CS3402 Database Systems

Homework

Question A. (50 marks)

A database schema consisting of three relations STUDENT, COURSE, and STAFF is created as follows:

```
CREATE TABLE STUDENT (STU_ID CHAR(4),
                      STUDENT_NAME CHAR(20),
                      ADDRESS CHAR(20),
                      BIRTHDATE DATE,
                      GENDER CHAR(6));

CREATE TABLE COURSE (COURSE_ID CHAR(6),
                     COURSE_TITLE CHAR(20),
                     STAFF_ID CHAR(3),
                     SECTION NUMBER(2));

CREATE TABLE STAFF (STAFF_ID CHAR(3),
                    STAFF_NAME CHAR(20),
                    GENDER CHAR(6),
                    DEPARTMENT CHAR(20),
                    BOSS_ID CHAR(3)
                    SALARY NUMBER(8,2));
```

Write down a SQL statement for each query below:

1) List the names of all male students who were born before 01-01-1995. [5 marks]

```
select student_name from student where gender='m' and birthdate < '01-JAN-1995';
```

Note: using the date format 01-01-1995 is also fine.

2) List the names of all students whose name is at least 6 characters long and whose birthdate falls within 01-01-1995 and 01-01-2000. Order the results alphabetically. [5 marks]

```
SELECT student_name FROM student WHERE length(student_name) >= 6 AND
birthdate BETWEEN '01-JAN-1995' and '01-JAN-2000'
```

3) List the names of all courses that are neither taught by SMITH nor by JONES. [6 marks]

```
1 SELECT course_title from COURSE
2 WHERE STAFF_ID NOT IN
3 (SELECT staff_ID
4   FROM staff
5   WHERE staff_name IN ('JONES','SMITH'))
```

4) Retrieve the number of courses for each section number. Your output should be in ascending order of section numbers. [6 marks]

```
SELECT section, COUNT(*)
FROM course
GROUP BY section
ORDER BY section
```

5) Find the name of every staff member who teaches the exact same number of courses as his/her boss. [7 marks]

```
SELECT staff_name
FROM STAFF S1
WHERE
(SELECT COUNT (*) FROM COURSE WHERE STAFF_ID=S1.STAFF_ID)
=
(SELECT COUNT (*) FROM COURSE WHERE STAFF_ID=S1.BOSS_ID);
```

6) List all staff members whose salary is higher than the average salary. [7 marks]

```
SELECT *
FROM staff
WHERE salary > (SELECT AVG(SALARY) FROM staff);
```

7) Find the names of the staff members who are the boss of some staff member in the Accounting department. [7 marks]

```
SELECT staff_name
FROM staff
WHERE staff_id IN
(SELECT boss_id
 FROM staff
 WHERE department='Accounting');
```

8) Find the average number of courses taught per staff member. [7 marks]

```
SELECT COUNT(DISTINCT course_id) / COUNT(DISTINCT staff_id)
FROM course;
```

Note: It is fine to write COUNT(course_id), i.e. without 'DISTINCT' for the first count expression assuming that course_id is a key attribute of course. Also, it is fine to divide by the total number of staff members (i.e. even the ones that do not teach any courses.)

Question B. [50 marks]

Specify the following queries on the COMPANY relational database schema shown in Figure 5.5 below using relational algebra expressions. Also show the result of each query as it would apply to the database state of Figure 5.6. (Note that some of the queries may return an empty result set.)

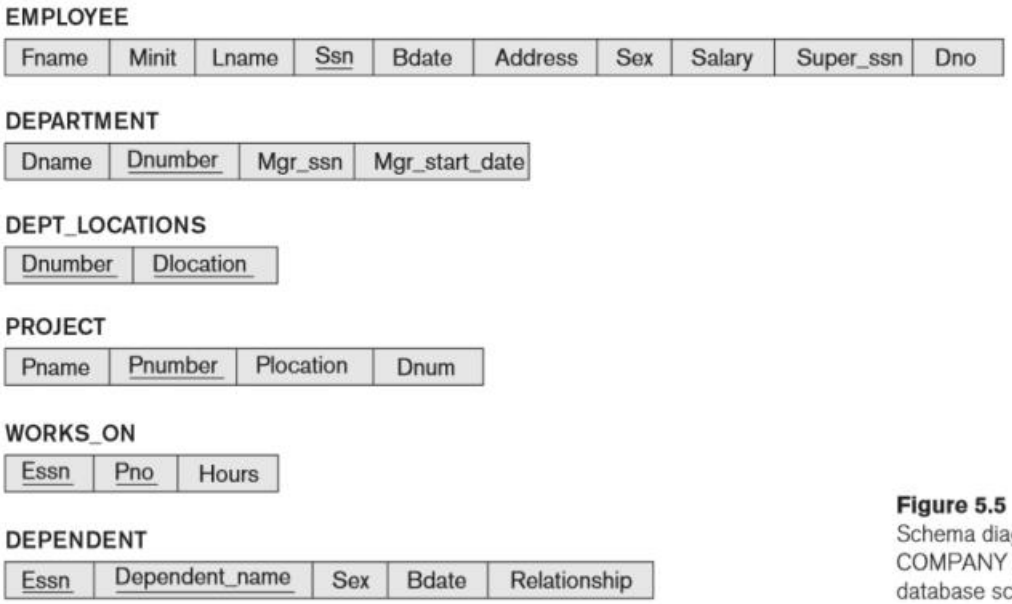

Figure 5.5
Schema diagram for the COMPANY relational database schema.

