No book or other document is allowed.

Duration of the exam: 2.5 hours.

The total number of points is 100.
QUESTION 1 [10 points]
The Newark Engineering Group (NEG) has contacted you to create a conceptual schema that meets the database requirements for its training program. The NEG administrator gives you the following description for the training group's operating environment:

NEG has a number of instructors and can handle up to 30 trainees per class. Each trainee is identified by an ID, and his/her name, phone number and address are also recorded. NEG offers a number of advanced technology courses, each of which may generate several classes. Courses are identified by a course code. The classes of a specific course can be identified by a Class number but the same class number can be used for classes of different courses. For courses, a name and a description are also recorded, and for classes a room number, a duration and the possibly multiple date/times the class is offered. If a class has fewer than 10 trainees enrolled, it is canceled. It is, therefore, possible for a course not to generate any classes during a session. Each class is taught by one instructor. Each instructor is identified by an ID and his/her name and address are also recorded. Each instructor may teach up to two classes or no class at all. Each trainee may take up to two classes per session. The database records information for a single session.

Draw an ER diagram for the database of NEG assuming that the entity types are COURSE, CLASS, INSTRUCTOR, and TRAINEE. Specify also (a) keys, (b) structural constraints that is, participation constraints (single/double lines) and cardinality ratios (numbers across the diamond boxes), for every relationship type (c) (min, max) constraints on every participation of an entity type into a relationship type, and (d) roles (only if this is necessary).

State clearly any additional constraints you impose. The additional constraints (if any) should not contradict with the collected specifications.

QUESTION 2 [14 points]
Consider the following database schema.

The primary key attributes are underlined. Thus, IID, CID and SDAY together form the key for SALE. Foreign keys are shown with arrows. For instance, SALE.CID is a foreign key that refers to the primary key CID of CUSTOMER.

Write the following queries in **Relational Algebra**:  
1. Find the names of customers who bought a red and a green item.  
2. Find the item IDs of items which are not bought by any customer older than 50.  
3. Find the ids of items sold at a price higher than the average price of items of the same color.  
4. Find the names of customers who have bought all and only yellow items.
QUESTION 3  [28 points]

Consider the database schema of Question 2. Write the following queries and update operations in SQL.

1. Find the names of customers who did not buy any red item.
2. For each city, find the id(s) of its youngest customers. The answer of the query should have two columns: CITY and CUST_ID.
3. Find the average price of green items sold to less than 5 different customers along with the IDs and the names of the items (the schema of the answer should have three attributes).
4. For every customer who bought at least 5 different green items, print her CID and her name and the total number of purchases that she made.
5. Create a view that computes the average rating of customers per age (the schema of the view should have two attributes). Then use this view to write a query that finds all customer names whose rating is higher than the average rating for their age. No duplicates should appear in the answer.
6. Raise by 10% the rating of every customer who is younger than 50 and bought a green item.
7. Find the names of customers who bought all and only red items.

QUESTION 4  [8 points]

Consider the database schema of Question 2. Write SQL to enforce the following constraints. If a constraint cannot be expressed in SQL explain why.

1. No customer can buy more than 12 different red items.
2. The number of purchases of the customer(s) with the highest number of purchases must be at least twice as big as the number of purchases of the customer(s) with the fewest number of purchases.

QUESTION 5  [30 points]

(Please justify your answers. Answers without justification will not get any points!)

1. Consider the relation schema STUDENT_INFO(STUDENT#, NAME, BIRTHDAY, AGE, ADVISOR, DEPARTMENT, SEMESTER, COURSE, GRADE) with the key {STUDENT#, SEMESTER, COURSE} and the following functional dependencies:

   STUDENT# → NAME, BIRTHDAY, AGE, ADVISOR, DEPARTMENT
   BIRTHDAY → AGE
   ADVISOR → DEPARTMENT

(a) Show that STUDENT_INFO is not in third normal form.

(b) Give a decomposition of STUDENT_INFO into relations in third normal form such that the decomposition is lossless join and preserves dependencies.

(c) Are the relations you obtained as the result of the decomposition in (b) in Boyce-Codd normal form? Prove your answer.
2. Prove or disprove the following inference rules for functional dependencies where A, B, C, D, E, F, G are attributes of a relation schema. A proof can be made either by a proof argument or by using Armstrong’s inference rules including reflexive rule, augmentation rule and transitive rule. A disproof should be performed by demonstrating a relation instance that satisfies the conditions and functional dependencies in the left-hand side of the inference rule but does not satisfy the dependencies in the right-hand side.

   a. \{A \rightarrow BC, B \rightarrow DF, D \rightarrow E, F \rightarrow G\} \models \{A \rightarrow EFG\}
   
   b. \{AB \rightarrow C, A \rightarrow C\} \models \{B \rightarrow C\}
   
   c. \{A \rightarrow BCD, D \rightarrow BE, E \rightarrow F\} \models \{A \rightarrow BCDEF\}
   
   d. \{A \rightarrow BC\} \models \{B \rightarrow C\}
   
   e. \{ABC \rightarrow D\} \models \{A \rightarrow D\}
   
   f. \{A \rightarrow B\} \models \{B \rightarrow A\}

3. Consider the following set of functional dependencies on the schema (V, W, X, Y, Z):
   
   X \rightarrow YV
   
   Y \rightarrow Z
   
   X \rightarrow Y
   
   Z \rightarrow V
   
   XYW \rightarrow Z
   
   (i) Find a minimal equivalent set of functional dependencies (you need not use any special algorithm to find this set).
   (ii) Find a key for the schema. You can use the relevant algorithm. Otherwise, explain how you found the key and show that it is a key.

**QUESTION 6** [10 points]

(Please justify your answers. Answers without justification will not get any points!)

1. Explain how the two phase locking protocol works. Is this protocol deadlock-free? Either prove that the protocol can’t encounter deadlock or show by example that it can.

2. Suppose you are told that no two transactions in a set of transactions will deadlock if they are the only transactions running concurrently using two phase locking. For example, T1: w(X) w(Y) and T2: w(Y) w(Z) will never deadlock together. Call such a set **pairwise deadlock-free**. Consider a pairwise deadlock-free set having exactly six transactions. Can such a set encounter deadlock when it executes? Either prove that the set can’t encounter deadlock or show by example that it can.