1. You are given two relation schemas, EMPLOYEE and DEPARTMENT. EMPLOYEE has the attributes ENUM (employee number), ENAME (employee name), DNUM (department number of the department in which the employee works in) and SALARY (employee salary in pounds). DEPARTMENT has the attributes DNUM (department number), DNAME (department name) and MGR (employee number of department manager).

   Answer the following three parts.

   a. Define the concepts of foreign key and referential integrity using the EMPLOYEE and DEPARTMENT schemas to illustrate your definitions.

   [11 marks]

   b. Define the concept of an ISA relationship used in an Entity-Relationship Diagram (ERD) using the EMPLOYEE and DEPARTMENT schemas to illustrate your definition.

   What is the connection between ISA relationships and referential integrity?

   [11 marks]

   c. An **inclusion dependency** (IND) is a statement of the form R[X] $\subseteq$ S[Y], where R and S are relation schemas and X and Y are equal length sequences of attributes in R and S, respectively. The IND R[X] $\subseteq$ S[Y] is satisfied in a database containing relations r over R and s over S, if for all tuples t in r, whenever the values of t are non-null then there exists a tuple u in s such that t[X] = u[Y].

   How can ISA relationships be defined in terms of inclusion dependencies ?

   How can the semantics of foreign keys be expressed in terms of inclusion dependencies ?

   [11 marks]

   [Total 33 marks]

2. Answer the following three parts.

   a. Define the concepts of primary key and entity integrity using the EMPLOYEE and DEPARTMENT relation schemas of Question 1 to illustrate your definitions.

   [11 marks]

   b. Give an incremental algorithm for testing entity integrity of a relation r over a schema R, where X $\subseteq$ schema(R) is the primary key of R.

   (Hint: Such an incremental algorithm assumes that prior to the insertion of a new tuple t into r, the relation r satisfies entity integrity. It then checks that entity integrity is maintained after t is inserted into r only with respect to t and the relation r prior to inserting t into it.)

   [11 marks]

   c. Prove (or give a convincing argument for) 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 cardinality of π_X(r) is the same as that of r.

   [11 marks]

[Total 33 marks]

3. A view is a relation which is defined by a relational algebra query (or equivalently an SQL query) with respect to the conceptual database. For example, we can create a view containing employee names and the names of their managers.

   Answer the following three parts.

   a. A view is **materialised** if it physically stored in the database system, or **virtual** if it is computed each time the user poses a query over it.

      What are the advantages of materialised views versus virtual views? Use the EMPLOYEE and DEPARTMENT relation schemas of Question 1 to illustrate the points.

      [11 marks]

   b. The **view update problem** is the problem of appropriately updating the conceptual database when an update is performed on a view.

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

      Argue for restricting such views so that the projected attribute list contains a key for the relation schema R. Use the EMPLOYEE relation schema of Question 1 to illustrate the points.

      [11 marks]

   c. Given a fixed relation, say r, over a relation schema, R, whose attributes are PARENT and CHILD give a simple algorithm which computes the transitive closure of r.

      You are allowed to use the following relational algebra operators in your algorithm: projection, natural join, union and renaming.

      [11 marks]

   [Total 33 marks]

4. Answer the following three parts.

   a. An **Armstrong** relation over a relation schema R for a set of FDs F over R is a relation which satisfies exactly those FDs in F (i.e. it satisfies all the FDs in F and only those in F).

      Argue that given a set of FDs F over R an Armstrong relation over R always exists for F.

      Give an Armstrong relation for the set of FDs containing the single FD, ENAME $\rightarrow$ $\{$DNAME, SALARY$\}$, over a relation schema EMP having the three attributes ENAME, DNAME and SALARY. (Hint: Four tuples suffice.)

      [11 marks]

   b. Define the concept of a database schema being in Boyce-Codd Normal Form (BCNF) with respect to a set of FDs using the EMPLOYEE and DEPARTMENT schemas of Question 1 to illustrate your definition.

      [11 marks]

   c. Why is BCNF a good relational database design objective ?

      Argue your case using example relations over a relation schema EMP having the three attributes ENAME, PROJECT (the projects an employee works on) and LOCATION (unique location of a project).

      [11 marks]

[Total 33 marks]

5. Answer the following three parts.

   a. Give a brief comparison of the Hierarchical and the Nested Relational data models.

   [11 marks]

   b. Explain the NEST and UNNEST operators of the nested relational algebra using an example relation over a relation schema EMP having the three attributes ENAME, PROJECT (the projects an employee works on) and LOCATION (unique location of a project).

   [11 marks]

   c. Suggest why the nested relational data model may be a good internal level between the physical and conceptual levels.

   Also, suggest why the nested relational model may useful in restructuring user views. Illustrate your answer using the relation schema EMP of part (b) of this question.

   [11 marks]

[Total 33 marks]

END OF PAPER