COP 4710 Database Systems

Spring 2013

Midterm Exam

March 5, 2013

Time: 75 minutes

Your Name: _____

USF ID: U_____

Problem I. (20pts) True or false, mark your choice clearly. Ambiguous answers will get zero points. You do not have to explain how you reached your conclusion.

1. SQL is an implementation of tuple relational calculus. T(✓) F( )
2. In relational algebra, if we have r x s = T, the **degree** of table T is the **product** of the degrees of tables r and s. T( ) F(✓)
3. A candidate key must also be a superkey. T(✓) F( )
4. In relational algebra, the expression $r \cup s$ is meaningless unless the two operands r and s have exactly the same schema. T( ) F(✓)
5. Relational algebra is a declarative language. T( ) F(✓)
6. A foreign key is not necessarily a key in the **referencing** table. T(✓) F( )
7. In a relation, a candidate key is not necessarily a primary key. T(✓) F( )
8. **Cardinality** of a table is defined as the number of tuples in the table. T(✓) F( )
9. By default, the resulting table of an SQL query is duplicate free. T( ) F(✓)
10. The EXISTS keyword in SQL tests if a following SQL statement returns an empty set. T(✓) F( )

Problem I. Query languages (50pts, 5 pts each)

Given the following relational database that stores Hollywood movie information

M	Movies (title, year, length, genre, studioName, ProducerName)
S	StarsIn (movieTitle, starName, pay)
MS	MovieStar (name, address, gender, birthdate)
ME	MovieExecutive(name, address, certificate_num, networth)

where the underlined attributes are the primary keys. The following foreign keys also exist:

StarsIn.starName --> MovieStar.name
StarsIn.movieTitle --> Movies.title
Movies.ProducerName --> MovieExecutive.name

I-A. Write the following queries in SQL (Queries written in the wrong language will get zero points)

1. Print the names and addresses of movie stars who were born before 08/15/1945;

```
SELECT MS.name, MS.address
FROM MovieStar MS
WHERE birthdate < '08/15/1945';
```

2. Print the title and year of movie(s) in which there is a movie star with a pay over 40 million dollars, and also print the movie star's name and birthdate.

```
SELECT MS.name, MS.birthdate, M.title, M.year
FROM MS, S, M
WHERE M.title = S.movieTitle AND
      MS.name = S.starName AND
      S.pay > 40M
```

3. Print the names of actors who starred in both an MGM (studio name) movie and a Universal (studio name) movie;

```
SELECT S1.starName
FROM S1, M1
WHERE $M1.studioName = 'MGM'
      AND M1.title = S.movieTitle

INTERSECT

                                  'Universal'
```

M	Movies (title, year, length, genre, studioName, ProducerName)
S	StarsIn (movieTitle, starName, pay)
MS	MovieStar (name, address, gender, birthdate)
ME	MovieExecutive(name, address, certificate_num, networth)

4. What is the average length of movies produced by Kevin Costner?

```
SELECT  AVG(length)
FROM  M
WHERE  ProducerName = 'Kevin Costner'
```

5. For each movie genre, find the total number of movies, but only return this number for those genres with at least 5 movies in it.

```
SELECT  COUNT(title)
FROM  M
GROUP BY Mgenre
HAVING  COUNT(title) >= 5
```

6. Print the names of movie stars who have acted in EVERY movie produced by Spielberg from 1990 to 1995.

Solution #1

```
SELECT MS.~~starName~~ name
FROM MS
WHERE  NOT EXISTS  ( SELECT  title
                     FROM  M1
                     WHERE M.ProducerName = 'Spielberg'
                        AND M.year ≥ 1990 AND
                        M.year ≤ 1995 )
MINUS
( SELECT  M2.title
  FROM  M2, S
  WHERE  ~~M2.name = S.starName~~
         M2.title = S.movieTitle
  ∧  MS.name = S.starName
```

