PLEASE HAND IN

PLEASE HAND IN

UNIVERSITY OF TORONTO
Faculty of Arts and Science

Midterm Test

CSC 343H1F, 2016
Section L5101/L2501 (Miller)

Duration — 50 minutes
No aids allowed

Student Number: |_|_|_|_|_|_|_|_|_|_|

Last Name: _____

First Name: _____

Section normally attended (Circle one): L0101/L2001(1pm) L0201/L2003 (3pm) L5101/L2501(6pm)

Do **not** turn this page until you have received the signal to start.
In the meantime, please fill out the identification section above,
and read the instructions below.

This test consists of 4 questions on 6 pages (including this one). When you receive the signal to start, please make sure that your copy of the test is complete.

If you use any space for rough work or need to scratch out an answer, circle the part that you want us to mark.

You may write in pencil, however, work written in pencil will not be considered for remarking.

Good Luck!

Question 1. [12 MARKS]

Twitter is a social media platform where users post messages called "tweets". Users can "follow" other users, and can "like" a tweet. Consider this relational schema for Twitter data. Keys are underlined.

As this question considers relational algebra, assume all relations are sets containing no nulls.

User(userID, name, email)

A Twitter user.

Tweet(tweetID, userID, content, date)

The user with *userID* made a tweet containing *content* on *date*.

Follows(a, b)

User a follows user b on Twitter, which means that
a has subscribed to b's tweets.

Likes(who, tweetID, userID)
User *who* liked tweet *tweetID, userID*.

Tweet[userID] $\subseteq$ User[userID]

Follows[a] $\subseteq$ User[userID]

Follows[b] $\subseteq$ User[userID]

Likes[who] $\subseteq$ User[userID]

Likes[tweetID, userID] $\subseteq$ Tweet[tweetID, userID]

Part (a) [2 MARKS]

Does the scheme enforce this constraint: a user cannot like his or her own tweet? Circle one:

Yes No.

If yes, explain; if not write a new constraint to enforce it. You cannot change any of the existing constraints, rather write a new constraint to enforce it:

Part (b) [2 MARKS]

Does the schema enforce this constraint: every userID in *Tweet* must have a name and email (recorded in *User*)? Circle one:

Yes No.

If yes, explain; if not write a new constraint to enforce it:

Part (c) [2 MARKS]

Suppose relation *Tweet* has 100 tuples and $\Pi_{TweetID}(Tweet)$ has 10 tuples. How many tuples could *Users* have? Circle all that apply:

0 1 100 10,000 100,000

Part (d) [4 MARKS]

Which of the following pairs of queries are equivalent? Circle each pair that returns the same results on all database instances.

1. $\Pi_{tweetID}(Likes \bowtie Tweet) = \Pi_{tweetID}(Tweet)$
2. $User = User \times User$
3. $\Pi_{userID}(\sigma_{tweetID=1}(Tweet) = \Pi_{userID}(\sigma_{tweetID=1}(Likes))$
4. $\Pi_{name}(User \bowtie Tweet \bowtie Likes) = \Pi_{name}(User \bowtie Tweet \bowtie Likes \bowtie Follows)$

Part (e) [2 MARKS]

Which of the following queries can be expressed using the same form of relational algebra that we used in class and on Assignment 1, that is assignment, and the operators $\Pi, \sigma, \bowtie, \bowtie_{condition}, \times, \cap, \cup, -, \rho$? Circle all that apply.

1. Find all users who have never tweeted.
2. Find dates on which every tweet that was posted was liked by at least two users.
3. Find the email of the most popular user (the user whose tweets are liked by the most people).
4. Find the first date on which a tweet was made.
5. Find the users with the fewest followers.

Question 2. [8 MARKS]

Here is the schema from Assignment 1. A few attributes and relations have been omitted for simplicity.

Relations

Product(DIN, manufacturer, name, form, schedule)
A tuple in this relation represents a drug product.

Price(DIN, price)
The price of a drug product.

Prescription(RxID, date, patient, drug, doctor)
A prescription for *drug* was written on *date* for *patient* by *doctor*. Attribute *patient* is the patient's OHIP number.

Filled(RxID, date, pharmacist)
Prescription *RxID* was filled by *pharmacist* on *date*. Attribute *pharmacist* is the pharmacist's OCP number.

Integrity constraints

$\text{Price}[\text{DIN}] \subseteq \text{Product}[\text{DIN}]$

$\text{Prescription}[\text{drug}] \subseteq \text{Product}[\text{DIN}]$

$\text{Filled}[\text{RxID}] \subseteq \text{Prescription}[\text{RxID}]$

$\Pi_{\text{schedule}}\text{Product} \subseteq$ {"prescription", "narcotic", "OTC"}

For every drug product that has been prescribed, report its DIN, the date on which the first prescription for it was written, and the date on which the last prescription for it was written.

Use only the basic operators $\Pi, \sigma, \bowtie, \times, \cap, \cup, -, \rho$, and assignment.

Continue your answer here if more space is needed.

Question 3. [6 MARKS]

Suppose we have implemented the Twitter schema from Question 1 in SQL, and the tables currently contain the following:

User:

userid	name	email
adele	Adele Laurie Blue Adkins	
drizzy	Drake	
potus	Barack Obama	potus@gov.us
rjm	Renee Miller	rjm@cs

Follows:

a	b
potus	drizzy
drizzy	rjm
drizzy	adele
adele	drizzy

Tweet:

tweetid	userid	content	date
15	adele	Hello	2016-10-16
61	adele	It's me	2016-10-16
33	potus	6 weeks	2016-10-11
28	rjm	in the 6	2016-10-10

Likes:

who	tweetid	userid
drizzy	15	adele
drizzy	61	adele
potus	33	potus
potus	61	adele

Show the output of each of the following queries.

```
SELECT who
FROM Likes natural full join Tweet
WHERE tweetID < 30;
```

```
SELECT t.userID, count(*), max(date)
FROM Tweet t, Likes l
where t.userID = l.who
GROUP BY t.userID;
```

```
SELECT count(*)
FROM User, Likes;
```

```
SELECT DISTINCT date
FROM User NATURAL JOIN Tweet;
```

```
SELECT T.TweetID, max(date)
FROM Tweet T, Likes L
WHERE T.userid = L.userid
GROUP BY T.TweetID
HAVING count(*) < 1;
```

```
(SELECT userID
FROM Likes)
  EXCEPT ALL
(SELECT userID
FROM Tweet);
```

Question 4. [4 MARKS]

Write a query to find the name of users who have only liked tweets posted on the same date that they have made a tweet. Ensure that your query would work on any instance of the database, not simply the one above.

[Use the space below for rough work. This page will not be marked unless you clearly indicate the part of your work that you want us to mark.]

1: _____/12

2: _____/ 8

3: _____/ 6

4: _____/ 4

TOTAL: _____/30