Assignment #3: Relational Algebra

Release: Jan. 30, 2020   Due: 8:00pm Feb. 5, 2020

Goal

Gain familiarity on Relational Algebra. We will continue with the *P_Mobile* Database Schema of Assignment #2.

Description

- Assume the following relational database schema that supports a cell phone company, *P_Mobile*, that was used in HW2 and shown below:

    - CUSTOMERS = (SSN, fname, lname, <u>cell_pn</u>, home_pn, street, city, zip, state, free_min, DOB, free_SMS)
    - RECORDS (<u>from_pn</u>, to_pn, <u>start_timestamp</u>, duration, type)
    - STATEMENTS (<u>cell_pn</u>, <u>start_date</u>, end_date, total_minutes, total_SMS, amount_due)
    - PAYMENTS (<u>cell_pn</u>, <u>paid_on</u>, amount_paid)
    - DIRECTORY (<u>pn</u>, fname, lname, street, city, zip, state)

- Answer the following questions [for a total of 100 points]:

    1. [5 points each] Assuming that the relations CUSTOMERS and PAYMENTS have 10 and 17 tuples, respectively, find the *arity* and *cardinality* of the following relations: (For those whose accurate values can not be determined, give the min and max values)

        (a) $\pi_{cell\_pn}(Customers)$

        (b) $\pi_{cell\_pn}(Payments)$

        (c) $Customers * Payments$

        (d) $\sigma_{paid\_on='2019-09-01'}(Payments) \bowtie_{Payments.cell\_pn=Customers.cell\_pn} Customers$

    2. [8 points total] Optimize the following relational algebra expression to be more efficient. Please explain your answers.

        (a) $\Pi_{total\_minutes,total\_SMS}(\sigma_{city='philadelphia' \, \wedge \, Statements.cell\_pn=Customers.cell\_pn}$
        $(Statements \times Customers));$

3. [12 points each] Write the *relational algebra* expression to answer each of the following queries in **nesting** notation. You can use any date format: '2020-01-30' or '01-30-2020'.

   (a) List the first and last names of customers who live in Pittsburgh.

   (b) Retrieve the phone numbers of customers who made calls to people in Pittsburgh.

   (c) List the SSNs of all customers that have ever paid more than 100 in a single payment, and have ever had an amount due more than 50.

4. [12 points each] Write the *relational algebra* expression to answer each of the following queries in **sequence** notation:

   (a) List only once every pair of cell phone numbers which use the same number of SMS in July 2019.

   (b) Find the SSNs of all customers who received calls from people in Pennsylvania, where they have at least one call duration more than 20.

   (c) List the SSNs for all customers that live in Pittsburgh city and received calls from New York state, but never made calls to New York state.

To submit your assignment:

1. Create a single file named `hw3-<username>` in PDF (.pdf) or Microsoft Word (.doc) format, containing your answers to all questions. **Do not forget to include your name and username (account name) in the file**.

2. Submit your assignments through the Web-base submission interface you have used for homework #1. **It is your responsibility to make sure the assignment was properly submitted.**

3. Submit your assignment by the due date (**8:00pm February, 5th, 2020**). **There is no late submission.**

4. If you have trouble generating any of the relational algebra symbols, for example R $\bowtie_{A=B}$ S, use a descriptive word, e.g., R Join(A=B) S.

Academic Honesty

The work in this assignment is to be done *independently*. Discussions with other students on the assignment should be limited to understanding the statement of the problem. Cheating in any way, including giving your work to someone else will result in an F for the course and a report to the appropriate University authority.