CS482/CS502 – Test 1

8:55pm – 10:10 pm

November 11, 2005

Note:

- There are 5 questions (100 points for 25% of the total grade).
- Be short and precise in your answers.
- There are some questions that are different for **CS482** and **CS502**. Be sure that you do the correct ones.

Name:

Signature:
Registered for:

1. (20 points) Consider the following relation schema:

```
Student(sid : integer, major : string, sname : string ,
                              dob : date, gpa : number)
Reservation(sid : integer, mid : integer, when : date)
Movies(mid : integer, title : string, type : string)
```

The key of these relations are `sid`, $\{\texttt{sid}, \texttt{mid}, \texttt{date}\}$, and `mid` respectively. Attributes with the same name in different relations do have the same meaning and those in `Reservation` represent referential constraints (e.g., `sid` in `Reservation` refers to `Sid` in `Students`, etc.) Write a relational expression for the following queries:

(i) Find the name of students who have reserved an *action* movie and a *fiction* movie. (*action, fiction* are different movie types.)

List of sids who reserved action movie:

$$R1 = \pi_{sid}(S \bowtie (\sigma_{type='action'}(R \bowtie M)))$$

List of sids who reserved fiction movie:

$$R2 = \pi_{sid}(S \bowtie (\sigma_{type='fiction'}(R \bowtie M)))$$

Answer

$$\pi_{name}(S \bowtie (R1 \cap R2))$$

(ii) Find the name of students who have reserved at least two movies.

$$\pi_{name}(S \bowtie (\sigma_{mid \neq mid1}(R \bowtie R1[sid, mid1, when1])))$$

(iii) Find the name of students with GPA less than 2.0 who have not reserved a movie.

$$\pi_{name}(\sigma_{gpa<2.0}(S \bowtie (\pi_{sid}(S) - \pi_{sid}(R))))$$

(iv) **(CS 502 only)** Find the name of students who have reserved all movies.

$$\pi_{name}(S \bowtie (\pi_{sid,mid}(R) / \pi_{mid}(M)))$$

2. (25 points) Consider the sanme relation schema in question 1. Write SQL commands to

(a) answer the questions in **1**. (**CS482**: please do also (iv)).

(b) create a list of students with the number of movies they have reserved.

```
select s.sname from student s, reservation r, movie m
where s.sid = r.sid and m.mid = r.mid and m.type='action'
intersect
select s.sname from student s, reservation r, movie m
where s.sid = r.sid and m.mid = r.mid and m.type='fiction';

select s.sname from student s
where s.sid in
   (select r.sid from reservation r
    group by r.sid having count(*) >= 2);

select s.sname from student s
where s.gpa < 2 and
      s.sid not in (select r.sid from reservation r);

select s.sname from student s
where not exists
    (select m.mid from movie m
     except
     select r.mid from reservation r
     where s.sid = r.sid
    );

select s.sname, s.sid, count(*) as noofMR
from student s, reservation r
where s.sid = r.sid
group by s.sid, s.sname
union
select s.sname, s.sid, 0 as noofMR
from student s
where s.sid not in (select r.sid from  reservation r);
```

3. (15 points) Let us consider a relation schema $\mathcal{R} = (ABC, \{A \rightarrow C\})$ where A, B, and C are integers.

(a) Write a SQL command to create a table in Oracle which stores valid instances of $\mathcal{R}$.

(b) Assume that you have create the table named R in (a) and the table is empty. What will happen if the following SQL commands are issued, in the order listed below:

```
insert into R values(1,2,3);
insert into R values(1,4,4);
insert into R values(2,4,4);
```

You can answer the question by listing the content of R after each command.

```
create table R(a int, b int, c int, primary key (ab));
alter table R add constraint fd_constraint
   check (not exists
             (select * from R as R1, R as R2
              where R1.a = R2.a and R1.c != R2.c));
```

After first insertion: the table has (1,2,3)

After second insertion: the table has (1,2,3) since the second tulpe violates the functional dependency

After third insertion: the table has (1,2,3) and (2,4,4)

4. (20 points) Consider the schema

$$\mathcal{R} = (ABCDE, \mathcal{F})$$

where $\mathcal{F} = \{AB \to C, DE \to C, B \to D\}$.

