Exercise-5: Solution SQL-1 DBS

1. The schema of a database containing university-type data is given below.

Primary key is bold for each relation.

STUDENT(**Sid**, Sname, Sex, Age, Year, GPA)
DEPT(**Dname**, Numphds)
PROF(**Pname**, Dname)
COURSE(**Cno**, Cname, Dname)
MAJOR(**Dname, Sid**)
SECTION(**Dname, Cno, Sectno**, Pname)
ENROLL(**Sid**, Grade, **Dname, Cno, Sectno**)

Write the following queries in SQL.

(i) Find the names of professors who work in departments that have fewer than 50 PhD students.

Solution:

```
SELECT    Pname
FROM      PROF, DEPT
WHERE     DEPT.Dname = PROF.Dname AND Numphds < 50;
```

(ii) Find the names and majors of students who have taken the 'Database System' course.

Solution:

```
SELECT  Sname, Dname
FROM    COURSE C, ENROLL E, MAJOR M, STUDENT S
WHERE   C.Cname = 'Database Systems'
AND     C.Dname = E.Dname
AND     C.Cno = E.Cno
AND     E.Sid = M.Sid
AND     E.Sid = S.Sid;
```

(iii) Find the ids, names, and GPAs of the students who have taken all courses from the 'Civil Engineering' department.

Solution:

```
SELECT    Sid, Sname, GPA
FROM      STUDENT S
WHERE     NOT EXISTS
          (SELECT C.CID
          FROM   COURSE C
          WHERE Dname = 'Civil Engineering'
          EXCEPT
          SELECT E.CID
          FROM ENROLL E
          WHERE Dname = 'Civil Engineering' AND E.Sid = S.Sid);
```

2. Suppose we are maintaining a database of articles published in our newspaper, the Straits Times. We have the following schema (where keys are underlined):

Article (issueID, articleID, author, title)

Citation (articleID, issueID, citedArticleID, citedIssueID)

WordAppears (wordID, issueID, articleID, position)

WordIs (wordID, wordText)

Issue (issueID, date, howManyDistributed)

Assume that dates can be compared using comparison operators (<, >, =). Assume that position is an index specifying where the word appears (1 = first word, 2 = second, etc.). Write the following query in SQL.

Find the documents in which the words "politician" and "corruption" appear.

Solution:

```
SELECT    DISTINCT wa1.issueID, wa1.articleID
FROM      WordAppears wa1, WordIs wi1, WordAppears wa2, WordIs wi2
WHERE     wa1.issueID = wa2.issueID AND wa1.articleID = wa2.articleID
          AND wa1.wordID = wi1.wordID AND wa2.wordID = wi2.wordID
          AND wi1.wordText = 'politician' AND wi2.wordText = 'corruption';
```

3. Consider the relation R(A,B,C,D) with candidate keys AC and D. What will be the output of the following query? Justify your answer.

```
SELECT A, B
FROM R
WHERE C > (SELECT D FROM R WHERE C = 3);
```

Solution:

This will be an error. The inner query is a row subquery and it can return multiple D's. Hence, there is a mismatch of operators.

4. Let R=(A, B, C), S=(C, D, E) be two relational schema. Let q and r be relations (i.e., tables) on schema R; and s be a relation (i.e., a table) on schema S. Convert the following relational algebra queries to SQL.

(i) $q - r$

(ii) $\Pi_{A, C}(r) \bowtie \Pi_{C, D}(s)$

Solution:

(i)

```
SELECT * FROM q
EXCEPT
SELECT * FROM r;
```

(ii)

```
SELECT r.A, r.C, s.D
FROM r, s
WHERE r.C = s.C;
```

5. The schema of a database containing university-type data is given below.

Primary key is bold for each relation.

