CSE4701 Introduction to Databases – Midterm Exam

Name: ____________________________

<table>
<thead>
<tr>
<th>Problem</th>
<th>Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td></td>
</tr>
</tbody>
</table>

Please show all work to receive ANY credit!!!!

Roughly equate 15 points to fifteen (15) minutes of time/effort.

Exam is Open Book, Open Notes, Open Course Web Page.

Write on only one side of the paper.
1. (15 points) Relational Algebra

This question requires you to write relational algebra queries. Note that you are allowed to use variables (relations) to hold intermediate results if your answer is in multiple steps.

(a) (3 points) List the full names of all mens players that have/had the uniform number 5.

(b) (4 points) List the full names and points per game of all mens players that averaged more than 20 points per game and at least 5 rebounds per game in one year on teams with winning regular season records.
(c) (4 points) List the full names and uniform numbers of all womens players that played for Geno Auriemma in 1996.

(d) (4 points) List the full names of all mens players who played on a team with a losing regular season record coached by Jim Calhoun.
2. (15 points) Reverse Engineering: Relational to ER Diagram

One of the tasks often associated with database maintenance is documenting the structure of a relational schema that has been in use for a long period of time (prior to EER and UML design techniques). In such a situation, an EER diagram can be reverse engineered from a set of relational tables. Consider the modified BBALL schema below:

PLAYER(PLName, PFName, StartYear, NumYears, UniformNumber, PersonID);
COACH(CLName, CFName, StartYear, EndYear, PersonID);
TEAM(TeamID, Year, Squad);
ROSTERS(TeamID, PlayerPersonID, CoachPersonID);
RSRECORD(TeamID, Wins, Losses);
PORECORD(TeamID, Wins, Losses);
STATISTICS(PlayerPersonID, TeamID, PPG, RPG, APG);
TITLES(TeamID, TitleType);

Note that a PersonID has been added to players and coaches, and is unique across the two relations with PersonID replacing the name in all of the tables. This changes the assumptions as follows: Players and Coaches can have the same name; and, a player one year can be a coach in a later year (but not in the same year) - that player would have the same PersonID if he or she was later a coach.

Reverse engineer this schema into a EER diagram. You must have one example of disjoint inheritance and one example of overlapping inheritance in your solution!
Problems 3 and 4 are based on a revised version of the BBALL Schema with all keys removed. PLAYER and COACH have been combined using boolean flags (PFlag, CFlag) to distinguish unique attributes; TEAM, ROSTERS, and TITLES have been combined into a single relation; and, RSRECORD and PSRECORD have been combined into a RECORD relation, with the attributes changed to represent wins and losses for the regular season (RSWins and RSLosses) and the entire season (TTLWins and TTLLosses), requiring playoff wins and losses to now be calculated. Other than this change, all other assumptions of the original BBALL Schema still apply.

PLAYERCOACH(LName, FName, StartYear, PFlag, NumYears, UniformNumber, CFlag, EndYear);
TEAMROSTERSTITLES(TeamID, Year, Squad, PLName, CLName, TitleType);
RECORD(TeamID, RSWins, RSLosses, TTLWins, TTLLosses);
STATISTICS(PLName, TeamID, PPG, RPG, APG);

3. (15 points) Relational Table Analysis

The revised version of the BBALL relational schema was designed in an ad-hoc manner, without the process of EER design, EER to relational conversion, and normalization. As a result, there are many aspects of the schema that cause it to be poorly designed. For this question, you are to analyze the schema and identify design problems. To do so, please consider the issues raised regarding “What is a Good DB Schema?” (Slide 15&16-6 and the subsequent four Guidelines (Slides 15&16-11, 15&16-14, 15&16-26/27, and 15&16-39) that are focused on, respectively: Represent a Single Entity, Redundant Information and Anomalies, If anomalies do not exist in a relation, explain why (1 or 2 sentences). If anomalies do exist in a relation, identify which type(s) exist, namely, insertion, deletion, or modification. For each type that does exist, describe an example that causes the anomaly. Focus on anomalies within a single relation rather than across multiple relations. (insert, update, delete), Null Values, and Spurious Tuples.

Make sure that as part of your answer, you review every relational table of the schema. If the table is not poorly designed, justify why not; if the table is poorly designed, be specific as to the problem that exists and describe an example that causes the problem.

PLAYERCOACH
TEAM

RECORD

STATISTICS
4. (15 points) Functional Dependencies

(a) (8 points) Define all functional dependencies (FDs) for the revised BBALL schema. List your results on a relation-by-relation basis. Be very specific - do not simply specify that a single attribute determines all others. Make sure that you use arrow notation for FDs, i.e., \(X \rightarrow Y\), \(SSN \rightarrow EmpName\), \(EmpAddr\).

PLAYERCOACH

TEAMROSTERSTITLES

RECORD

STATISTICS
(b) (7 points) Multi-valued dependencies occur when one attribute can determine multiple values of another attribute. For example, in the relation

EMPLOYEE(EmpName, ProjName, DependentName)

there are multi-valued dependencies EmpName → ProjName (an Employee works on multiple projects) and EmpName → DependentName (an Employee has multiple dependents). Identify all multi-valued dependencies for the revised BBALL schema.

PLAYERCOACH

TEAMROSTERSTITLES

RECORD

STATISTICS
5. (15 points) Normalization

Consider the NBAPLAYER relation in First Normal Form given below with key indicated:

\[
\text{NBAPLAYER}(\text{Name, Year, Coach, Team, State, Salary})
\]

with the functional dependencies:

A. \{Name, Year, Coach\} \rightarrow \{Team, State, Salary\}
B. \{Name, Year\} \rightarrow \text{Team}
C. Team \rightarrow \text{State}
D. Coach \rightarrow \text{Salary}

and sample data shown:

<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Coach</th>
<th>Team</th>
<th>State</th>
<th>Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones</td>
<td>2011-2012</td>
<td>Reagan</td>
<td>Miami</td>
<td>Florida</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Jones</td>
<td>2011-2012</td>
<td>Clinton</td>
<td>Miami</td>
<td>Florida</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Jones</td>
<td>2011-2012</td>
<td>Nixon</td>
<td>Miami</td>
<td>Florida</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Bird</td>
<td>2010-2011</td>
<td>Clinton</td>
<td>Boston</td>
<td>Mass</td>
<td>800,000</td>
</tr>
<tr>
<td>Malone</td>
<td>2008-2009</td>
<td>Clinton</td>
<td>Miami</td>
<td>Florida</td>
<td>700,000</td>
</tr>
</tbody>
</table>

Apply normalization to convert the NBAPLAYER relation into second normal form, Clearly indicate both the keys and the functional dependencies in each relation after normalization.

Is the resulting table in 3NF? If so, then argue why the table is in 3NF. If not, then convert your 2NF solution into 3NF. Again, clearly indicate both the keys and the functional dependencies in each relation after normalization.
Extra Page for Problem 5 if Needed