Answer any three out of the following five questions

1. Answer the following three parts.

   a. Discuss the unique identification of entities in a database by means of primary key values versus their unique identification by means of object identity.

   [11 marks]

   b. Answer the following two parts.

      i. Recall that the manipulative part for a *nested relational database* includes the two operators $NEST$ and $UNNEST$. The operator $NEST$ transforms a nested relation into a "more deeply" nested relation while the operator $UNNEST$ transform a nested relation into a "flatter" nested relation. Explain with the aid of an example how these operators are used, given that you have available a flat relational database.

      [5 marks]

      ii. Let R be a relation schema R with schema(R) = $\{$A, B$\}$. Give an example of a nested relation r over R such that

      $$NEST_A(UNNEST_{(A)*}(r)) \neq r.$$

      [6 marks]

      [Total 11 marks]

   c. Write two to three sentences on each of the following:

      1. physical level of the database
      2. physical data independence
      3. conceptual level of the database
      4. conceptual data independence
      5. view level of the database

      [11 marks]

      [Total 33 marks]

2. Answer the following three parts.

a. Let r_1, r_2 and r_3 be three relations over relation schemas R_1, R_2 and R_3, respectively, with schema(R_1) = schema(R_2) = $\{$EN, DN$\}$ and schema(R_3) = $\{$DN, PN$\}$. (EN stands for Employee Name, DN stands for Department Name and PN stands for Project Name.)

Express the following queries both in the relational algebra and in SQL:

1. Output the tuples of employees that are either in r_1 or in r_2.
2. Output the tuples of employees that are in r_1 but not in r_2.
3. Output the names of employees in r_1 and projects they work on in r_3.
4. Output the names of employees in r_1 working on the "database" project in r_3.

[11 marks]

b. Give an incremental algorithm for testing whether a relation r over a relation schema R satisfies entity integrity, where X $\subseteq$ schema(R) is the primary key of R. Such an incremental algorithm assumes that prior to the insertion of a new tuple t into r, the relation r satisfies entity integrity.

[11 marks]

c. Give an efficient (i.e. polynomial time) algorithm for finding one key for a relation schema R with respect to a set of functional dependencies F over R.

[11 marks]

[Total 33 marks]

CONTINUED

3. Answer the following three parts.

   a. Suppose that we have a database schema consisting of two relation schemas EMP and DEPT such that schema(EMP) = {EN, DN} and schema(DEPT) = {DN, MN}. (EN stands for Employee Name, DN stands for Department Name and MN stands for Manager Name.) The primary key of EMP is EN and the primary key of DEPT is MN.

   Let r be a relation over EMP and s be a relation over DEPT. Discuss different policies of maintaining referential integrity given that the tuple <Jack, History> is inserted into r and the tuple <Computing ,Jill> is deleted from s.

   [11 marks]

   b. A view is **materialised** if it is physically stored in the database system. The **view maintenance problem** is the problem of appropriately updating a materialised view when an update is performed on the conceptual database.

   A materialised view is *self-maintainable* with respect to an update if it can be maintained without accessing the conceptual database relations.

   Suppose that a materialised view is defined as a projection of a single relation onto a subset of the attributes of its relation schema, say R.

   1. Show that such a views is self-maintainable with respect to insertions.
   2. Show that such a view is self-maintainable with respect to deletions from a relation r over R, if the primary key K for R is included in the view definition,

   [11 marks]

   c. Prove the assertion that a set of attributes X $\subseteq$ schema(R) is a superkey for a relation schema R if and only if for all relations r over R the number of tuples in π_X(r) is the same as the number of tuples in r.

   [11 marks]

[Total 33 marks]

4. Answer the following three parts.

a. Define the concept of a relation schema being in Boyce-Codd Normal Form (BCNF) with respect to a set of functional dependencies (FDs) over a relation schema R. Give an example of a relation schema which is in BCNF with respect to a set of FDs and another relation schema which is *not* in BCNF with respect to a set of FDs.

[11 marks]

b. In what sense is BCNF the definitive normal form for relational databases when the integrity constraints are specified in terms of a set of functional dependencies? Illustrate your answer with an example.

[11 marks]

c. i. What is the use of Armstrong's axiom system for functional dependencies?

[5 marks]

ii. Prove the soundness of the decomposition inference rule using the reflexivity, augmentation and transitivity inference rules of Armstrong's axiom system. Recall that the decomposition rule states that: if $F \vdash X \rightarrow YZ$ then $F \vdash X \rightarrow Y$ and $F \vdash X \rightarrow Z$, where F is a set of functional dependencies over R and X, Y, Z are subsets of schema(R).

[6 marks]

[Total 11 marks]

[Total 33 marks]

5. Answer the following three parts.

   a. Given a relation schema R, with schema(R) = $\{$A,B,C$\}$, together with a set of functional dependencies, F = $\{$AB $\rightarrow$ C, AC $\rightarrow$ B, BC $\rightarrow$ A$\}$. Is R in Second Normal Form (2NF)? Is R in Third Normal Form (3NF)? Is R in Boyce-Codd Normal Form (BCNF)? Justify your answers.

   [11 marks]

   b. A **simple key** is a key which which consists of only a single attribute.

   Give a proof of the assertion that a relation schema R that is in Third Normal Form (3NF) with respect to a set of functional dependencies F over R and such that all the keys for R with respect to F are simple is also in Boyce-Codd Normal Form (BCNF) with respect to F.

   [11 marks]

   c. Argue using an example that if a relation schema R is in Third Normal Form (3NF) with respect to a set of functional dependencies F but is not in Boyce-Codd normal Form (BCNF) with respect to F, then it must have at least two distinct keys with respect to F which overlap, i.e. such that their intersection is nonempty.

   [11 marks]

[Total 33 marks]

END OF PAPER