NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 2 EXAMINATION 2022-2023

SC2207/CZ2007 – INTRODUCTION TO DATABASES

Apr/May 2023 Time Allowed: 2 hours

INSTRUCTIONS

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

---

1. Consider a database comprising six tables as shown. The six tables record information about visitors and restaurants located in theme parks. Some visitors bring their children together to theme parks and dine at restaurants in theme parks. Keys are underlined.

VISITOR(**vID**, country, ageGroup)
PARENT(**parentVisitorID, childVisitorID**)
RESTAURANT(**rID**, capacity, pID)
THEMEPARK(**pID**, address, capacity)
VISIT(**vID, pID, datetime**, dayOfWeek)
DINE(**vID, rID, datetime**, dayOfWeek)

(a) Construct an ER diagram corresponding to the eight tables. You may make suitable assumptions in order to construct the ER diagram.

(9 marks)

(b) Using relational algebra, find parents who brought their children to theme parks on Sundays.

(8 marks)

(c) Moreover, for theme parks visited by those families in Q1(b), find out which ones are the most frequently visited. (Note: You can make use of the intermediate relations created when solving for Q1(b).)

(8 marks)

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

$$AB \rightarrow C, \quad CD \rightarrow E, \quad C \rightarrow D$$

(a) Draw a directed graph to determine the key (not superkey) of R. Explain the steps in deriving the key.

(5 marks)

(b) Which functional dependencies violate the BCNF criteria? Apply a BCNF decomposition on R, and then verify whether your BCNF decomposition preserves all functional dependencies. Show your decomposition tree.

(10 marks)

(c) State the definition of 3NF. Verify whether R is in 3NF. If R is not in 3NF, apply a 3NF decomposition on R.

(10 marks)

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

PROFESSOR(**PID**, name, office, department, email)
COURSE(**CID**, title, department)
ENROLL(**CID, semester**, PID, totalEnrollment)
STUDENT(**SID**, sName, department)
STUDENT-ENROLL(**SID, CID, semester**, grade)

Attribute totalEnrollment describes the total number of enrollments for a specific course in a semester. Attribute PID is the professor id, attribute CID is the course id, and attribute SID is the student id. The CID attribute of ENROLL is a foreign key into the COURSE relation. Attribute semester describes calendar year and semester, and has distinct value for every semester. Attribute department for each department has distinct value.

Answer each of the following queries with a single SQL statement.

(a) Find the names and ids of students who have received the highest grade for "Database Systems" course on semester "2022S2".

(5 marks)

(b) Finds names of professors who taught both "Database Systems" course and "Introduction to Algorithm" course in the same semester.

(5 marks)

Note: Question Q3 continues on Page 3

(c) Find names of professors who taught all the courses in their departments.

(5 marks)

(d) Find names of students who have been taught by the greatest number of professors.

(5 marks)

(e) For each course, find the total number of **distinct** students enrolled in the course over the years, and the average grade of students.

(5 marks)

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

(i) Using SQL, create a view CourseAverageGrade(CID, coursetitle, maxgrade, number_of_enrollments), to record for each course, the maximum grade and the number of enrollments.

(3 marks)

(ii) Using SQL, create a temporary view ProfessorStudent(professorid, professorname, semester, number_of_student), to record for each professor and each semester, the number of students taught by the professor. Note if a student is enrolled in two courses of a professor, it is counted as twice.

(4 marks)

(b) For the database schema in Q3, create SQL triggers for the following constraint: Table STUDENT-ENROLL should be consistent with Table ENROLL in terms of the total number of enrollments for a specific course in a semester. Triggers should be able to maintain the records in Table ENROLL based on the records in Table STUDENT-ENROLL. Write all the necessary triggers to implement this constraint.

(8 marks)

(c) Consider relation STUDENT(**SNO**, name, phone, address, email, gender, age, CGPA, year_of_study).

(i) Build an index for relation STUDENT that can best speed up the following query and explain your answer:

```
SELECT name, phone FROM STUDENT WHERE age > 15
AND sno = 's12345678'
```

(2 marks)

Note: Question Q4 continues on Page 4

(ii) Build an index for relation CUSTOMER that can best speed up the following query, and explain your answer:

SELECT name, phone FROM STUDENT WHERE age > 20 AND CGPA = 4.5

(3 marks)

(d) Answer the following questions:

(i) Describe two differences between XML data and relational data.

(2 marks)

(ii) Give a solution to store XML data using relational databases.

(3 marks)

END OF PAPER

CZ2007 INTRODUCTION TO DATABASES
SC2207 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.