Figure 5.6
One possible database state for the COMPANY relational database schema.

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

(a) Retrieve the names of employees that have at least one dependent. [5 marks]

$\pi_{\text{Fname,Minit,Lname}}$ (EMPLOYEES $\bowtie_{\text{Essn = Ssn}}$ DEPENDENT)

Result: John B Smith, Franklin T Wong, Jennifer S Wallace

(b) List the first names of employees who work in the Research department and whose salary is below 39000. [5 marks]

π Fname (σsalary<39000 AND Dname='Research' EMPLOYEE $\bowtie$Dno=Dnumber DEPARTMENT)

Results: John, Ramesh, Joyce

(c) Find the first names of male employees that are not supervised by 'Jennifer Wallace'. [5 marks]

$Jennifer \leftarrow \sigma_{Fname='Jennifer' \; AND \; Lname='Wallace'}(EMPLOYEE)$
$Jennifer\_emp \leftarrow \pi_{EMPLOYEE.Fname}(EMPLOYEE \bowtie_{Super_{ssn}=Ssn} Jennifer)$
$Male\_emp \leftarrow \pi_{Fname}(\sigma_{Sex='M'}(EMPLOYEE))$
Result $\leftarrow Male\_emp - Jennifer\_emp$

Results: John, Franklin, Ramesh, James

(d) List the first names of all employees who work on the project 'Newbenefits' and who have a daughter as a dependent. [5 marks]

Relation $\leftarrow \sigma$ relationship='Daughter' (EMPLOYEE $\bowtie$Essn=Ssn DEPENDENT)

Employees$\leftarrow \sigma$Pname='Newbenefits' (EMPLOYEE $\bowtie$Dno=Dnum PROJECT)

π Fname (Relation $\cap$ Employees)

Results: (empty)

(e) Retrieve the first names of employees who work on every project. [5 marks]

Ssn_pno $\leftarrow \rho_{(Ssn,\ Pnumber)}$ (πEssn,Pno ((WORKS_ON))

PNos $\leftarrow$ πPnumber (PROJECT)

SSn_all_pnos $\leftarrow$SSn_pno $\div$ PNos

πFname (EMPLOYEE * SSn_all_pnos)

Result: (empty)

(f) Retrieve the first names of employees who do not work on any project. [7.5 marks]

ALL_EMPS <-- πSsn (EMPLOYEE)

WORKING_EMPS<-- ρSsn (πEssn (WORKS_ON))

NON_WORKING_EMPS <-- ALL_EMPS - WORKING_EMPS

πFname (EMPLOYEE * NON_WORKING_EMPS)

Result: (empty)

(g) Retrieve the salary of all male employees working in the department with department number 4. [5 marks]

π salary (σ sex='M' and DNo=4 (EMPLOYEE))

Results: 25000

(h) Find the first names and addresses of employees who work on at least one project located in Houston but whose department has its location in either Sugarland or Bellaire. [7.5 marks]

Houston_projects $\leftarrow \sigma_{\text{PLocation='Houston'}}$ (PROJECTS $\bowtie_{\text{PNumber=PNo}}$ WORKS_ON)

Dptmts $\leftarrow \sigma_{\text{DLocation='Sugarland' or DLocation='Bellaire'}}$ (DEPARTMENT * DEPT_LOCATION)

Employee_Dptmts $\leftarrow$ Dptmts $\bowtie_{\text{DNo=DNumber}}$ EMPLOYEE

$\pi_{\text{Fname,Address}}$ (Houston_projects $\bowtie_{\text{Essn =Ssn}}$ Employee_Dptmts)

Results:
Ramesh, 975 Fire Oak, Humble TX
Franklin, 638 Voss, Houston TX

(i) List the last name of each department manager who has a spouse as a dependent. [5 marks]

Mgr_ssn_spouse $\leftarrow \sigma_{\text{relationship='Spouse'}}$ (DEPARTMENT $\bowtie_{\text{Mgr\_ssn=Ssn}}$ DEPENDENT)

π_{Lname} (Mgr_ssn_spouse $\bowtie_{\text{Mgr\_ssn=Essn}}$ EMPLOYEE)

Results: Wallace, Wong