CS 338 Computer Applications in Business: Databases (Fall 2013)

Midterm Exam

Exam date: 8 October 2013
Appeal deadline: 22 October 2013

Student Id: _____

Student Name: _____

Instructions:

1. This is a closed book examination.
2. Exam duration is 75 minutes (11:30AM-12:45PM).
3. Answer all the questions.
4. Answer each question in the space provided.
5. The exam is 7 (seven) pages; make sure you have all of the pages.

Question	Points	Score
1	5	
2	10	
3	4	
4	2	
5	2	
6	2	
Total:	25	

Q1. (5 points)

Choose the best answer for each of the following questions.

(a) If nothing is known about a particular value in a relation, we assign the value:

A. 0
B. 999999
C. '' (i.e., the empty string)
D. NULL

(a) _____

(b) Declaring that attribute A is a foreign key is an example of a/an _____ constraint.

A. inherent model-based
B. schema-based
C. application-based
D. key

(b) _____

(c) A view defined over a relational database corresponds to a/an _____ schema

A. external
B. conceptual
C. internal
D. classical

(c) _____

(d) The business role responsible for ensuring that a database system runs efficiently is held by one or more persons acting as a/an.

A. database administrator
B. database designer
C. application programmer
D. sophisticated end-user

(d) _____

(e) Which of these four relational expressions is **not** like the others (gives a different result for some input R)?

A. $\sigma_{<condition_1>}(\sigma_{<condition_2>}(\sigma_{<condition_3>}(R)))$
B. $\sigma_{<condition_2>}(\sigma_{<condition_1}$ AND $<condition_3>}(R))$
C. $\sigma_{<condition_2>}$ AND $<condition_3$ AND $<condition_1>}(R)$
D. $\sigma_{<condition_2>}$ **OR** $<condition_3$ **OR** $<condition_1>}(R)$

(e) _____

Solution: Correct answers are in bold.

Marking scheme: Straightforward.

Q2. (10 points)
Consider the following relations:

Emp(eno, ename, title, city)
Proj(pno, pname, budget, city)
Works(eno, pno, resp, dur)
Pay(title, salary)

where the primary keys are underlined, and Emp.title is a foreign key to Pay.title, Works.eno is a foreign key to Emp.eno, and Works.pno is a foreign key to Proj.pno.

For each part of this question (considered independently of the other parts), write a single SQL statement that accomplishes the given requirements.

(a) (2 points) Retrieve the names of employees whose salaries are greater than $100,000.

Solution:

```
SELECT ename
FROM Emp, Pay
WHERE salary > 100000
AND Emp.title = Pay.Title
```

Marking scheme:

- selection condition: 0.5
- join condition: 0.5
- select attributes: 0.5
- overall correctness: 0.5

(b) (2 points) For each city, how many projects are located in that city and what is the total budget over all projects in the city?

Solution:

```
SELECT city, count(pno) AS totproj,
    sum(budget) AS totbudget
FROM Proj
GROUP BY city
```

Marking scheme:

- use “count”: 0.5
- use “sum”: 0.5
- group by: 0.5
- select attributes: 0.5

(c) (2 points) Retrieve the name and title of employees who work for more than 12 months on a project whose budget is greater than $525,000.

Solution:

```
SELECT name, title
FROM Emp, Proj, Works
WHERE budget > 525000
AND dur > 12
AND Emp.eno = Works.eno
AND Proj.pno = Works.pno
```

Marking scheme:

- two joins: 1.0
- two selection constraints: 0.5
- correct select clause: 0.5

(d) (2 points) For each project that employs more than 10 employees, retrieve the project number and number of employees that it employs. Sort your answers in ascending project number.

Solution:

```
SELECT pno, COUNT(eno)
FROM Works
GROUP BY pno
HAVING COUNT(eno) > 10
ORDER BY pno
```

Marking scheme:

- count: 0.5
- group by: 0.5
- order by: 0.5
- having clause: 0.5

(e) (2 points) Retrieve the employee number and responsibility of *all* persons assigned to any project for which more than any 2 persons share the same responsibility for that project.

