NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 1 EXAMINATION 2016-2017

CZ2007 – INTRODUCTION TO DATABASES

Nov/Dec 2016 Time Allowed: 2 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 7 pages.
2. Answer **ALL** questions.
3. This is a closed-book examination.
4. All questions carry equal marks.

---

1. (a) Consider a database for a hotel chain, Notlih, with the following requirements:

- Notlih has a number of hotels, each of which has an ID, a name, an address, and a hotel rating.
- Each hotel room has a unique number within the hotel. Information recorded for each hotel room includes whether it is a non-smoking room and whether it is wheelchair accessible.
- Rooms in a hotel are categorized into different types. Information recorded for each type includes type name, room size, number of single beds, and number of double beds.
- A hotel room may be reserved 12 months in advance. Notlih assigns a price for each room type for each day in those 12 months. For example, the price for a room type may be more expensive than usual during Christmas.
- Each hotel customer has an ID, a name, a date of birth, and a gender.

Question No. 1 continues on Page 2

- A customer may make one or more reservations at each hotel. Each reservation has a room type, a check-in date, and a check-out date. The price of the room type for each date in the stay period is also recorded.

- When a customer checks into a hotel, Notlih retrieves the reservation made by the customer, and assigns hotel room(s) to the customer based on the room type reserved. Note that the room assigned to the customer may change during his/her stay. Notlih records information about the actual room assignment for each day.

(i) Construct an ER diagram that captures the above requirements.

(10 marks)

(ii) Convert the ER diagram into a set of tables. Indicate the primary keys.

(4 marks)

(b) Consider a train database with the following schema:

```
TRAIN(TID, Source, Destination, Date, Time,
      NumberOfSeats)
PASSENGER(PID, Name, Age, Gender)
TICKET(TID, PID, TicketClass)
```

`TRAIN` and `PASSENGER` record information about trains and passengers, respectively. Whenever a passenger books a ticket for a train, the `TICKET` table records the passenger's ID and the train's ID. `TICKET.TID` and `TICKET.PID` are foreign keys referencing `TRAIN.TID` and `PASSENGER.PID`, respectively.

Answer each of the following queries with relational algebra. You may use the following operators: σ (selection), Π (projection), $\cup$ (union), $\cap$ (intersection), $-$ (difference), γ (grouping and aggregation), δ (duplicate elimination), $\div$ (division), $:=$ (assignment), ρ (rename), $\bowtie$ (join), $\dot{\bowtie}_L$ (left outerjoin), $\dot{\bowtie}_R$ (right outerjoin), $\dot{\bowtie}$ (full outerjoin).

(i) Find the number of passengers who have booked a train ticket from Seoul to Busan.

(2 marks)

Question No. 1 continues on Page 3

(ii) Find the IDs of the passengers who have booked a train from Seoul to Busan as well as a train from Busan to Seoul on the same date.

(2 marks)

(iii) Find the IDs of the trains that have been fully booked (i.e., the number of seats in the train equals the number of passengers who have booked the train).

(3 marks)

(iv) Let's define a *frequent traveller* as a passenger who has booked more than 100 tickets. Find the "source, destination" pair that has been booked the largest number of times by frequent travelers.

(4 marks)

2. Consider a relation R(A, B, C, D, E, F) with the following functional dependencies: AF→D, B→C, BE→D, BDA→F, CE→A, D→E.

(a) Verify whether R is in BCNF. If R is not in BCNF, use the BCNF decomposition algorithm to normalize R, and then verify whether your BCNF decomposition preserves all functional dependencies. If not all functional dependencies are preserved, specify one functional dependency that is not preserved and explain why.

(15 marks)

(b) Verify whether R is in 3NF. If R is not in 3NF, use the 3NF decomposition algorithm to normalize R.

(10 marks)

3. Consider the following relational schema:

```
Reader( RDNR, Surname, Firstname, City, Birthdate )
Book(   ISBN,   Title,   Author,   NoPages,   PubYear,
     PublisherName )
Publisher( PublisherName, PublisherCity )
Category( CategoryName, BelongsTo )
Copy( ISBN, CopyNumber, Shelf, Position )
Loan( ReaderNr, ISBN, Copy, ReturnDate )
BookCategory( ISBN, CategoryName )
```

Question No. 3 continues on Page 4

Each book may have multiple categories, which are stored in the `BookCategory` table. A category can have multiple subcategories, but a category may have only one parent category – this information is stored in the `Category` table. The `Reader` table stores information about individual readers. The `Book` table stores information about books. Each book can have multiple copies, which are stored in the `Copy` table. Information about publishers are stored in the `Publisher` table. When a reader borrows a copy of a book, the corresponding information is stored in the `Loan` table.

