NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 2 EXAMINATION 2020-2021

CZ2007 – INTRODUCTION TO DATABASES

Apr/May 2021 Time Allowed: 2 hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 5 pages.
2. Answer **ALL** questions.
3. This is a closed-book examination.
4. All questions carry equal marks.

---

1. Consider a database comprising eight tables as shown below. The eight tables record information about drivers, student drivers, professional drivers, passengers, travels done by passengers, trains, trips driven by drivers, and whether a driver is licensed to drive a particular type of train. Primary keys are underlined.

DRIVERS(**d#**, age, salary)
STUDENT-DRIVERS(**d#**, learning_license)
PROFESSIONAL-DRIVERS(**d#**, professional_license)
PASSENGER(**g#**, seat_preference)
TRAVELS(**g#**, **t#**, date)
TRAIN(**n#**, type, year_built)
TRIPS(**t#**, d#, n#, origin_station, final_station, date, day_of_week)
LICENSED(**d#**, **n#**, **year_attained**)

(a) Construct an ER diagram corresponding to the eight tables.

(10 marks)

(b) Using relational algebra, find all 2010-built trains driven by all student drivers. The student drivers are licensed to drive the trains from 2018 onwards.

(5 marks)

Question No. 1 continues on Page 2

(c) Using relational algebra, find all pairs of 50-year-old drivers who have driven trains that depart from Jurong Station the same number of times every Monday. Each pair of drivers must be different.

(10 marks)

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

$$A \rightarrow D, \quad AE \rightarrow H, \quad DF \rightarrow BC, \quad E \rightarrow C, \quad H \rightarrow E$$

(a) What are the keys (not superkeys) of R?

(5 marks)

(b) Using the closure method, decide whether each of the following functional dependencies is implied by the above functional dependencies.

(i) $EC \rightarrow AD$

(ii) $ADF \rightarrow E$

(iii) $A \rightarrow DH$

(iv) $AED \rightarrow C$

(v) $DH \rightarrow C$

(10 marks)

(c) What are BCNF and 3NF? Consider three relational schema R1(A, D), R2(E, C), and R3(A, B, E, F, H). They form a decomposition for R. Do the original functional dependencies apply in R1, R2, and R3? Is this decomposition in BCNF? Is this decomposition in 3NF? Explain your answers.

(10 marks)

3. Consider the following schema of a database (primary keys are underlined).

ITEM(**itemid**, name, price, producerid, category)
CUSTOMER(**customerid**, name, address, phone)
PURCHASE(**purchaseid**, customerid, date)
PURCHASE_ITEM(**purchaseid**, **itemid**, quantity)
PRODUCER(**producerid**, name, address, country)

The ITEM table records the information of all items in a library. The CUSTOMER table records the information of all customers. The PURCHASE table records the item purchasing information of a customer. The PURCHASE_ITEM table records what items a customer purchases in one transaction in the PURCHASE table. The PRODUCER table records the producer information of items.

Answer each of the following queries with an SQL statement.

(a) Find the names and ids of customers who have purchased items on 3 Dec. 2019 that are produced in Denmark.

(5 marks)

(b) Find the ids and names of the customers who bought both an item of category 'diary' produced in 'AUS' and an item of category 'coffee' produced in 'SIN' in a single purchase.

(5 marks)

(c) Return the ids, names, and prices of items that have the largest amount of sales in 2019, which is computed by the product of price and quantity (if more than one item have the largest sales, return them all). Note that you can use function YEAR(*date*) to return the year of a *date*.

(5 marks)

(d) Find the ids and names of customers who have purchased all items in the 'diary' category but not purchased any item in the 'coffee' category.

(5 marks)

(e) Find the ids and names of customers who have spent the **second biggest** amount of money in purchasing items.

(5 marks)

4. (a) Consider the database schema in Q3.

(i) Using SQL, create a view CategoryCountrySales(category, country, numberofitem, sales), to record for each category and country the number of distinct items that are produced in the country and belong to the category, as well as the sales for these items.

(4 marks)

(ii) Using SQL, create a temporary view CustomerCatItem(customerid, customername, category, numberofitem), to record for each customer the number of distinct items in each category purchased by the customer.

(4 marks)

(b) Consider the following schema of a database (primary keys are underlined):

EMPLOYEE (**person-name**, street, city)
WORKS (**person-name**, company-name, salary)
COMPANY (**company-name**, city)
MANAGES (**person-name**, manager-name)

(i) Create SQL triggers for the following constraint: When the company-name in an inserted record in WORKS table does not exist in COMPANY table, insert a record for the company in COMPANY table, where missing value is replaced by NULL if applicable.

(4 marks)

(ii) Create SQL triggers for the following constraint: Each manager can manage at most 5 employees and an employee cannot be his/her own manager. If this constraint is violated, report an error.

(4 marks)

(c) Consider relation CUSTOMER (**NRIC**, name, phone, address, email, sex, age).

(i) Build an index for relation CUSTOMER to best speed up the following query and explain your answer:

```
SELECT name, phone FROM CUSTOMER WHERE age > 30
AND sex = 'male'
```

(2 marks)

Note: Question Q4 continues on Page 5

(ii) Build an index for relation CUSTOMER to best speed up the following two queries, and explain your answer:

- SELECT name, phone FROM CUSTOMER WHERE age > 30 AND NRIC = 'S9999'
- SELECT name, phone FROM CUSTOMER WHERE NRIC = 'S8888' AND sex = 'male'

(2 marks)

(d) Consider the first four tables of the database in Q3, which are listed below again for convenience.

ITEM(**itemid**, name, price, producerid, category)
CUSTOMER(**customerid**, name, address, phone)
PURCHASE(**purchaseid**, customerid, date)
PURCHASE_ITEM(**purchaseid**, **itemid**, quantity)

Design a DTD for exporting this data as an XML view. The data in these tables satisfies the following constraint: Each item is purchased by at least one customer. In the XML view, the data is grouped by customers, and then by the items purchased by each customer. It is required that all the information in the database is reflected in the XML view.

(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.