CS 1555 HW 3

Mathew Varughese - mav120

Wednesday, Feb 5

I assume that Payments.cell_pn is a foreign key to the primary key Customers.cell_pn .

1. (a) Arity = 1 Cardinality = 10

   (b) Arity = 1 Min Cardinality = 1 Max Cardinality = 10

   (c) Arity = 14 Min Cardinality = 10 Max Cardinality = 17

   It depends on what the foreign key of cell_pn comes from. I assume the cell_pn attribue in PAYMENTS is a foreign key to CUSTOMERS.

   (d) Arity = 15 Min Cardinality = 0 Max Cardinality = 10

   Because cell_pn and paid_on are a unique key, you cannot have multiple cell_pn with the same date of '2019-09-01', so at the maximum, you can have 10 rows on PAYMENTS that have different cell phone numbers, that join with cell_pn on CUSTOMERS.

2. (a) $\Pi_{total\_minutes, total\_SMS}(\sigma_{city='philadelphia}(Customers) * Statements)$

   By selectivity, we can reduce the rows in the Customers, which will make the natural join with Statements more efficient. The time of the projection does not matter for time complexity.

3. (a) $\Pi_{fname,lname}(\sigma_{city='Pittsburgh'}(Customers))$

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

   $\Pi_{cell\_pn}(Customers \bowtie_{cell\_pn=from\_pn} ((\sigma_{city='Pittsburgh'}Directory) \bowtie_{Directory.pn=Records.to\_pn} RECORDS))$

   (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.

   $\Pi_{SSN}(Customers * (\sigma_{amount\_paid>100}(Payments) * \sigma_{amount\_due>50}(Statements)))$

4. (a) List only once every pair of cell phone numbers which use the same number of SMS in July 2019. (I am assuming that ≤ '2019-07-31' can find all records before the last day in July)

   $JULY \leftarrow \sigma_{start\_date \geq '2019-07-01' \wedge end\_date \leq '2019-07-31'}Statements$

   $JSMS \leftarrow \Pi_{cell\_pn,total\_SMS}(JULY)$

   $JSMS2 \leftarrow \rho_{cell\_pn2,total\_SMS}(JSMS)$

$BOTH \leftarrow JSMS * JSMS2$

$REMOVEDUPS \leftarrow \sigma_{cell\_pn>cell\_pn2}BOTH$

$RSLT \leftarrow \Pi_{cell\_pn,cell\_pn2}REMOVEDUPS$

(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.

Assuming this means, a customer is shown if they have received a call from Pennsylvania and at least one of those calls was longer than 20.

$PA \leftarrow (\sigma_{state='PA'}(Directory))$

$DUR \leftarrow (\sigma_{duration>20}(Records))$

$PARECORDS \leftarrow PA \bowtie_{Directory.pn=Records.from\_pn} DUR$

$CUS \leftarrow PARECORDS \bowtie_{Records.to\_pn=Customers.cell\_pn} CUSTOMERS$

$RSLT \leftarrow \Pi_{SSN}(CUS)$

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

$PGH \leftarrow \sigma_{city='Pittsburgh'}(Customers)$

$TO\_PGH \leftarrow PGH \bowtie_{Customers.cell\_pn=Records.to\_pn} Directory$

$FROM\_PGH \leftarrow PGH \bowtie_{Customers.cell\_pn=Records.from\_pn} Directory$

$NY \leftarrow \sigma_{state='NY'}Directory$

$NOT\_NY \leftarrow \sigma_{state \neq 'NY'}Directory$

$FROMNY \leftarrow TO\_PGH \bowtie_{Records.from\_pn=Directory.pn} NY$

$TO\_NOT\_NY \leftarrow FROM\_PGH \bowtie_{Records.to\_pn=Directory.pn} NOT\_NY$

$SSN\_FROMNY \leftarrow \Pi_{SSN}FROMNY$

$SSN\_TO\_NOT\_NY \leftarrow \Pi_{SSN}TO\_NOT\_NY$

$RSLT \leftarrow SSN\_FROMNY \cap SSN\_TO\_NOT\_NY$