(a) Formulate the following queries in SQL.

(i) Which categories do not have any subcategory?

(5 marks)

(ii) Which readers have borrowed all the books (by ISBN, not by copies) from the author "Jiawei Han"?

(5 marks)

(iii) For which of the books there is at least one copy available?

(5 marks)

(b) Enforce that a reader can only borrow up to 20 books. Give a solution using CHECK and a solution using Trigger.

(6 marks)

(c) Formulate in SQL the following modifications. Change the return date of all the books in the category "Databases" that should be returned before 15.03.2013 so that they can be kept for 30 days longer (Assume that you can add days to dates in SQL).

(4 marks)

4. (a) Express left outer join, right outer join, and full outer join using the following operations of relational algebra: σ (selection), Π (projection), $\cup$ (union), $\cap$ (intersection), $-$(difference), $:=$ (assignment), ρ (rename), and $\bowtie$ (join).

(6 marks)

Question No. 4 continues on Page 5

(b) Consider the following relation R and its decomposition into R1 and R2.

```
R( Student_ID, Date_Enrolled, Course_ID, Room NR,
     Professor )
R1( Student_ID, Date_Enrolled, Course_ID )
R2( Date_Enrolled, Room NR, Professor )
```

For this question, assume that each course is taught only by one professor and takes place in just one room. Also, when Student_ID and Course_ID are given, Date_Enrolled is unique.

(i) Show that the decomposition of R to R1 and R2 is a lossy decomposition.

(4 marks)

(ii) Find and explain a lossless decomposition of the same relation R.

(3 marks)

(c) Consider the simple relation Employee(ID, salary) storing the employee IDs and salaries, where ID is a key. Consider the following two triggers over this relation:

Trigger T1:

```
CREATE TRIGGER T1
AFTER INSERT ON Employee
REFERENCING NEW as New_Emp
FOR EACH ROW
UPDATE Employee
SET salary = 1.1 * (SELECT max(salary)
                    FROM Employee)
WHERE ID = New_Emp.ID
```

Question No. 4 continues on Page 6

Trigger T2:

```
CREATE TRIGGER T2
AFTER INSERT ON Employee
REFERENCING NEWTABLE AS New_Emp
FOR EACH STATEMENT
UPDATE Employee
SET salary = 1.1 * (SELECT max(salary)
                    FROM Employee)
WHERE ID IN (SELECT ID FROM New_Emp)
```

Assume that relation `Employee` has no tuple initially. Suppose that we had inserted the following four rows into the `Employee` table as the result of a single SQL statement:

```
1  1000
2  2000
3  3000
4  4000
```

(i) Show the final database state after trigger execution if only trigger `T1` is defined.

(3 marks)

(ii) Show the final database state after trigger execution if only trigger `T2` is defined.

(3 marks)

(d) Consider the XML document **result.xml** in Figure Q4d. Write XPath expressions for the following queries posed on this document.

(i) Find all choice elements where the choice number (`choiceNum`) is 1 and the merit score (`meritScore`) is greater than 800.

(3 marks)

(ii) Find all codes (`code`) of the applicant *"Doreen"*.

(3 marks)

Question No. 4 continues on Page 7

```
<results>
   <applicants>
     <applicant name="Doreen" appNum="a1"/>
     <applicant name="Dilwyn" appNum="a2"/>
     <applicant name="Suzanne" appNum="a3"/>
   </applicants>
   <choices>
     <choice applicant="a1" code="MPSOF"
     choiceNum="1" meritScore="750"/>
     <choice applicant="a1" code="MPALG"
     choiceNum="2" meritScore="750"/>
     <choice applicant="a1" code="MPCSN"
     choiceNum="3" meritScore="800"/>
     <choice applicant="a2" code="MPALG"
     choiceNum="1" meritScore="700"/>
     <choice applicant="a3" code="MPCSN"
     choiceNum="1" meritScore="850"/>
     <choice applicant="a3" code="MPALG"
     choiceNum="2" meritScore="850"/>
   </choices>
</results>
```

Figure Q4d

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.