CS 1555/2055: Database Management Systems (Fall 2018)

Department of Computer Science, University of Pittsburgh

Midterm Review

We will use the following relational model:

- **Employee** (SSN, fname, lname, DOB, address)
  PK(SSN)
- **Project** (PNumber, Location)
  PK(PNumber)
- **Assignment** (SSN, PNumber)
  PK(SSN, PNumber)
  FK(SSN) $\rightarrow$ Employee(SSN)
  FK(PNumber) $\rightarrow$ Project(PNumber)

1. Assuming that the relations **Employee**, **Project** and **Assignment** have 50, 5 and 100 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). Recall that $\bar{\bowtie}$ is the left outer natural join operator and $\bowtie\bar{}$ is the right outer natural join operator.

   $a)$ $Employee \bowtie\bar{} \; Assignment$
   Arity: 6
   Cardinality: 100

   $b)$ $Employee \; \bar{\bowtie} \; Assignment$
   Arity: 6
   Cardinality: Min = 100; Max = 149

2. Consider the database schema mentioned in the description, write the **relational algebra** expression(s) in **SEQUENCE notation** for each of the following queries.

   $a)$ List the first and last names of all the employees who have been assigned to the same projects as Mike Smith.

   $$Mike\_SSN(SSN) \leftarrow \Pi_{SSN}(\sigma_{fname='Mike' \; \wedge \; lname='Smith'}(Employee))$$

   $$Mike\_Proj(PNumber) \leftarrow \Pi_{PNumber}(Mike\_SSN * Assignment)$$

   $$Coworker\_SSN(SSN) \leftarrow Assignment \div Mike\_Proj$$

   $$RESULT \leftarrow \Pi_{fname,\, lname}(Coworker\_SSN * Employee)$$

b) List the first and last names of the employees who have been assigned the max number of projects.

$Proj\_Count(SSN, PCount) \leftarrow {}_{<SSN>}\mathcal{F}_{<\text{COUNT } PNumber>} Assignment$

$Highest\_Count(MaxCount) \leftarrow \mathcal{F}_{<\text{MAX } PCount>} Proj\_Count$

$Highest\_Proj\_Count(SSN, PCount) \leftarrow Proj\_Count \bowtie_{PCount=MaxCount} Highest\_Count$

$RESULT \leftarrow \Pi_{fname,\, lname}(Highest\_Proj\_Count * Employee)$

3. Express in SQL each of the following queries. For the time, use the standard format 'YYYY-MM-DD HH:mm:ss'.

   a) List the SSN, first name, last name and project count of employees who were born later than Year 1990 and have been assigned to more than 2 projects. List the result in descending order of project count.

```
SELECT E.SSN, E.FNAME, E.LNAME, COUNT(A.PNUMBER)
FROM EMPLOYEE E JOIN ASSIGNMENT A ON E.SSN=A.SSN
     JOIN PROJECT P ON A.PNUMBER=P.PNUMBER
WHERE E.DOB>='1991-01-01 00:00:00'
GROUP BY E.SSN
HAVING COUNT(A.PNUMBER) > 2
ORDER BY COUNT(A.PNUMBER);
```

   b) List the SSN of employees who work in a City whose name's the third letter is P and contains at least two P's.

```
SELECT SSN
FORM (EMPLOYEE E NATURAL JOIN ASSIGNMENT) NATURAL JOIN PROJECT
WHERE LOCATION LIKE '__P%P%';
```