Solution:

```
SELECT W1.eno, W1.resp
FROM Works W1
WHERE pno in
        (SELECT W2.pno
         FROM Works W2
         GROUP BY W2.pno, W2.resp
         HAVING COUNT(*)>2)
```

Marking scheme:

- has the where condition “in” or “= any” etc.: 0.5
- use subquery: 0.5

- group by: 0.5
- having: 0.5

Q3. (4 points)

Formulate the following SQL queries in relational algebra. These are expressed over the same schema as given in the previous question

(a) (2 points)

```
SELECT pname, budget
FROM Proj, Works, Emp
WHERE title='Programmer'
AND Works.eno = Emp.eno
AND Works.pno = Proj.pno
```

Solution:

$$\Pi_{\text{pname,budget}}(\text{Proj} \bowtie \text{Works} \bowtie (\sigma_{\text{title='Programmer'}}, \text{Emp}))$$

You can put the selection outside all of the joins if you wish; that would work too:

$$\Pi_{\text{pname,budget}}(\sigma_{\text{title='Programmer'}}, (\text{Proj} \bowtie \text{Works} \bowtie \text{Emp}))$$

Marking scheme: This is straightforward to mark so I indicate where you lose points:

- If you used cross product rather than join, you lose 0.5 marks.
- If you insert a projection that eliminates columns that you need in the outer projection, you lose 0.5 mark.
- If you missed one of the joins, you lose 0.5 mark
- If you got the join generally, but then made other errors (such as doing two joins and then trying to union them) you lose 0.5 marks.

(b) (2 points)

```
SELECT ename
FROM Emp
WHERE City = 'Waterloo'
```

Solution:

$$\Pi_{\text{ename}}(\sigma_{\text{City='Waterloo'}}, (\text{Emp}))$$

Marking scheme:

- selection operation: 1 mark
- projection operation: 1 mark

Q4. (2 points)
Briefly describe the difference between a *relation schema* and a *relation instance* (or *relation state*).

Solution: A relation schema is a description of of a relation. It includes the relation name, attribute names, domains for the attributes, and possibly additional constraints on valid relations (e.g., keys). A relation instance (or state) is a set of tuples that conform to a relation scheme at some moment in time. It is a finite subset of the Cartesian product of the domains corresponding to the attributes in the schema.

Marking scheme: Describing each of the two is worth 1 mark.

Q5. (2 points)
Write an appropriate DDL statement to declare the following relation in SQL:

`Car (LicenceNum: string, Province: string, Make: string, Model: string, Year: integer, VIN: string)`

where values of `LicenceNum` has no more than 8 characters; `Province` has exactly 2 characters; `Make` and `Model` both have up to 20 characters; `Make` is a foreign key into a relation called Manufacturer (having an attribute also named `Make` as its primary key); `Year` must be between 1900 and 2013; and `VIN` has exactly 15 characters and cannot be identical for two or more `Car` tuples.

Solution:

```
CREATE TABLE Car (
    LicenceNum VARCHAR (8) NOT NULL,
    Province CHAR(2) NOT NULL,
    Make VARCHAR(20),
    Model VARCHAR(20),
    Year INTEGER,
    VIN CHAR(15) UNIQUE,
    CONSTRAINT CarPK PRIMARY KEY (LicenceNum, Province),
    CONSTRAINT CarFK FOREIGN KEY (Make) REFERENCES Manufacturer(Make),
    CONSTRAINT CarYear CHECK (Year BETWEEN 1900 AND 2013)
);
```

Notes:

- Do not pay attention to the bold typefaces – the typesetting system is emphasizing SQL keywords automatically and sometimes gets confused.
- I don't care if you declare the PRIMARY KEY, FOREIGN KEY, and CHECK constraints as named constraints. They can simply be written as, e.g., PRIMARY KEY LicenceNum,Province).

Marking scheme:

- Attribute definitions: 0.5 marks – partial marks if you get some right and some wrong.
- Primary key definition: 0.5 marks.
- Foreign key definition: 0.5 marks.
- CHECK constraint: 0.5 marks.

Q6. (2 points)

Consider the instance of the Sailors relation given in Figure 1 where the key is underlined. Show the result of the following query.

```
SELECT S.sid
FROM Sailors S
WHERE S.rating >= ALL
        (SELECT S2.rating
         FROM Sailors S2)
```

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

Figure 1: Database instance for Question 6

Solution:

sid
58
71

Marking scheme: If you compute the right answer, you get the marks. However:

- If you did not specify column headings, you lost 0.5 marks. The result that is returned is a table, so you need to specify the column headings.
- If you made other minor errors, you lost 0.5 marks.