Comp 521 Midterm 1

9 October 2018

Instructions

Don't PANIC!. Focus on getting your answer as close to correct as you can. There will be an opportunity for partial credit on the SQL questions that will depend on how much you have to change to make your answer correct.

We will **not** answer questions about course content, SQL syntax, etc. We will only deal with issues related to the exam implementation.

You should use the latest **Google Chrome** browser to take this exam; it may work in other browsers but I can't help you if it doesn't.

If your browser hangs, for example because of a bad SQL query, simply kill the page and refresh. It should restore all of your work.

You may use **only** a single 8.5 by 11 sheet of paper, possibly double sided, as a cheatsheet.

You **must expand your web browser window to cover the full screen** and keep it that way for the duration of the exam. You must **not** access anything besides this page before you submit the exam. I will know if you use other browser pages or programs.

You may **NOT** leave before you submit your exam. When you submit you must enter the code displayed on the screen at the front of the class or given to you by ARS. Only your **first** submission with a correct submit code will be graded.

After you submit you must leave the room. Do not use your phone or computer until after you leave the room.

No submissions will be accepted after the allotted time.

Questions

1. (6 points) Which of the following SQL operations, used in the normal ways we discussed, will change the schema?

none of these 6 points

2. (6 points) Which of the following is characteristic of a database schema?

- [ ] It determines the degree
- [ ] It determines cardinality
- [ ] It must specify a primary key

3. (6 points) What is the value of this SQL expression: (true or true and null)?

Choose an answer 6 points

4. What is the size of the relational algebra result? **(2 points each, 16 total)**

Consider the relation A with degree $A_d > 0$ and cardinality $A_c > 0$. A has a column named x of type integer and possibly many other columns.

For each of the expressions in parts 4.1 through 4.4 below you are to choose the best description of the size of result, C

Which expression best describes the cardinality of C?

a. $C_c = 0$
b. $C_c = 1$
c. $C_c = A_c$
d. $1 <= C_c <= A_c$
e. $0 <= C_c <= A_c$
f. $0 <= C_c < 5$

Which expression best describes the degree of C?

A. $C_d = 0$
B. $C_d = 1$
C. $C_d = A_d$
D. $1 <= C_d <= A_d$
E. $0 <= C_d <= A_d$
F. $0 <= C_d < 5$

4.1 $C = \pi_x A$

Cardinality: Choose an answer 2 points Degree: Choose an answer 2 points

4.2 $C = \sigma_{x<5} A$

Cardinality: Choose an answer 2 points Degree: Choose an answer 2 points

4.3 $C = \pi_x \sigma_{x<5} A$

Cardinality: Choose an answer 2 points Degree: Choose an answer 2 points

4.4 $C = \sigma_{x<5} \pi_x A$

Cardinality: Choose an answer 2 points Degree: Choose an answer 2 points

Database info

You are to write a single SQL query to answer each question. You will not write any Python code below.

The schema of the database is below. It represents book reading activity on Tar Heel Reader over a short period.

The **ip** address is assumed to correspond to a single location. For reasons you'll learn in Comp 431 they sometimes don't but for the purposes of the exam assume they do.

```
create table Reads ( -- a record for each book that was read
    time date,       -- time the book was read
    ip integer,      -- ip of the computer
    bid integer,     -- book id
    foreign key (bid) references Books,
    foreign key (ip) references Locations)

create table Books (
    bid integer primary key,
    title text,
    aid integer,           -- author id
    reviewed integer,      -- 1 if the book is reviewed, 0 otherwise
    foreign key (aid) references Authors)

create table Locations (
    ip integer primary key, -- ip address of the computer
    country text)           -- country determined from the ip address

create table Authors (
    aid integer primary key,
    login text,    -- user id like Gary
    birthday text) -- month and day of the author's birthday like May 29
```

I'll dump the first few rows of each table so you can get a better idea of the contents.

Reads

```
(time, ip, bid)

('2017-10-01 00:06:48', 3023614022, 124363)
('2017-10-01 00:07:35', 3023614022, 124884)
('2017-10-01 00:07:53', 3023614022, 124366)
```

Books

```
(bid, title, aid, reviewed)

(148, 'Lullaby', 9, 1)
(171, 'Balloons Everywhere!', 14, 1)
(202, 'Baa Baa Black Sheep', 61, 1)
```

Authors

```
(aid, login, birthday)

(1, 'Gary', 'May 29')
(2, 'DLM', 'February 24')
(5, 'Jenny', 'January 18')
```

Locations

```
(ip, country)

(32591211, 'United States')
```

```
(40693121, 'United States')
(60836422, 'Canada')
```

More questions

5. (8 points) How many different ip addresses read books?

```
SELECT COUNT(DISTINCT L.ip)
FROM Reads R, Locations L
WHERE R.ip = L.ip
```

Execute 8 points

6. (10 points) Which countries read books?

List the country names in alphabetical order. Note: ip addresses may be included in the Locations table that never actually read a book.

```
SELECT L.country
FROM Locations L, Reads R
WHERE L.ip = R.ip
GROUP BY L.Country
```

Execute 10 points

7. (12 points) Which authors have written more than 10 reviewed books?

List the author login and number of books. List them with the largest count first; in case of ties on count put them in alphabetical order by login. Only include books with the reviewed flag equal 1.

```
SELECT A.login, COUNT(*)
FROM Authors A, Books B
WHERE A.aid = B.aid AND B.reviewed = 1
GROUP BY A.login
HAVING COUNT(*) > 10
ORDER BY COUNT(*) DESC, A.login
```

Execute 12 points

8. (12 points) What was the maximum number of times any book was read?

```
WITH Temp AS
(SELECT COUNT(*) as count
FROM Reads R
GROUP BY R.bid
ORDER BY COUNT(*) DESC
)
```

Execute 12 points

9. (12 points) Which book(s) were read in the most different countries?

List the login of the author, title of the book, and number of countries. In case of a tie for max list them all.

```
WITH Temp AS
(SELECT L.country as country, B.title as title, A.login as login, COUNT(*) as count
```

```
FROM Locations L, Books B, Reads R, Authors A
WHERE L.ip = R.ip AND R.bid = B.bid AND A.aid = B.aid
GROUP BY B.title
HAVING count > 5
```

Execute 12 points

10. (12 points) Which books were read at least 5 times by the same ip address?

For each book, list the title, the number of times it was read by the same ip address and the country corresponding to that ip address. Order your result by title in alphabetical order, then by count in decreasing order and then by country in alphabetical order. Hint: you can group by more than one value.

```
SELECT B.title, COUNT(*) as count, L.country
FROM Locations L, Reads R, Books B
WHERE L.ip = R.ip AND B.bid = R.bid
GROUP BY (L.ip)
HAVING count > 5
ORDER BY B.title, count DESC, L.country
```

Execute 12 points

Honor Pledge

I certify that no unauthorized assistance has been received or given in the completion of this work. Fill in your full name here: Enter your full name

Submit your exam

You must enter the exam submit code displayed on the screen (or given to you by ARS) immediately before submitting. Your submission **will not be graded** if you use an incorrect or old code. The system will warn you if your code is invalid. If that happens simply use the back button on your browser to go back and enter the correct code.

Only your **first submission** with the correct code will be counted. Do not enter the code before you are ready to make your final submission.

Enter submit code:

Submit