German University in Cairo
Media Engineering and Technology Faculty
Prof. Dr. Slim Abdennadher

January 10, 2017

CSEN501: Databases I
Winter Semester 2016-2017
Final Exam

Bar Code

Instructions: Read carefully before proceeding.

1) Please tick your major

	Major
	CSEN
	DMET
	BI

2) Duration of the exam: 3 hours (180 minutes).

3) No books or other aids are permitted for this test.

4) This exam booklet contains 16 pages, including this one. Three extra sheets of scratch paper are attached and have to be kept attached. **Note that if one or more pages are missing, you will lose their points. Thus, you must check that your exam booklet is complete**.

5) Write your solutions in the space provided. If you need more space, write on the back of the sheet containing the problem or on the four extra sheets and make an arrow indicating that. **Scratch sheets will not be graded unless an arrow on the problem page indicates that the solution extends to the scratch sheets**.

6) When you are told that time is up, stop working on the test.

Good Luck!

Don't write anything below ;-)

Exercise	1	2	3	4	5	6	7	8	$\sum$
Possible Marks	6	10	10	16	14	9	8	10	83
Final Marks									

Exercise 1 (6 Marks)

a) Suppose we have a table called Student with the following content:

name	age	grade
Sarah	null	8
Ali	null	8
Mina	null	1
Menna	null	null
Nada	null	1
Mohamed	null	2

1. What is the output of the query below?

```
select min(grade), max(grade), sum(grade),
       avg(grade), count(grade), count(*)
from Student;
```

Solution:

```
min | max | sum | avg | count | count
-----+-----+-----+--------------------+-------+-------
1 | 8 | 20 | 4.0000000000000000 | 5 | 6
```

b) We have tables R and T. The content of R is

a	b
1	2
8	7
5	null
null	6

The content of T is

b	c
2	5
2	9
1	4
null	18
6	88

What is the result of this query:

```
select * from R natural join T;
```

Solution:

```
b | a | c
---+---+----
2 | 1 | 5
2 | 1 | 9
6 | null| 88
```

Why are tuples (5, null) and (null, 18) excluded from the result, yet tuples (null, 6) and (6, 88) are included? (5, null) and (null, 18) are excluded because when the natural join compares the 2 null values, they are not considered equal. It makes sense if you think of natural join as involving a Cartesian product and a WHERE, and if you remember that the truth value of a comparison involving nulls is always unknown and that WHERE is picky. In contrast, (null, 6) and (6, 88) are included because we are comparing on attribute b, which is 6 for both of these tuples, not null.

c) Suppose we have this table:

```
create table names(first text, last text, unique (first, last));
```

What will the result of the following be?
Doing this twice: `insert into names values ('Nada', 'Sharaf');`

Exercise 2 (3+3+2+2=10 Marks)

Consider the relational schema:

```
Student(sID, surName, firstName, campus, email, cgpa)
Course(dept, cNum, name, breadth)
Offering(oID, dept, cNum, term, instructor)
Offering.(dept,cNum) references Course
Took(sID, oID, grade)
Took.sID references Student
Took.oID references Took
```

Note that `breadth` is a boolean indicating whether or not a course satisfies the breadth requirement for degrees in the Faculty of Media Engineering and Technology

a) Write a relational algebra query for each of the following:

1. The names of courses taught by two different instructors in two different terms.
2. The sID of the students who took all the breadth courses.

b) Given the following tables (where columns with the same names represent foreign keys):

R:

A	C	B
2	3	5
7	5	2
2	8	2
7	5	5

S:

B	D
5	3
2	3

T:

B	C	D
2	5	3
7	2	3
2	8	3
5	5	3

1. What is the output of:

$$\pi_{A,B,D}(R \bowtie S) \cap T$$

2. Which of the following queries has the following output:
Choose all the correct answers. | 7 | 5 |

   i. $\frac{R \bowtie T}{S}$

   ii. $\pi_{A=7}(R \bowtie S)$

   iii. $(R \times S) - (T \bowtie S)$

   iv. None of the above

Exercise 3 (10 Marks)

Consider the relational schema:

```
Student(sID, surName, firstName, campus, email, cgpa)
Course(dept, cNum, name, breadth)
Offering(oID, dept, cNum, term, instructor)
Offering.(dept,cNum) references Course
Took(sID, oID, grade)
Took.sID references Student
Took.oID references Took
```

Note that `breadth` is a boolean indicating whether or not a course satisfies the breadth requirement for degrees in the Faculty of Media Engineering and Technology

Write a relational calculus query for each of the following:

a) The names of all students who have earned a grade over 80 in a breadth course with Professor Slim.

b) sID of students who have a grade of 100 exactly twice.

c) sID of the students who took all the breadth courses.

Exercise 4 (16 Marks)

Consider the relational schema:

```
Student(sID, surName, firstName, campus, email, cgpa, major)
Course(dept, cNum, name, breadth)
Offering(oID, dept, cNum, term, instructor)
Offering.(dept,cNum) references Course
Took(sID, oID, grade)
Took.sID references Student
Took.oID references Took
```

Note that `breadth` is a boolean indicating whether or not a course satisfies the breadth requirement for degrees in the Faculty of Media Engineering and Technology

a) Does any of these queries give an error? Justify your answers by stating what the query does in English if it is legal. Otherwise, you have to state what the problem is.

