Time: Three Hours   **Max. Marks: 70**

1. Consider a database system for a baseball organization such as the major leagues. The data requirements are summarized as follows: **(8)**

   o The personnel involved in the league include players, coaches, managers, and umpires. Each is identified by a unique personnel id. They are also described by their first and last names along with the date and place of birth.

   o Players are further described by other attributes such as their batting orientation (left, right, or switch) and have a lifetime batting average (BA).

   o Within the players group is a subset of players called pitchers. Pitchers have a lifetime ERA (earned run average) associated with them.

   o Teams are uniquely identified by their names. Teams are also described by the city in which they are located and the division and league in which they play (such as Central division of the American League).

   o Teams have one manager, a number of coaches, and a number of players.

   o Games are played between two teams with one designated as the home team and the other the visiting team on a particular date. The score (runs, hits, and errors) are recorded for each team. The team with the most runs is declared the winner of the game.

   o With each finished game, a winning pitcher and a losing pitcher are recorded. In case there is a save awarded, the save pitcher is also recorded.

   o With each finished game, the number of hits (singles, doubles, triples, and home runs) obtained by each player is also recorded.

   Design an Enhanced Entity-Relationship diagram for the BASEBALL database.

2. Map the given EER diagram into a relational schema; give reasons for your decisions.   **(8)**
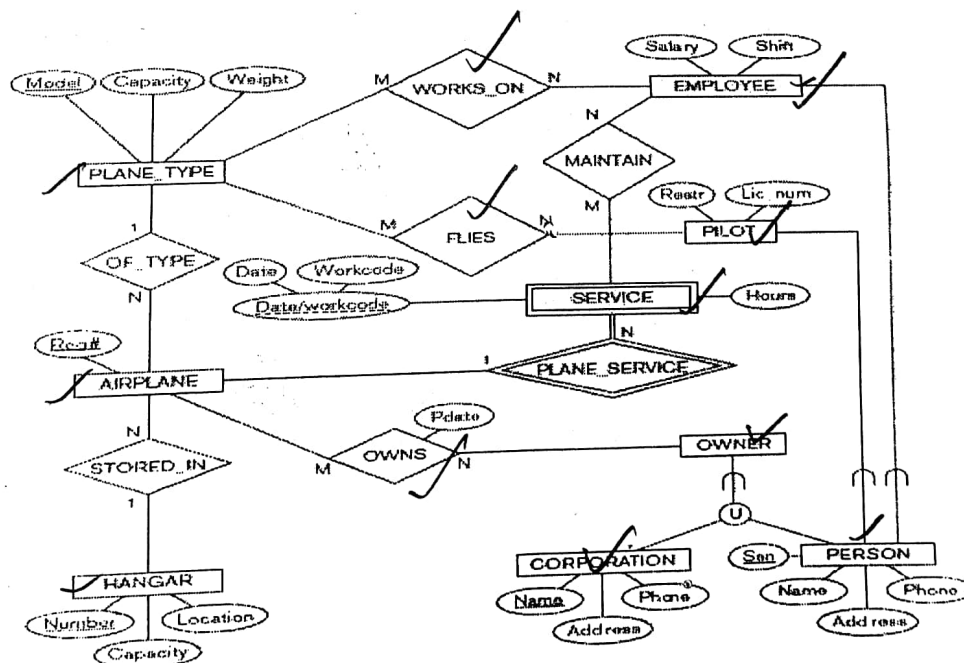
3. Consider the following relation for book club:

Members (**Member_Id**,Name,Designation,Age)

Books (**Book_Id**,BookAuthor,BookTitle,BookPublisher,BookPrice)

Reserves (**Member_Id,Book_Id**,Date)

Write relational algebra queries for the following: (3*3=9)

   i. List the names of members who are professors older than 50 years.

   ii. Find the authors and titles of book reserved on '20-9-2012'.

   iii. Find the names of members who have reserved all books.

4. Consider a database that consists of the following relations.

SUPPLIER(**Sno**, Sname)

PART(**Pno**, Pname)

PROJECT(**Jno**, Jname)

SUPPLY(**Sno, Pno, Jno**)

Write SQL queries for the following: (3*3=9)

   i. Retrieve the part numbers that are supplied to exactly two projects.

   ii. Retrieve the part numbers that are supplied by every supplier.

   iii. Retrieve the names of suppliers who supply at least two different parts each to at least two different projects.

5.

a. Consider a relation Drinkers (**name**, **addr**, phones, **beersLiked**), FD: {name ->addr} and MVD's: {name ->->phones, name ->->beersLiked}. Is it in 4NF? If not decompose it into 4NF. (3)

b. Consider a relation R (A,B,C,D,E,F) and the set of functional dependencies on R,

   F = {ABD->AC, C->BE, AD->BF, B->E}

   Find minimum cover for this set of FDs. (5)

c. Consider a relation R (N,C,HR,E,HN,HL) and set of functional dependencies on R, F= {NC->HR, N->E, C->HNHL,HN->HL} and answer the following questions. (1+5=6)

   I. In what normal form is R?

   II. Apply normalization until you cannot decompose the relations further. State the reasons behind each decomposition.

d. Consider the following decomposition for the relation schema R = {A, B, C, D, E, F, G, H, I, J} and the set of functional dependencies F = { {A, B}→{C}, {A}→{D, E}, {B}→{F}, {F}→{G,H}, {D}→{I, J} }. Determine whether each decomposition has (1) the dependency preservation property, and (2) the lossless join property, with respect to F.

   D = {R1, R2, R3, R4, R5}; R1 = {A, B, C, D}, R2 = {D, E}, R3 = {B, F}, R4 = {F, G, H}, R5 = {D, I, J}. (2+3=5)

6.

a. Consider the three transactions T1, T2, and T3, and the schedule S given below. Draw the serializability (precedence) graphs for S and state whether schedule is serializable or not. If a schedule is serializable, write down the equivalent serial schedule(s). (3)

   T1: r1 (X); w1 (X); r1(Y); w1(Y);

   T2: r2 (Z); r2 (Y); w2 (Y); r2(X); w2 (X);

   T3: r3 (Y); r3 (Z); w3 (Y); w3 (Z)

   S: r3 (Y); r3 (Z); r1 (X); w1(X); w3 (Y); w3 (Z); r2 (Z); r1 (Y); w1 (Y); r2 (Y); w2(Y); r2(X); w2(X);

b. Which of the following schedules is (conflict) serializable and why? (2)

   i. r3(X); r2(X); w3(X); r1(X); w1(X);

   ii. r3(X); r2(X); r1(X); w3(X); w1(X);

7. Answer the following questions: (3*4=12)

a. What is the system log used for? What is transaction commit point, and why it is important?

b. Discuss the entity integrity and referential integrity constraints. Why each is considered important?

c. Compare binary locks to exclusive/shared locks. Why the latter type locks preferable?

d. Describe the three schema architecture. Why do we need mappings between schema levels?