CS 1555 - Assignment #3: Relational Algebra & Relational Algebra

Gordon Lu

9/24/2020

Problem 1:

[15 points: 5 points each] Assuming that the relations TICKETS and ASSIGNMENT have 6 and 12 tuples, respectively, find the arity or degree and cardinality of the following relations and explain your answer: (For those whose accurate values cannot be determined, give the min and max values)

$$\pi_{\text{machine\_name}}(TICKETS);$$

Solution:

Arity, |R| = 1

Arity is 1 in this case, as we are projecting on a single attribute, `machine_name`.

Cardinality, |r(R)| = 6

Cardinality will be 6, or the cardinality of the `TICKETS` relation. The number of values inside the tuples will just be 1, but we will retrieve the same number of tuples, as we have not applied a filter that would remove any tuples.

$$TICKETS * ASSIGNMENT;$$

Arity, |R| = 12

Arity is 12 in this case, as when we Natural Join the two tables, we will get the sum of the table's attributes, minus the common attributes. In this case, the only commonality was `ticket_number`, thus we take their sum of attributes, 13 and subtract 1 to get 12.

Cardinality, |r(R)| = 12

Cardinality would be 12 in this case as since there is a foreign key relationship, every tuple in the assignment table will be included in the Natural Join.

$$TICKETS \bowtie ASSIGNMENT;$$

Arity, |R| = 12

Arity is 12 in this case, as the Left Outer Join takes the sum of the attributes, and eliminates any redundant common attributes. Thus, our arity here would be the same as the arity of the Natural Join.

Cardinality, min |r(R)| = 12, max |r(R)| = 17

The cardinality in this case is not definitive. The minimum cardinality would be 12, where every ticket number in the Assignments table would match to a ticket number in the Tickets table. The maximum cardinality on the other hand would be when there is a mismatch for every ticket number between Tickets and Assignments. Due to the foreign key relationship, there has to be at least one match, so the maximum cardinality occurs when one ticket is assigned to 12 people, thus the cardinality would be 12 tuples, and the 5 tuples from the left over tuples in the Tickets table, making the cardinality 17.

Problem 2:

Write the relational algebra expression to answer each of the following queries. Use the nesting notation.

a) [7 points] List the first and last names of all the Tech personnel whose expertise is hardware.

$$\pi_{fname,lname}(\sigma_{expertise='hardware'}(TECH\_PERSONNEL))$$

b) [7 points] List the number of machines located in Sennott Square ('SENSQ') building

$$f_{COUNT\ machine\_name}(\sigma_{location\_name="SENSQ"}(INVENTORY * LOCATIONS))$$

Problem 3:

Write the relational algebra expression to answer each of the following queries. Use the sequence notation. For dates, use the DD-MON-YYYY format.

a) [9 points] List all the ticket information for all the tickets with status assigned which have been assigned to Bob Hoffman.

$$Hoffman\_PSID \leftarrow \pi_{pplSoft}(\sigma_{fname=\text{Bob},lname="Hoffman"}(TECH\_PERSONNEL))$$

$$BH\_ASSIGN \leftarrow ASSIGNMENT * Hoffman\_PSID$$

$$Hoffman_{TICKETS} \leftarrow TICKETS \bowtie_{TICKETS.ticket\_number=BH\_ASSIGN.ticket\_number} BH\_ASSIGN$$

$$ALL_{BH_{TICKS_{INFO}}} \leftarrow \pi_{\text{ticket number,owner pplSoft,date submitted,date closed,days worked on,category id,machine name,description}}(\sigma_{status="assigned"}(Hoffman_{TICKETS}))$$

b) [9 points] List all the users' pplSoft numbers along with the most recent submitted ticket number(s) within the month December 2019.

$$USER\_TICKS \leftarrow USERS \bowtie_{USERS.pplSoft=TICKETS.owner\_pplSoft} TICKETS$$

$$TICKS\_IN\_DEC \leftarrow \sigma_{date\_submitted \geq "01\text{-}DEC\text{-}2019" \wedge date\_submitted \leq "31-DEC-2019"}(USER\_TICKS)$$

$$RSLT \leftarrow \pi_{pplSoft,date\_submitted}(TICKS\_IN\_DEC)$$

c) [9 points] List the first and last names of users who submitted more than 5 tickets during the month of February 2019.