STUDENT(**Sid**, Sname, Sex, Age, Year, GPA)
DEPT(**Dname**, Numphds)
PROF(**Pname**, Dname)
COURSE(**Cno**, Cname, Dname)
MAJOR(**Dname, Sid**)
SECTION(**Dname, Cno, Sectno**, Pname)
ENROLL(**Sid**, Grade, **Dname, Cno, Sectno**)

Write the following query in SQL. Find the name(s) of student(s) with the lowest GPA.

Solution:

```
SELECT    Sname
FROM      STUDENT
WHERE     GPA  IN
(SELECT MIN(GPA)
FROM STUDENT);
```

6. Consider the following relational schema:

Reader(RDNR, Surname, Firstname, City, Birthdate)
Book(ISBN, Title, Author, NoPages, PubYear, PublisherName)
Publisher(PublisherName, PublisherCity)
Category(CategoryName, BelongsTo)
Copy(ISBN, CopyNumber, Shelf, Position)
Loan(ReaderNr, ISBN, Copy, ReturnDate)
BookCategory(ISBN, CategoryName)

BelongsTo refers to which parent categories the current category belongs to. Each book has a specific ISBN, and many copies of a book might be available under the same ISBN. A reader may borrow the same copy for multiple times, and each instance is recorded by its ReturnDate. All the parent categories that a book belongs to are stored in the table BookCategory.

Formulate the following queries in SQL.

(a) Which categories do not have any subcategories?

Solution:

```
SELECT C1.CategoryName
FROM Category C1
WHERE NOT EXISTS
(SELECT CategoryName
FROM Category C2
WHERE C2.BelongsTo = C1.CategoryName);
```

(b) For which of the books there is at least one copy available?

Solution:

```
SELECT Title
FROM Book
WHERE ISBN IN
(SELECT ISBN FROM
((SELECT CopyNumber, ISBN FROM Copy)
EXCEPT
(SELECT Copy, ISBN FROM Loan)));
```

(c) Which books have more pages than twice the average of the number of pages of all books?

Solution:

```
SELECT ISBN
FROM Book
WHERE NumberOfPages>= 2* (SELECT AVG(NumberOfPages)
                          FROM Book);
```

(d) What are the surnames of the readers from the city "New York"?

Solution:

```
SELECT DISTINCT Surname
FROM Reader
WHERE City = 'New York'
```

7. For the following relational schema:

employee (employee-name, street, city)

works (employee-name, company-name, salary)

company (company-name, city)

manages (employee-name, manager-name)

Give an expression in SQL for each of the following queries:

(a) Find the names of all employees who earn more than the average salary of all employees of their company. Assume that all people work for at most one company.

Solution:

```
SELECT employee-name
FROM works t
WHERE salary > (SELECT AVG(salary)
                FROM works s
                WHERE t.company-name = s.company-name);
```

(b) Find the names of all employees in the database who live in the same cities and on the same streets as do their managers. Assume that all people work for at most one company. Each company has at most one manager, who is also an employee of the same company.

Solution:

```
SELECT p.employee-name
FROM employee p, employee r, manages m
WHERE p.employee-name = m.employee-name
AND m.manager-name = r.employee-name
AND p.street = r.street AND p.city = r.city;
```

8. Consider the following schema containing airport flight information. Primary Keys are in bold.

FLIGHTS(**flno:integer**, from:string, to:string, distance:integer, departs:time, arrives:time)

AIRCRAFT(**aid:integer**, aname:string, cruisingrange:integer)

CERTIFIED(**eid:integer**, **aid:integer**)

EMPLOYEES(**eid:integer**, ename:string, salary:integer)

Note that the Employees relation describes pilots and other kinds of employees as well; every pilot is certified for some aircraft (otherwise, he orshe would not qualify as a pilot), and only pilots are certified to fly.

Give an SQL expression for the following query. Your solution should be only one SQL statement.

Find the eids of employees who make the second highest salary

```
SELECT E.eid
FROM   Employees E
WHERE  E.salary = (SELECT MAX (E2.salary)
                   FROM   Employees E2
                   WHERE  E2.salary ≠ (SELECT MAX (E3.salary)
                                       FROM   Employees E3 ))
```