NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 1 EXAMINATION 2019-2020

CZ2007 – INTRODUCTION TO DATABASES

Nov/Dec 2019 Time Allowed: 2 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 6 pages.
2. Answer **ALL** questions.
3. This is a closed-book examination.
4. All questions carry equal marks.

---

1. Consider a database system comprising six relation schemas as shown below. The six relations record information about customers dining at restaurants and shopping at shopping malls. Primary keys are underlined.

CUSTOMER(custName, address, ageGroup)
RESTAURANT(restName, address, capacity)
SHOPPING_MALL(mallName, address, openingHours)
SHOP_AT(custName, mallName, date, time, dayOfWeek)
DINE_AT(custName, restName, date, time, dayOfWeek)
LOCATED_AT(restName, mallName, levelNumber, unitNumber)

(a) Construct an ER diagram corresponding to the six relational schemas.

(5 marks)

(b) Using relational algebra, find those restaurants that are frequented by the largest number of repeat customers in the 30s-40s age group. Repeat customers of a restaurant are customers who have dined in the restaurant at least once.

(10 marks)

Question No. 1 continues on Page 2

(c) Find those shoppers in the 20s-30s age group who have never shopped at Nanyang Shopping Mall on Fridays between 7pm to 10pm. Also find out which streets these shoppers live in.

(10 marks)

2. Consider the relation R(A, B, C, D, E) with the following functional dependencies:

$$AB \rightarrow C \qquad B \rightarrow D \qquad DE \rightarrow A$$

(a) Is the schema R in BCNF? Explain your answer by first stating the definition of BCNF.

(10 marks)

(b) Use one step of the BCNF decomposition algorithm to decompose R into two sub-relations. You only need to do the first decomposition.

(5 marks)

(c) Is your decomposition in Part 2(b) lossless and dependency preserving? Explain your answer by first stating the definitions of lossless and dependency-preserving decomposition.

(10 marks)

3. (a) Consider a database of social groups that allows people to become members of groups: a person can be a member of several groups and each group maintains a list of pictures that are accessible to all members. In addition to the groups, the database also maintains a list of friends. The schema is:

MEMBER(personName, groupName)
PICTURE(groupName, picture)
FRIEND(personName1, personName2)

Question No. 3 continues on Page 3

PICTURE stores for each picture the name of the group that owns that picture. The FRIEND table is symmetric, i.e., if X is friend with Y then Y is friend with X. Every person is a member of at least one group.

Give an expression in SQL for each of the following queries:

(i) Write an SQL query that computes for every person the total number of pictures that she can access through her group memberships. That is, a person X can access a picture Y if X is a member of some group Z, and Z owns the picture Y. You need to write a SQL query that returns a result like this:

'Andy', 12
'Sara', 7
'Sue', 0
'Raj', 9
...

(5 marks)

(ii) The marketing department has decided to recommend people to subscribe to additional groups. They would like to recommend to a person X to subscribe to a group Y if all X's friends are members of the group Y, but X is not a member of Y. Write a SQL query that computes for each person X the set of groups to recommend. You need to write an SQL query that returns a list of (person, group) pairs.

(5 marks)

(iii) Create a new table ACCESS(personName, picture) that lists for each person the list of pictures that he or she can access. A person X can access a picture Y either if X belongs to a group that owns Y, or if X has a friend Z who belongs to a group that owns Y. Write SQL statements that insert the corresponding tuples in ACCESS. You need to write one or more INSERT queries.

(5 marks)

(b) Suppose that the table EMPS(<u>eid</u>, name, salary, bonus) contains only two tuples:

(10, 'Joe', 1000.0, NULL)
(20, 'Sally', NULL, 2000.0)

Question No. 3 continues on Page 4

For each of the following SQL queries, show the output that it would produce if runs against EMPS, and explain your answer.

(i) SELECT salary + bonus AS answer FROM EMPS WHERE name = 'Joe';

(5 marks)

(ii) SELECT SUM(salary) AS answer FROM EMPS;

(5 marks)

4. (a) Consider the following schema for the Australian Competitive Team Boomerang League (ACTBL). The database tracks information about teams, players, and the outcomes of games they have played. Each game consists of a "home" team and an "away" team. The ACTBL requires that no games can end in a tie. The winner of a game is (obviously) the team with the most points. Assume that every team has played at least one game at home and at least one game away so far. The schema is given below.

TEAMS (TName, TCity)

PLAYERS (PName, Team, Salary)
[where Team is a foreign key referencing TEAMS]

GAMES (HomeTm, AwayTm, Date, HomePts, AwayPts)
[where HomeTm and AwayTm are foreign keys referencing TEAMS]

Write SQL statements to print the name of the team that has won the most home games. (Hint: you might find it easier if you use views).

(5 marks)

(b) Consider a shop for flowers on the World Wide Web. It offers a selection of different bouquets (bunches of flowers) described in the table "FLOWERS".

FLOWERS(Prod, Name, Description, Price)

The table "CUSTOMER" contains information about customers.

Question No. 4 continues on Page 5

CUSTOMER(Cust, CName, CAddress, CCity, CZip, CState)

Next, there is a table with information about orders and recipients (to whom the flowers should be delivered). "Prod" identifies the product, i.e. which bouquet should be delivered, "Cust" identifies the customer (who ordered the flowers), and "Date" is the date when the flowers should be delivered. The rest of the information is for the person receiving the flowers.

ORDER(Ord, Prod, Cust, Date, RName, RAddress, RCity, RZip, RState)

Some of the flowers are sent via FedEx and other orders are given to local flower shops. Only flowers that have not yet been delivered are kept in these two following tables.

MAILED(Ord, MDate, Tracking-No)
SUBCONTRACTED(Ord, Shop, Phone)

(i) Use assertion constraint on the field CZip in the CUSTOMER table to ensure that the input value has five digits (between 10000 and 99999).

(5 marks)

(ii) Write a trigger to delete all the orders, mails, and subcontracts of a customer if a person is deleted from the CUSTOMER table.

(5 marks)

(c) Recall that database indexes provide faster access to data for operations that return a small portion of a table's rows. You can create an index on any column; however, if not created wisely, it may not improve the performance, and the index takes up resources unnecessarily. Specify two scenarios when creating an index on a column is recommended.

(5 marks)

(d) A database with the following schema stores a collection of Webpages and the words they contain, and a collection of dictionaries in several languages and the words in those languages:

Occurs(url, word)
Dictionary(language, word)

Question No. 4 continues on Page 6

- url represents a Webpage.
- Every Webpage may contain several words, and every word may occur in several Webpages.
- Every language may contain several words, and every word may occur in several languages
- There are no nulls in the database.

Write an XML DTD, representing the same schema as above.

(5 marks)

END OF PAPER

CZ2007 INTRODUCTION TO DATABASES

Please read the following instructions carefully:

1. **Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**

2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.

3. Please write your Matriculation Number on the front of the answer book.

4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.