$$USER\_TICKS \leftarrow USERS \bowtie_{USERS.pplSoft=TICKETS.owner\_pplSoft} TICKETS$$

$$TICKS\_IN\_FEB(\leftarrow \sigma_{TICKETS.date\_submitted \geq \text{"01-FEB-2019"} \wedge TICKETS.date\_submitted < \text{"01}-MARCH-2019\text{"}}(USER\_TICKS)$$

$$COUNTS(pplSoft, num\_tickets) \leftarrow TICKETS.owner\_pplSoft\ f_{COUNT\ ticket\_number}(TICKS\_IN\_FEB)$$

$$\text{JOIN\_ME} \leftarrow TICKS\_IN\_FEB \bowtie_{TICKETS.owner\_pplSoft=COUNTS.pplSoft} COUNTS$$

$$RSLT \leftarrow \pi_{fname,lname}(\sigma_{JOIN\_ME.num\_tickets > 5}(JOIN\_ME))$$

d) [11 points] Find the tech personnel who has contributed the most in solving problems (i.e., all tech staff who worked on resolved tickets more than any other tech personnel) and list their first and last names.

$$RESOLVED\_TICKETS \leftarrow \sigma_{status=\text{closed\_unsuccessful} \vee status=\text{closed\_sucessful}}(ASSIGNMENT)$$

$$SUCCESSFUL\_TECH(success_{pplSoft}, num\_resolved) \leftarrow tech\_pplSoft f_{COUNT\ status}(RESOLVED\_TICKETS)$$

$$MAX\_closed(\text{max\_close}) \leftarrow f_{MAX\ num_{resolved}}(SUCCESSFUL\_TECH)$$

$$RSLT \leftarrow \pi_{fname,lname(}\pi_{success\_pplSoft}(SUCCESSFUL\_TECH \bowtie_{num\_resolved=\text{max}\ \_close}\ (MAX\_CLOSE)) * TECH\_PERSONNEL)$$

e) [11 points] Identify the most problematic machines during the months of June and August 2019 and list as "Device Name" the names of these machines (Hint: these machines had the maximum number of tickets in the given period).

$$SUB\_DATE \leftarrow \sigma_{date\_submitted \geq \text{"01}-JUN-2019\text{"} \wedge date\_submitted \leq \text{"31}-AUG-2019\text{"}}(TICKETS)$$

$$NUM\_MACHINES(mname, ticks) \leftarrow machine\_name\ f_{COUNT\ ticket_{number}}(SUB\_DATE)$$

$$MAX\_MACHINE(\text{ticks}) \leftarrow f_{MAX\ ticks}(NUM\_MACHINES)$$

$$RES(Device\ Name) \leftarrow \pi_{machine_{name}}(NUM\_MACHINES * MAX\_MACHINE)$$

f) [11 points] Find the expertise of each tech staff by listing the most frequent categories of tickets assigned to them.

$$ASSIGNED\_TECH\_TICKETS \leftarrow \pi_{tech_{pplSoft}, category_{id}}(TICKETS * \pi_{ticket\_number, tech\_pplSoft}(\sigma_{status="assigned"}(ASSIGNMENT)))$$

$$NUM\_ASSIGNED(assigned\_pplSoft, category, num\_tickets) \leftarrow tech\_pplSoft, category\ f_{COUNT\ category}(CATEGORIES * ASSIGNED\_TECH\_TICKETS)$$

$$HIGHEST(maxTickets) \leftarrow assigned\_pplSoft\ f_{MAX\ num\_tickets}(NUM\_ASSIGNED)$$

$$PPLSOFT\_WITH\_MAX \leftarrow NUM\_ASSIGNED\ \theta_{NUM\_ASSIGNED.assigned\_pplSoft=HIGHEST.assigned\_pplSoft \wedge maxTickets=num\_tickets} HIGHEST$$

$$MAX\_EXPERTISE(pplSoft, category) \leftarrow \pi_{NUM\_ASSIGNED.assigned\_pplSoft, category}(\text{PPLSOFT\_WITH\_MAX})$$

$$RSLT(fname, lname, expertise) \leftarrow \pi_{fname, lname, category}(MAX\_EXPERTISE * TECH\_PERSONNEL)$$

g) [11 points] For all the possible unique categories of submitted tickets during September 2019, list the first name, last name, and pitt ID of users who submitted tickets for all unique categories during that month.

$$SUBMIT \leftarrow \sigma_{date\_submitted \geq "01-SEPT-2019" \wedge date\_submitted \leq "30-SEPT-2019"}(TICKETS)$$

$$TECH\_SUBMIT \leftarrow \pi_{owner\_pplSoft, category\_id}(SUBMIT)$$

$$NUM\_TECHS(category\_id, num\_techs) \leftarrow category\_id\ f_{COUNT\ count\_id}(TECH\_SUBMIT)$$

$$RSLT \leftarrow \pi_{fname, lname, pittID}(\pi_{owner\_pplSoft}(\sigma_{num_{techs}=1}(TECH\_SUBMIT * NUM\_TECHS)) * USERS)$$