CompSci 316: Intro to Databases

HW-2 : Relational Algebra

Score: 100 (+ 15 extra credit)

Release Date: Tue, 01/21/2020

Due Date: Thurs, 01/30/2020, 11:59 pm

(Part of HW2 appears in Lab1 that is separately graded.)

Please check collaboration and late submission policy from the course website.

Description

For Q1, you will use gradience. **Gradiance** is an online testing system that provides immediate feedback to your answers, and allows you to retry a problem multiple times until you get it right.

Please read the instructions sent on Sakai regarding Gradiance to get started and find the code of the class there. There is no need to turn in anything for these problems; your scores will be tracked automatically.
http://www.newgradiance.com/services/servlet/COTC

For Q2, consider a database “beers” containing information about bars, beers, and drinkers.

drinker(name, address)
bar(name, address)
beer(name, brewer)
frequents(drinker, bar, times_a_week)
likes(drinker, beer)
serves(bar, beer, price)

Write the queries in Q2 in relational algebra using RA, our homegrown relational algebra interpreter built and maintained by Prof. Jun Yang and group (started as a 316 project long back!).

Remember that in RA, you can see list the relations in our small sample database and inspect their contents (by simply issuing a query for a relation with no operators). Other useful features of RA include comments (which allow you to document and explain your queries) and views (which allow you to write complex queries in multiple steps). RA also supports various extensions to standard relational algebra, notably grouping and aggregation (which we will learn later in this course in the context of SQL); **for this homework, however, do NOT use grouping and aggregation features of RA (only select-project-join/cross-diff-rename are allowed).**

Some helpful Documentation: How to use RA

Q0 (5 points)

Set up your VM and RA following these two instructions:

- Readying VM for the course
- Using RA on the VM

To get points, indicate that you have finished both tasks by submitting the corresponding Gradescope assignment.

Q1 (15 points)

Solve the Gradiance problem set HW2 - Q1 (Relational Algebra)

Q2 (80 points)

Write the queries in Q2 in relational algebra using RA, our homegrown relational algebra interpreter. **There must be a semicolon at the end of the query.**

a) Find names of beers that *'Ben'* likes. (10) [Output columns: beer]

b) Find names of all bars that serve Amstel but not Budweiser (10) [Output columns: bar]

c) For each beer that Dan likes, find the names of bars that serve it at the highest price.

   Format your output as list of (beer, bar) pairs (10) [Output columns: beer, bar]

d) Find all beers that is served in exactly 1 bar (10) [Output columns: beer]

e) Find all bars that both Eve and Ben frequent (10) [Output columns: bar]

f) Find drinker pairs *(drinker1, drinker2)* such that the beers liked by *drinker1* is a subset of beers liked by *drinker2.* Format your output as list of *(drinker1, drinker2)* pairs (15)

[Output columns: drinker1, drinker2]

g) Find names of all drinkers who frequent *only* bars that serve *only* beers they like. (15)

[Output columns: name]

Extra credit (15 points)

h) Find names of all drinkers who frequent *only* bars that serve *every* beer they like.

[Output columns: drinker]

How to test your answer on a small database

- Run RA queries on your own VM that you installed in Q0: https://www2.cs.duke.edu/courses/spring20/compsci316/instructions/radb/

How to test your answer on the test database that is used for grading

- Directly submit your answer to autograder (see below) and see whether it is correct or wrong.
- Find the online RA tester on https://ratest.cs.duke.edu/
    - RAtest is a research tool that can show a small counterexample if your answer is wrong showing why it does not work.
    - Note that the use of RAtest is optional, RAtest is a research tool, and it would ask for your consent to collect data for its improvement that would be only used anonymously and in the aggregated form.

Submission Instructions

Submit your solutions on Gradescope by uploading one file for each question. **There must be a semicolon at the end of the query.** Name the files a.txt, b.txt, etc. Each file is a plain text file and only contains the RA query as the answer. Make sure the filename matches your answer.

You need to submit all answers at the same time so you can start with blank files named a.txt, b.txt etc..

If you received any help or collaborate with others, please submit **one more file as collaboration.txt** together with your solutions and put collaborators' names and types of collaborations there. You must mention any students' names who you discussed with (and it is good to mention the course staff’s help as well). If you are unsure what to write, remember that providing more information is better than less information.

Your submission will be auto-graded. The db0 test database it uses is identical to the sample one you have, while db1 is a hidden test database. The order of output rows does not matter but please make sure **the columns are in the correct order**. Multiple re-submissions are allowed.