

CS3431 Homework 2 (Fall 2012 - B term)

Homework Out: Monday, Nov 5th, 2012 (in class)

Homework Due: Monday, Nov 12th, 2012 (in class and via mywpi).

Homework Submit: In hardcopy at start of class Monday!

Homework Points: A maximum of 50 pts.

Purpose: Getting started with Oracle and Relational Database Design. But don't give us in hardcopy the actual long output of the SQL query runs! You can deposit those into mysql, if you wish.

Problem 1: To Get You Started With Oracle. [10 pts]

You need to do the following: Create a table GameScores with attributes playerNumber of type integer, and score of type integer.

```
CREATE TABLE GameScores (  
  playerNumber INTEGER,  
  score INTEGER);
```

We provide a script to randomly generate data to be loaded into your GameScores database. For that, save the script available at:

<http://www.cs.wpi.edu/~cs3431/b12/hw/>

in your local directory. Execute the above script as **./createData.pl**. Make sure that this script is executable, otherwise change permissions using the command **chmod 755 createData.pl**. The execution of this script will create sample data in a file named **sample.dat**. Now load this data into your table. You do not want to be entering all those tuples by hand. The number of tuples will be huge. Thus you are motivated to give the data bulk loader a try (described in the lecture notes).

From SQLPLUS, if we execute the command **SET TIMING ON**, for every SQL statement, you will get the time it took for the database server to execute that statement.

For each of the following SQL statements, run them, and then report the time it took for executing that statement. Don't report the actual output tuples, as there will be many too many of them - wasting paper and trees and trees. If you wish to submit this, you can submit the runs themselves via mywpi.

However, you still need to provide the analysis of these runs in hardcopy in class. Namely, indicate the time of how long each query ran. Then observe which query ran the quickest, which the slowest, and so on. Lastly, indicate if you have any thoughts about why this difference in running time may be the case.

- **SELECT COUNT(playerNumber) FROM GameScores;**

- UPDATE GameScores SET score=score + 100;
- SELECT playerName, MIN(score), MAX (score) FROM GameScores GROUP BY playerName;
- SELECT COUNT(T1.playerNumber) FROM GameScores T1, GameScores T2 WHERE T1.playerNumber = T2.playerNumber;

Problem 2: Map ER Design to Relational Design [40 pts]

Consider the ER design that we have designed for our target sports application in homework 1. You need to utilize the schema that we have provided to you as solution, even if your own model solution is different. This will allow us to keep this homework 2 consistent among all students.

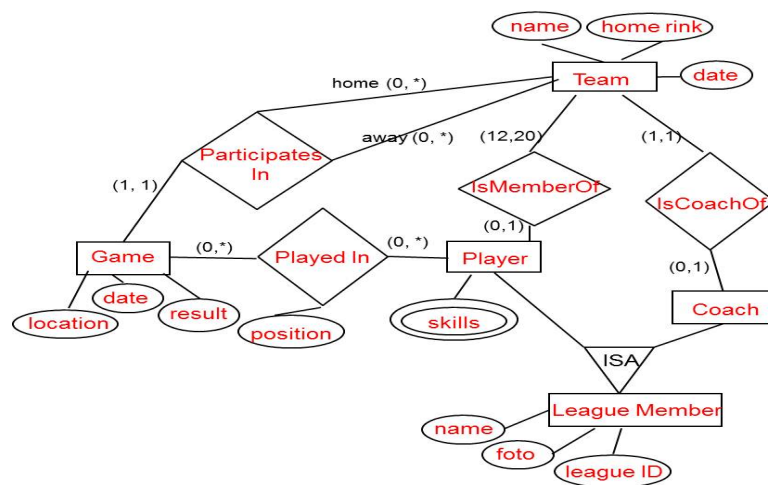


Figure 1: My Favorite Sport.

1. Please translate the ER schema into SQL DDL statements, reducing the number of relations whenever appropriate.
2. State any additional assumptions that you have made about this mapping. In particular, state what assumptions you make about the IS-A relationship being covering or non-covering, being overlapping or non-overlapping.
3. Provide a brief discussion justifying your chosen mapping strategy for the different cases of mapping that you encounter, such as for the reduction of the number of relations, if any, for your handling of IS-A relationships, and so on.
4. Provide a listing of any information, including constraints and cardinalities, that could not be represented in the relational model (based on your knowledge of the relational model as covered in CS3431 thus far).

5. Enter your DDL statements into ORACLE to verify that they are correct. Develop some data content for your database application, and load it into your schema. Show us the log (using spool off/on) of the above script of you working with ORACLE DBMS, i.e., of creating your schema and loading data into it.