1. ```
   SELECT dept
   FROM Took, Offering
   WHERE Took.oID = Offering.oID
   GROUP BY dept
   HAVING avg(grade) > 75;
   ```
2. ```
   SELECT Took.oID, count(grade)
   FROM Took, Offering
   WHERE Took.oID = Offering.oID
   GROUP BY Took.oID
   HAVING avg(grade) > 75;
   ```

b) State in English what do the following queries compute.

1.
```
SELECT s1.sID, t1.oID
FROM Student s1 LEFT OUTER JOIN Took t1 ON s1.sID = t1.sID
WHERE EXISTS
        (SELECT o.oID
         FROM Offering o INNER JOIN Took t2 ON o.oID = t2.oID
         WHERE t1.oID = t2.oID
         GROUP BY t2.oID
         Having COUNT(t2.sID) =
            (SELECT MAX(students)
             FROM (SELECT count(t3.sID) AS students
                   FROM Took t3
                   GROUP BY t3.oID)
            )
        )
```

2.
```
SELECT s1.sID, s2.sId
FROM Student s1, Student s2
WHERE s1.sID <> s2.sID AND
NOT EXISTS
        (SELECT t1.oID
        FROM Took t1
        WHERE t1.sID = s2.sID

        EXCEPT

        SELECT t2.oID
        FROM Took t2
        WHERE t2.sID = s1.sID
        )
```

c) Write a SQL query to create a trigger when a record is added in the table Took, the gpa of the corresponding student is updated. Assume the gpa is composed of the sum of all the grades.

d) Write a SQL procedure that returns the names of students that have taken, at some point, every course Carmen has taught (but not necessarily taken them from Carmen i.e. even if they didn't attend the course with Carmen).

Exercise 5 (14 Marks)

Answer the following questions.

a) Suppose the functional dependency $BC \rightarrow D$ holds in R. Create an instance of R that violates this FD.

Solution:

In order to violate this FD, we need two tuples with the same value for B and the same value for C (both!), yet different values for D.

```
A B C D
1 3 6 4
2 3 6 5
```

b) Are the sets $\{PQ \rightarrow R\}$ and $\{P \rightarrow Q, P \rightarrow R\}$ equivalent? If yes, explain why. If no, construct an instance of the relation that satisfies one set of FDs but not the other.

Solution:

These are not equivalent, as demonstrated by this instance that satisfies PQ ! R but not P ! Q; P ! R: P Q R 1 2 4 1 8 9

c) Suppose we have a relation on attributes $ABCDEF$ with these FDs:

$$\{AC \rightarrow F, CEF \rightarrow B, C \rightarrow D, DC \rightarrow A\}$$

Use the below Armstrong Axioms to answer the following:

$$Reflexivity \quad if \quad Y \subseteq X \; then \; X \rightarrow Y.$$

$$Augmentation \quad if \quad X \rightarrow Y \; then \; XZ \rightarrow YZ.$$

$$Transitivity \quad if \quad X \rightarrow Y \; Y \rightarrow Zthen \; X \rightarrow Z.$$

1. Does it follow that $C \rightarrow F$? Prove with steps or disprove with a counter example.
2. Does it follow that $ACD \rightarrow B$? Prove with steps or disprove with a counter example.

Solution:

We use the closure test to check whether an FD follow from a set of FDs. C+ = CDAF. Therefore, C ! F does follow. ACD+ = ACDF. Therefore, ACD ! B does not follow.

Exercise 6 (9 Marks)

a) Suppose we have a relation with attributes `cdf, name, grade`. Here is an instance of that relation:

`cdf`	`name`	`grade`
g3tout	Amy	91
g4foobar	David	78
c0zhang	David	85

1. Suppose we were to decompose this into two new relations: `R1(cdf, name)` and `R2(name, grade)`. Project the data onto those two new relations.
2. Now compute `R1` ⋈`R2` to rebuild the original table.
3. What was lost?

Exercise 7 (8 Marks)

a) Suppose we have a relation with attributes $movie, theatre, city$ and FDs $\{theatre \rightarrow city, movie, city \rightarrow theatre\}$.

1. What is the candidate key of the relation? Justify your answer.
2. Is the relation in 3NF? Justify your answer.
3. Is the relation in BCNF? if not, normalize it to BCNF.
4. Does the normalization preserve the functional dependencies? Justify your answer.

Exercise 8 (10 Marks)

a) Consider a relation schema R with attributes ABCDEFGH with functional dependencies FD:

$$FD = \{B \rightarrow CD, BF \rightarrow H, C \rightarrow AG, B \rightarrow F, FC \rightarrow H, CH \rightarrow B\}$$

1. Find a minimal cover for FD describing all the steps.
2. Employ the 3NF decomposition algorithm to obtain a lossless decomposition of R into a collection of relations that are in 3NF. Make sure it is clear which relations are in the final decomposition.
3. Is your decomposition dependency-preserving?

Scratch paper

Scratch paper

Scratch paper