Information and Database Management Systems I (CIS 4301)
(Spring 2017)

Instructor: Dr. Markus Schneider

TA: Yang Peng

Homework 2

<table>
<thead>
<tr>
<th>Name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>UFID:</td>
<td></td>
</tr>
<tr>
<td>Email Address:</td>
<td></td>
</tr>
</tbody>
</table>

Pledge (Must be signed according to UF Honor Code)

On my honor, I have neither given nor received unauthorized aid in doing this assignment.

_______________________________________________
Signature

For scoring use only:

<table>
<thead>
<tr>
<th></th>
<th>Maximum</th>
<th>Received</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise 1</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Exercise 2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Exercise 3</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Exercise 4</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
Exercise 1 (Knowledge Questions) [25 points]

1. [5 points] What is a DBMS? What are the processes that a DBMS facilitates?

2. Consider a database of employees in which we need to record information about employees’ addresses. Name at least two conditions that would cause you to make “address” an entity set of its own rather than an attribute of the employee entity set.

3. [5 points] What is the difference of generalization and aggregation?

4. [5 points] List the basic operations of the Relational Algebra with their name and their correct symbolic notation.

5. [5 points] What is a key? Should a key contain only a single value? If yes, explain why. If not, give an example.

Exercise 2 (Relational Algebra) [20 Points]

Consider the following database schema:

Flights (flightNumber: string, travelFrom: string, travelTo: string, distance: integer, departs: time)

Aircraft (planeId: string, planeName: string, range: integer)

Pilots (employeeId: string, planeId: string)

Employees (employeeId: string, employeeName: string, salary: integer)

Express the following colloquial queries in Relational Algebra:

1. [5 points] Find the employee Id’s of pilots who can operate the aircrafts of types "Boeing 747" and "Boeing 777".

2. [5 points] Find the id of the pilot, or pilots, with the highest salary.

3. [5 points] Find the airplane Id's of aircrafts that cannot fly non-stop from ATL to JFK.

4. [5 points] Find the names of pilots who can operate planes with a range greater than or equal to 1,500 miles but cannot operate "Boeing 747" and "Boeing 777" aircrafts.
Exercise 3 (Relational Algebra) [30 points]

Consider the following database schema for a simplified Twitter database (primary keys are underlined):

- Person(pname, city, ttext) – Assume the pname is unique
- Follows(pname1, pname2) – Person pname1 follows person pname2
- Tweets(tid, ttitle, ttext) – Tweet with tid has title ttitle and text ttext
- PersonTweets(pname, tid, ts) – Person pname posted tweet tid at timestamp ts
- TweetTag(tid, tagname) – Tweet tid had tagname in its list of tags

Users post ‘tweets’, that is short pieces of text. They may tag their tweets with zero or more tags of their own choice. For example, a user tweeting about the G20 summit may decide to use the tag ‘G20’. A user ‘u’ may follow zero or more other users, which means that their ‘tweets’ are visible to user ‘u’ when he/she logs in.

For your convenience, you can use the follow abbreviations: P for Person, F for Follows, T for Tweets, PT for PersonTweets, TT for TweetTag. Express the following queries in relational algebra assuming set semantics.

1. [5 points] Find all the people (pname) who posted a tweet with tag ‘Obama’.
2. [5 points] Find all the tags ‘Larry Page’ uses in his tweets. (i.e. Larry’s tweeting interests)
3. [5 points] Find all the tags ‘Larry Page’ reads in the tweets of the people he follows. (i.e. Larry’s reading interests)
4. [5 points] Find all the people (pname and city) who follow people who follow ‘Tim Cook’. (i.e. second-level followers).
5. [5 points] Find all the people (pname and city) who follow at least everyone that ‘Larry Page’ follows.
6. [5 points] Find all pairs of people (pname) who have at least one follower in common.
Exercise 4 (Relational Calculi) [25 points]

Consider the following database schema:
EMPLOYEE (emp_name, address_street, address_city, phone_no)
DEPARTMENT (dept_name, mgr_name, mgr_start_date)
WORKS_ON (emp_name, industry_name, salary)
INDUSTRY (industry_name, location_city)
MANAGES (emp_name, mgr_name)

Write Tuple Relational Calculus expressions for problems 1 and 2. Write Domain Relational Calculus expressions for problems 3, 4, and 5.
1. [5 points] List the names of all employees who don’t work for ‘Verizon Wireless’ and make more than $50,000 per year.

2. [5 points] Find the names and phone number of all employees who work for Intel and live in Santa Clara.

3. [5 points] Find the names of all employees who live in the same city and on the same street as their managers.

4. [5 points] Find the names of all employees in this database who live in the same city as the industry they work on is located in.

5. [5 points] Find the names of all industries located in every city in which Intel is located.