF. For each resulting relation, check if it is also in BCNF.

$R(A, B, C, D, E, F, G)$ $\mathcal{F} = \{AB \rightarrow C, CD \rightarrow EF, CF \rightarrow AG\}$

SUBMISSION INSTRUCTIONS. Submit a PDF document for this homework using Gradescope. No other format and no hand written homeworks please. No late submissions will be allowed.

The gradescope for homework submissions will become available by Tuesday September 18 the latest.