Remark Request Form   CSC343H, Fall 2017

If you are requesting a test remark, staple the test to this form. Otherwise, we can
access your work through MarkUs or PCRS.

Item (circle one)	A1	(midterm)	A2	A3	lecture prep #:
Given Name	Ruijie				
Family Name	Sun				
Student Number	1003326046				
Section (circle one)	L0101	(L0201)		L5101	
Date of request	2018.02.26				
UofT email address	jerry.sun@mail.utoronto.ca				
Teachin lab account	sunruij2				
Partner's name					
Partner's lab account					

Describe the reason for your request. Please be clear and concise! You **must** include a description of
the marking errors you have found. Please be aware that your entire item may be remarked.

There is calculation error for Q4 of midterm.

I should get 7.5/10 instead of 6.5/10.

UNIVERSITY OF TORONTO
Faculty of Arts and Science

Midterm Test

CSC 343H 2018, 9:10am
Duration — 50 minutes
No aids allowed

Student Number: 1, 0, 0, 3, 3, 2, 6, 0, 4, 6

Last Name: _Sun_

First Name: _Ruijie_

Enrolled Section (circle one): L0101/L2001/L2003(12:00) (L0201)/L2201(9:00) L5101/L2501(18:00)

Attending (for returning exam): L0101/L2001/L2003(12:00) (L0201)/L2201(9:00) L5101/L2501(18:00)

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 5 questions on 10 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.　　[15 MARKS]

As this question considers relational algebra, assume all relations are sets containing no nulls. Please write integrity constraints in relational algebra. Keys are underlined.

- Salesrep(<u>agent</u>, <u>region</u>, type)
  Indicates that a salesrep, identified by an agent name and region (like 'Canada') has a certain type.

- Customer(<u>cid</u>, name, agent, region)
  Each customer has a name and is assigned a salesrep (indicated by the agent and region).

- Product(<u>pid</u>, name)

- Service(<u>sid</u>, name, startDate, endDate)

- Purchase(<u>cust</u>, <u>item</u>, date)
  Indicates that customer cust purchased item.

- $\Pi_{agent,region}(Customer) \subseteq \Pi_{agent,region}(Salesrep)$

Part (a)　[3 MARKS]

Write the following integrity constraint in relational algebra: the cust attribute of <u>Purchase</u> is a foreign key of Customer.

$$Purchase[cust] \subseteq Customer[cid] \quad \checkmark \; ③$$

Part (b)　[3 MARKS]

Does the schema enforce this integrity constraint: a customer can only purchase an item at most once? Circle one and explain your answer:

Yes　　(No)

If yes, explain; if not write a new integrity constraint in relational algebra to enforce it.

$$Purchase(\; \underline{cust,\; item}\;,\; date)$$

Part (c)　[3 MARKS]

Does the schema enforce this constraint: the item in Purchase must either be a Product or a Service? Circle one and explain your answer:

Yes　　(No.) $\checkmark$

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

$$Purchase[item] - \left(Product[pid] \; \cup \; Service[sid] \right) = \phi$$

③

Part (d) [3 MARKS]

Suppose $|\Pi_{cid}(Customer)| = 100$ and $|\Pi_{region}(Salesrep)| = 4$. How many tuples could the relation Salesrep have? Circle all that apply and explain your answer:

0   ~~2~~   4   100   1000

(handwritten) $+3$ $+3-1 = 2$

$R1$
$R2$
$R3$
$R4$

$C100$

# of tuples in Salesrep is at least 1

most

Part (e) [3 MARKS]

Does your answer to Part(d) change if we add the constraint $\Pi_{agent}(Salesrep) \subseteq \Pi_{cid}(Customer)$? Explain.

(handwritten) Yes. Answer will be 2, 4, 100

# of tuple in Salesrep is at most 400. 3

Question 2. [9 MARKS]

Which of the following pairs of queries are equivalent? Circle each pair that returns the same results on all database instances. Consider only the schema and constraints that were given at the start of the questions (not any constraints added in Parts (a-e)). Explain for partial credit. For example, if they are not equivalent, you could give a small instance where the two queries differ. If they are equivalent, explain.

1. $\Pi_{name}(Customer \bowtie Salesrep) = \Pi_{name}(Customer \times Salesrep)$

(handwritten) Equivalent

Both list all name of customer

2. $\Pi_{pid}Product - \rho_{R(pid)}(\Pi_{item}Purchase) = \Pi_{pid}Product - (\Pi_{pid}Product \cap \rho_{R(pid)}(\Pi_{item}Purchase))$

(handwritten) Equivalent

Both list id which in Product but not in Purchase

3. $\Pi_{agent}(\sigma_{region='Canada'}Customer) = \Pi_{agent}(\sigma_{region='Canada'}Salesrep)$

(handwritten) No

Customer

cid	name	agent	region
1	Jerry	"A"	"Canada"

Salesrep

agent	region	type
"A"	"Canada"	"car"
"B"	"Canada"	"car"

CONT'D...

