CSCI 6333/6315 Database Systems
Test 1 – Sample Solutions

1. [80 points] Consider the employee database of Fig. 1, where the primary keys are underlined. For each of the following questions, give (1) a relational algebra expression, (2) a tuple relational calculus expression, (3) a domain relational calculus expression, and (4) and SQL query.
    a. Find the highest paid managers in every city in Texas. That is, you shall generate a relation *highPaidManagersInTexas(managerName, maxSalary, companyName, city).*

employee(employeeId, employeeName, streetNumber, street, city, state)
works(employeeId, companyId, salary)
company(companyId, companyName, city, state)
manages(employeeId, managerId)

Figure 1. Employee database

Solution:

a. Find the highest paid managers in every city in Texas. That is, you shall generate a relation

highPaidManagersInTexas(managerName, maxSalary, city).

a.1. A relational algebra expression

//find manager, city, company, salary

$$r_1 \leftarrow \prod_{\substack{E.city, E.employeeName, \\ W.salary}} (\sigma_{\substack{E.employeeId=W.employeeId \wedge E.state='Texas' \\ \wedge E.employeeId=M.managerId \\ \wedge W.companyId=C.companyId}} (\rho_E(employee) \times \rho_W(works) \times \rho_C(company) \times \rho_M(manages)))$$

//find max manager salary in every city in Texas

$$r_2 \leftarrow_{city} G_{max(salary)\ \text{as}\ maxSalary,\ \text{city}}(r_1)$$

//find highest paid managers in every city in Texas

$$\prod_{\substack{\text{employeeName as ManagerName}, \\ \text{maxSalary,compamayName} \\ r_1.\text{city}}} (\sigma_{r_1.\text{city}=r_2.\text{city} \wedge r_1.\text{salary}=r_2.\text{maxSalary}}(r_1 \times r_2))$$

a.2. A tuple relational calculus expression

$$\{t|\exists s \in \text{employee}, \exists u \in \text{works}, \exists w \in \text{company}, \exists m \in \text{manages}$$
$$(t[\text{managerName}] = s[\text{employeeName}] \wedge s[\text{state}] = ' \text{Texas}'$$
$$\wedge t[\text{city}] = s[\text{city}] \wedge t[\text{companyName}] = w[\text{companyName}]$$
$$\wedge t[\text{maxSalary}] = u[\text{salary}] \wedge s[\text{employeeId}] = u[\text{employeeId}]$$
$$\wedge u[\text{companyId}] = w[\text{companyId}]$$
$$\wedge s[\text{employeeId}] = m[\text{managerId}])$$
$$\wedge \forall q \in \text{employee}, \forall h \in \text{works}$$
$$(t[\text{city}] = q[\text{city}] \wedge q[\text{employeeId}] = h[\text{employeeId}]$$
$$\Rightarrow t[\text{salary}] \geq h[\text{salary}])\}$$

a.3. A domain relational calculus expression

$$\{< mc, mn, mcn, ms > | \exists xid, \exists xstn, \exists xst, \exists xcid, \exists yid, \exists yc$$
$$(< xid, mn, xstn, xst, mc, \text{Texas} > \in employee \wedge$$
$$< xid, xcid, ms > \in works \wedge < xcid, mcn, yc > \in company$$
$$\wedge < yid, xid > \in manages)$$
$$\wedge \forall \text{fid}, \forall \text{fn}, \forall \text{fstn}, \forall \text{fst}, \forall fs, \forall \text{fcid}$$
$$(< fid, fn, fstn, fst, mc > \in employee$$
$$\wedge < \text{fid,fcid,fs} > \in \text{works}$$
$$\Rightarrow ms \geq fs) \}$$

a.4. An SQL query

```
//find manager name, city, company name, salary
with r1 (managerName, salary, companyName, city) as
    select E.city, E.employee_name, W.salary, C.company_name
    from emplyee as E, works as W, companay as C, manages as M
    where E.employeeId = W.employeeId
        ∧ E.employeeId = M.managerId ∧ W.companyId
        = C.companyId ∧ E. state = 'Texas'

//find the max manager salary in every city in Texas
with r2 (maxSalary, city) as
    select max(salary), city
    from r1
    group by city

//find the highest paid managers in every city in Texas
select managerName, maxSalary, companyName, r1.city
from r1, r2
where r1.city = r2.city and r1.salary = r2.maxSalary
```

2. [20 points] Consider the employee database of Fig. 1, where the primary keys are underlined. Write an SQL query to find the highest paid families in every city. Assume that the total income of a family is the sum of salaries of employees living at the same address. Note that you shall generate a relation
highPaidManagerFamiliesInTexas(streetNumber, street, city, familyIncome,).

Answer: An SQL query is given below

```
//find manger addresses in Texas
with m(streetNumber, street, city) as
        select E.streetNumber, E.street, E.city
        from employees as E, manages as C
        where E.employeeId = manages.managerId and E.state = 'Texas'
//find salary per address in Texas
with r1(salary, street_number, street, city ) as
        select W.salary, E.street_number, E.street, E.city
        from employees as E, works as C
         where E.employeeId = works.employeeId and E.state = 'Texas'
//find manager family incomes in Texas
with r2(familyIncome, streetNumber, street, city) as
       select sum(salary), streetNumber, street, city
       from r1 natural join m
       group by streetNumber, street, city
//find max manager family incomes in every city in Texas
with r3(maxFamilyIncome, city) as
       select max(familyIncome) , city
       from r2
      group by city
//find highest paid manager families in every city in Texas
 select streetNumber, streetName, City, maxFamilyIncome as familyIncome
       from r2, r3
      where r2.city = r3.city
              and r2.familyIncome = r3.maxFamilyIncome
```