Solution #2

```
SELECT  S.starName
FROM  S, M
WHERE  S.movieTitle = M.title
~~WHERE~~ AND  ~ 1990 ~ 1995
GROUP BY  S.starName
HAVING  COUNT(M.title)
   IN  ( SELECT  COUNT(title)
         FROM  M1
         WHERE  Spielberg
                1990
                1995 )
```

M	Movies (title, year, length, genre, studioName, ProducerName)
S	StarsIn (movieTitle, starName, pay)
MS	MovieStar (name, address, gender, birthdate)
ME	MovieExecutive(name, address, certificate_num, networth)

II-B. Write the following queries in relational algebra:

7. Print the titles and years of movies in which Tom Hanks starred.

$$\Pi_{title, year}\left(\sigma_{M.title = S.movieTitle \wedge S.starName = 'Tom\ Hanks'}(M \times S)\right)$$

8. If, in any movie, a movie star's pay is more than 50% of the networth of the movie's producer, print the movie star's name and the producer's name.

$$\Pi_{S.starName,\ ME.name}\left(\sigma_{S.movieTitle = M.title \wedge M.ProducerName = ME.name \wedge S.pay > ME.networth \times 0.5}(S \times M \times ME)\right)$$

9. Print the names of movie stars who have acted in EVERY movie produced by Spielberg from 1990 to 1995.

$$\Pi_{movieTitle,\ starName}(S) \div \Pi_{title}\left(\sigma_{M.ProducerName = 'Spielberg' \wedge M.year >= 1990 \wedge M.year <= 1995}(M)\right)$$

10. Print the names of producers who produced exactly ONE movie in 2009;

Solution #1

$$\Pi_{ProducerName}\left(\sigma_{M.year = 2009}(M)\right) \rightarrow T_2$$

$$\Pi_{producerName}\left(\sigma_{M_1.ProducerName = M_2.ProducerName \wedge M_1.title <> M_2.title \wedge M_1.year = 2009 \wedge M_2.year = 2009}(M_1 \times M_2)\right) \rightarrow T_1$$

Answer: $T_2 - T_1$

Solution #2:

$$M.ProducerName\ \mathcal{G}_{CNT(title) \rightarrow CNT\_title}\left(\sigma_{M.year = 2009}(M)\right) \rightarrow T_1$$

Answer $\longrightarrow$

$$\Pi_{producerName}\left(\sigma_{CNT\_title = 1}(T_1)\right)$$

II. (30 pts, 5 pts each) Consider the following two relations:

Relation *r*

B	C	D
a	5	*b*
b	6	*a*
c	25	*b*
a	5	*c*

Relation *s*

D	E	F
b	10	6
c	25	3
b	10	5

give the resulting table of the following relational algebraic expressions. You should specify the schema as well as the values of all attributes for all tuples in your solutions. If you believe the expression does not make sense, put ϕ as your answer. Write down intermediate results for partial credits.

1. $\Pi_C(r)$

C
5
6
25

2. $\Pi_B \sigma_{r.C<6}(r)$

B	C	D
a	5	b
a	5	c

B
a

← Answer

3. $r \cap s$ = ϕ Not union compatible !!

4. r x s (hint: cross product)

	B	C	r.D	s.D	E	F
✓	a	5	b	b	10	6
	a	5	b	c	25	3
✓	a	5	b	b	10	5
	b	6	a	b	10	6
	b	6	a	c	25	3
	b	6	a	b	10	5
✓	c	25	b	b	10	6
	c	25	b	c	25	3
✓	c	25	b	b	10	5

5. $\Pi_{B,F}(r \bowtie s)$ (hint: this is natural join)

r ⋈ s will return rows with a checkmark shown above.

Answer:

B	F
a	6
a	5
c	6
c	5
a	3

↓ cont.

	a	5	c	b	10	6
✓	a	5	c	c	25	3
	a	5	c	b	10	5

6. $\Pi_{B,D}(r) \div [\Pi_{D}\sigma_{r.D != a}(r)]$

B	D
a	b
b	a
c	b
a	c

÷

D
b
c

=

B
a