THE UNIVERSITY OF HONG KONG
FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE
COMP3278A Introduction to Database Management Systems

Date: Dec. 21, 2021
Time: 9:30pm – 12:30pm

University No. (Student ID): _____
Name & type of calculator used: _____

- The total mark of all questions is 100.
- Please write in the space provided. If you need to write outside the provided space, please state clearly the question number in your answer, for example, "Question 1(a)".
- Candidates are permitted to bring one sheet of A4-sized paper with printed or written notes on both sides to the examination.

(For examiners only)

Question No.	Mark
Question 1	/24
Question 2	/20
Question 3	/16
Question 4	/30
Question 5	/10
Total	**/100**

Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is candidates' responsibility to ensure that their calculator operates satisfactorily, and candidates must record the name and type of the calculator used on the front page of the examination script.

Question 1 [Entity-Relation Design] (24%)

An HKU DBMS team is building an online payment system. Please read the following requirements and answer questions (a) and (b).

Requirements.

- The main purpose of the system is to support online payment in two different ways: immediate payment and installment payment.

- Each **user** in the system is identified by a unique user ID. Each user has a username and a profile attribute. The profile stores the records of birthday, email address, and the date the user joined the system.

- A user owns one or more **accounts**. Each account in the system is identified by a unique account ID, and an account must be owned by a user. The user can login the account via account name and password. The remaining cash stored in the account is denoted as balance.

- A **transaction** is an order that needs to be paid. Each transaction record consists of a unique transaction ID, date and amount. Besides, it is possible to leave (attach) several messages along with the transaction.

- There are two payment methods for the transaction: immediate payment and installment payment. The transaction is specialized to either *Immediate Transaction* or *Installment Transaction*.

- For immediate transaction, it will be immediately paid by an account.

- For installment transaction, it needs an attribute to specify the number of installments, and creates several **sub-transactions**. Each sub-transaction can be uniquely identified by the combination of transaction ID and its sub-transaction ID. Each sub-transaction transfers a sub-amount of the total amount. Each sub-amount can be derived according to the number of installments (we assume that each sub-transaction has equal amount).

- Each sub-transaction is paid by one account.

(a) **(12%)** Please model the above requirements by drawing an E-R Diagram.
   - If a requirement is unclear, a reasonable assumption could be made. You must **clearly state** all such assumptions.

(b) (**12%**) Please reduce your E-R Diagram into relational table schemas. For each relation, underline the primary key and identify all foreign key(s), if any.

Question 2 [SQL] (20%)

Please consider the following table schema and answers questions from (a) to (d) by writing SQL queries.

Restaurant(*restaurant_id*, name)

RestaurantCategory(*restaurant_id*, *category*)

  Foreign key: {restaurant_id} references **Restaurant**

Branch(*restaurant_id*, *branch_no*, location, seats)

  Foreign key: {restaurant_id} references **Restaurant**

Member(*member_id*, name, birthday, joined, points)

  (the joined column is the year the member joined)

Visits(*visit_id*, member_id, restaurant_id, branch_no, date, score)

  Foreign keys: {restaurant_id, branch_no} references **Branch**,

  {member_id} references **Member**

(a) (**5%**) List the name and location of all fast food joints (restaurant category is "Fast Food"), and the name of the restaurant must start with 'F' or 'S'. Order the list by the restaurant's name and then by the branch location.

(b) **(4%)** Add 10 points per year to every member since the member joined. The current year is 2021. For example, if a member joined in 2018, (2021-2018)*10=30 points should be added. (The joined column in the Member table stores the year the member joined the membership)

(c) **(5%)** Count the number of unique member IDs visiting each of the restaurants. List the result by showing the restaurant name and the number of unique member IDs. A restaurant should be listed even if it is not visited by any member, in that case the count should be zero.

(d) **(6%)** Find all branches with at least two visits scoring larger than 5. List the restaurant name, the branch location, and the best score among the visits to the branch.

Question 3 [Relational Algebra] (16%)

(a) **(4%)** Please write a relational algebra expression using only the basic operators (σ, π, $\cup$, $-$, $\times$) to implement the division function ($\div$).

(b) Please consider the following relational tables, and answer the below questions by writing relational algebra expressions.

Employee (*employee_name*, age, city_code)

Company (*company_name*, city_code)

WorksFor (*employee_name*, *company_name*, workhour)

CityInfo (*city_code*, city_name)

(b.1) (4%) Find the names of all employees who live in a city different from where they work in. (A worker can work for multiple companies.)

(b.2) (4%) Find the names of employees who work for all the companies in the Company table.

(b.3) (4%) Find all companies where the longest working hour among their employees is below 40.

Question 4 [Database Normalization] (30%)

Given $R = (A, B, C, D, E)$ and a set of FDs on R, $F = \{B \to D, C \to D, B \to E, DE \to B, AD \to E\}$.

Please answer questions (a), (b) and (c).

(a) (**4%**) Is R in BCNF? Why? Clearly state your reasons.

(b) (**10%**) Find all candidate keys of R and clearly **prove** your results.

(c) (**16%**) Decompose R into relations that satisfy all the below three requirements. **Prove** clearly your results.

  1) The decompositions must be a lossless-join decomposition;

  2) The decompositions must preserve functional dependence;

  3) All decomposed relations must be in BCNF.

Question 5 [Database and AI] (10%)

Open question: How database can help the development of AI and how AI can help the development of database? Please list your reasons in point forms (no more than 5 reasons).

END OF PAPER