(i) Find three FDs entailed by the set of FDs $\mathcal{F}$ which violate the BCNF condition. Provide the reasons for your answer.

(ii) Find a BCNF decomposition for $\mathcal{R}$. Show the steps in your decomposition.

(iv) Is your decomposition lossless? Why?

(v) **CS502 only:** Are all relations in your BCNF decomposition of $\mathcal{R}$ in 3NF? Justify your answer.

(i) Since $AB^+ = \{A, B, C, D\}$, AB is not a superkey. $AB \to C$ is not a trivial FD. So, $AB \to C$ violates the BCNF condition.

Since $DE^+ = \{E, C, D\}$, DE is not a superkey. $DE \to C$ is not a trivial FD. So, $DE \to C$ violates the BCNF condition.

Since $B^+ = \{B, D\}$, B is not a superkey. $B \to D$ is not a trivial FD. So, $B \to D$ violates the BCNF condition.

(ii) First, use $AB \to C$ to decompose, we get $\mathcal{R}_1 = (ABC, \{AB \to C\})$ and $\mathcal{R}_2 = (ABDE, \{B \to D\})$.

$\mathcal{R}_1$ is indeed in BCNF since AB is now the key of the schema $\mathcal{R}_1$.

$\mathcal{R}_2$ is not in BCNF since $B \to D$ violates the BCNF condition. Continue decomposing using this FD, we get $\mathcal{R}_3 = (BD, \{B \to D\})$ and $\mathcal{R}_4 = (ABE, \{\})$. Both are in BCNF.

So, a decomposition of the schema is (ABC, BD, ABE) (with the corresponding set of FDs!).

(iii) We follow the algorithm for BCNF decomposing, which guarantees the lossless.

(iv) Yes, since every relation in BCNF is in 3NF.

5. (20 points) Consider the schema

$$\mathcal{R} = (ABCDE, \mathcal{S})$$

where $\mathcal{S} = \{ACDE \bowtie AB, ABDE \bowtie ABC, A \to D, AB \to E\}$.

(a) Is $\mathcal{R}$ in 4NF? Why?

(b) Suppose that we decompose $\mathcal{R}$ into $\mathcal{R}_1 = (ACDE, \mathcal{S}_1)$ and $\mathcal{R}_2 = (AB, \mathcal{S}_2)$. Answer the following

(i) What will be $\mathcal{S}_1$? Is $\mathcal{R}_1$ in 4NF?

(ii) **CS502 only:** What will be $\mathcal{S}_2$? Is $\mathcal{R}_2$ in 4NF?

(a) Since $A^+ = \{A, D\}$ we know that A is not a supperkey.

Consider $ACDE \bowtie AB$, which is a non-trivial MVD, we have that $ACDE \cap AB = A$ which is not a supperkey. This means that the relation is not in 4NF.

(b.i) $\mathcal{S}_1$ should contain both MVDs and FDs. From $ACDE \bowtie AB$, we get $ACDE \bowtie A$ in $\mathcal{S}_1$.
From $ABDE \bowtie ABC$, we get $ADE \bowtie AB$ in $\mathcal{S}_1$.
From $A \to D$, we get $A \to D$ in $\mathcal{S}_1$.
So, $\mathcal{S}_1 = \{ACDE \bowtie A, ABDE \bowtie ABC, A \to D\}$.

Again, A is not a superkey of $\mathcal{R}_1$. We have that $ABDE \bowtie ABC$ is a non-trivial MVD with $ABDE \cap ABC = AB$ which is not a superkey. So, $\mathcal{R}_1$ is not in 4NF.

(b.ii) $\mathcal{S}_2$ should contain both MVDs and FDs. From $ACDE \bowtie AB$, we get $A \bowtie AB$ in $\mathcal{S}_2$.
From $ABDE \bowtie ABC$, we get $AB \bowtie AB$ (trivial) in $\mathcal{S}_2$.
From $A \to D$ and $AB \to E$, we get no FDs in $\mathcal{S}_2$.
So, $\mathcal{S}_2 = \{A \bowtie AB\}$. No nontrivial MVDs exist for $\mathcal{R}_2$, so it is in 4NF.