CMPT 354: Database Systems I
Summer 2019
Midterm Exam (June 27, 2019)

First Name:
Last Name:
Student Number:

Instructions:

1. This is a closed book examination.
2. No electronic devices may be used.
3. Please write down your answers using a pen.
4. Answer each question in the space provided. You should not need more space.
5. Exam duration is 75 minutes (12:45 – 2:00pm)
6. This exam includes 10 questions. Answer all the questions.
7. The exam is 8 (eight) pages. Make sure you have all of the pages.

<table>
<thead>
<tr>
<th>Question</th>
<th>Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 (20 points)</td>
<td></td>
</tr>
<tr>
<td>Q2 (10 points)</td>
<td></td>
</tr>
<tr>
<td>Q3 (6 points)</td>
<td></td>
</tr>
<tr>
<td>Q4 (12 points)</td>
<td></td>
</tr>
<tr>
<td>Q5 (9 points)</td>
<td></td>
</tr>
<tr>
<td>Q6 (15 points)</td>
<td></td>
</tr>
<tr>
<td>Q7 (5 points)</td>
<td></td>
</tr>
<tr>
<td>Q8 (8 points)</td>
<td></td>
</tr>
<tr>
<td>Q9 (10 points)</td>
<td></td>
</tr>
<tr>
<td>Q10 (5 points)</td>
<td></td>
</tr>
<tr>
<td><strong>Total (100)</strong></td>
<td></td>
</tr>
</tbody>
</table>
Q1 (20 points). Design an E/R diagram describing the following domain:

- A **person** has attributes id (key), fname, and lname.
- A **company** has attributes id (key) and name.
- A **party** has attributes pid, date, and duration. The date attribute cannot be NULL. A Party cannot last longer than 4 hours.
- A party is **booked** by either a person or a company. We need to capture information about both when the party takes place and when the party was booked. The booking date must be earlier than the party date. A party is booked by zero or one organizers. A person or company can book many parties.
- A party occurs in either a house, an apartment, or a suite. All these locations are uniquely identified with a location id (lid). A house has an attribute street address. The address is an atomic string attribute. An apartment additionally has an apartment number. A suite has a street address and a suite number. A party occurs in zero or one place. A place can host multiple parties.
- A company occupies zero or one suite and a suite hosts zero or one company.
- A person lives in either a house or an apartment. A person lives in zero or one place. A given place can host many people.

Draw your E/R diagram for the described domain in the following space.
Q2 (10 points). Write the CREATE TABLE statements necessary to capture the subset of the above E/R diagram that corresponds to party, person, company, and booking. Include only the primary key, and foreign key constraints.

Q3 (6 points). Write down the relational algebraic representation of the expression represented by the tree below (3 points). Then write the SQL query that it represents (3 points).
Q4 (12 points). Parse the following SQL query using a parse tree (4 points). Then, write down the tree representation of the relational algebraic notation (4 points). Then, write down the RA representation of the query (4 points).

```sql
SELECT name FROM User
WHERE city = 'Vancouver' AND age='18'
```

Q5 (9 points). What is a dense index? (2 points) What is a sparse index? (2 points) Argue usefulness of a sparse index (3 points). What is a secondary index? (2 points)
Q6 (15 points). Derive the division operator on relations R and S based on base operations (Select, Project, Union, Set difference, and Cross Product) (6 points). Show your steps with an example (6 points). Define the properties the relations should hold for a division to have a meaningful result (3 points).

Q7 (5 points). We have a database including the following tables.

Movies (title, year, length, genre, director, studioName, producerCNum)
Director (name, address, gender, birthdate)
Made(MovieTitle, movieYear, directorName)
Studio (name, address, president)

Write a query to show movie title, movie year, and studio name of all movies made by “Paramount” studio. Group the results based on the movie year.
Q8 (8 points). What are the steps involved in query processing? What are the tasks at each step? At which step(s) do we perform query optimization?

Q9 (10 points). You are working with a customer who just designed their E/R diagram and started to create their schemas using the following statements.

CREATE TABLE Product (pid INT PRIMARY KEY, name VARCHAR(20))

CREATE TABLE Inventory (pid INT PRIMARY KEY, quantity INT,
FOREIGN KEY (pid) REFERENCES Product)

CREATE TABLE Supplier (sid INT PRIMARY KEY, name VARCHAR(20),
address VARCHAR(50))

CREATE TABLE Supplies (sid INT REFERENCES Supplier, 
pid INT REFERENCES Product, 
PRIMARY KEY (sid, pid))

CREATE TABLE PurchaseOrder (pid INT REFERENCES Product, quantity INT)

The customer would like to add certain constraints to his database. Advise your customer on the appropriate implementation for each of the following constraints. If appropriate, show how to modify the CREATE TABLE statements:
(a) Ensure that attribute quantity from table **Inventory** is always greater than or equal to zero **(5 points)**.

(b) Whenever someone deletes a tuple in **Supplier**, any tuple in **Supplies** that referred to it should also be deleted **(5 points)**.

**Q10 (5 points).** Depict the resulting B-Tree after completing the deletion of value 12 from the following B-Tree.