CSC 343H

MIDTERM TEST

Question 3. [8 MARKS] (b)

Here is part of the schema from Assignment 1. Assume there are no constraints on the schema other than those stated. Keys are underlined.

- User(uid, name, profile, email, phone, photo, lastSeen)

    A tuple in this relation represents a WhatsUpp user, uid is a unique identifier. name, profile, email, and phone are information about this user. photo is the url of the profile photo of this user. lastSeen is the last time this user used WhatsUpp. A larger time is more recent than a smaller time.

- Contact(user, contact, start)

    A tuple in this relation represents the fact that the user with identifier user may send messages to the user contact after the time start. Both user and contact are foreign keys for User.

- Group(gid, uid)

    uid is a member of group gid. The attribute uid is a foreign key for User.

Write a relational algebra query to return every user (her uid and name) that have no contacts with start value (Contact[start]) later than her lastSeen value. As in the assignment you can use $<, >, \leq$, etc. to compare all attribute values. A later date has a higher value.

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

user.uid=contact
·use

$$\text{All Possible } (uid, contact, start, lastSeen) = \Pi_{uid,} \left(\frac{User \times Contact}{\text{find correspo}} \right)$$

contact,
start, contact, not
lastSeen everything.

$$\text{Non User } (uid) = \Pi_{uid} \left(\sigma_{start > lastseen} (All\ possible) \right)$$

-ing

-2

$$\text{Target } (uid) = \Pi_{uid} (User) - \text{Non User}$$

$$\text{Answer } (uid, name) = \Pi_{uid,\ name} \left(Target \bowtie User \right)$$

Question 4. [10 MARKS]

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

Product

pid	name
1	ipad
2	ipod
10	macpro
11	blackberry

Service

sid	name	startDate	endDate
5	voice	2018-01-01	2018-06-01
10	data	2018-06-01	2018-12-31
11	data	2018-01-01	2018-06-01
15	text	2018-01-01	2018-12-31

Customer

cid	name	agent	region
1	Xi	Mo	Malta
2	Ty	Pa	Canada
3	Ma	Pat	US
4	Pam	Jo	Kyrgyzstan

Purchase

cust	item	date
1	10	2018-10-16
1	11	2018-06-01
2	5	2018-01-01
2	1	2018-01-01

Show the output of each of the following queries. The output schema is determined by the query, you should give the tuples. If a query is ill-formed (incorrect SQL), state this and explain why.

```
SELECT name
FROM Service, Purchase
WHERE date = startDate;
```

name
-----
voice
voice
data
data
data
text
text

2/2

```
SELECT date, count(*) as numTuples, count(distinct cust) as numCust
FROM Purchase P, Product Pd, Service S
WHERE item = pid or item = sid
GROUP BY date;
```

date	numTuples	numCust
2018-01-01	2 ~~8~~	1
2018-06-01	2 ~~7~~	1
2018-10-16	2 ~~7~~	1

1/2

```
SELECT item, cid, region, count(*)
FROM Customer, Purchase
WHERE cid = cust
GROUP BY item;
```

No

there is no cid, region in Group ✓

2/2

```
(SELECT date
FROM Customer, Purchase
WHERE cid = cust)
   EXCEPT ALL
(SELECT startDate
FROM Service);
```

2018-12-16 ✓

2/2

```
SELECT DISTINCT Customer.name, Product.name
FROM Customer, Product, Purchase
WHERE cid = cust and item = pid
ORDER BY Customer.name, Product.name;
```

Customer.name	Product.name
Ty	ipad ✓
X₁	blackberry
X₁	mac pro

0.5/2

CONT'D...

Student #:

Question 5.   [8 MARKS]

Write an SQL query that does the following. For every product that has been purchased more than twice by customers in the Canadian region, return the product id along with the first date and last date on which it was sold to a Canadian customer.

```
Creat View    Target Product  As
  select distinct p1.pid
  from    Product p1 Purchase p2   Customer C
  where   p1.pid = p2.item  and   C.cid = P2.cust  and C region
                                                        = Cana
                                                            da
  group by  p1.pid
  having    count(*) >= 2


  select  T.pid, min (P.date), max (p.date)
  from    Target Product T, Purchase P
  where   T.pid = P.item
  group by  T.pid
```

[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: _____11_____/15

\# 2: _____9_____/ 9

\# 3: _____6_____/ 8

\# 4: _____5,5_____/10

\# 5: _____8_____/ 8

TOTAL: _____79_____/50