CS561: Advanced Topics In Database Systems
Spring-2012

Final Exam

Total Points: 170
Release Date: 04/19/2012 @4:00pm
Due Date: 04/24/2012 @4:00pm

Submission: Submit a single zip (or tar) file containing your answers on Blackboard system by the due date.

Question 1 [50 Points]

Given the following database schema:

Department(<u>ID</u>, name, address) --- short name D
Faculty(<u>ID</u>, deptID, firstName, lastName, joinYear) --------------- short name F
Course(<u>ID</u>, deptID, name, numOfCredits) ----------------------------- short name C
Student(<u>ID</u>, deptID, firstName, LastName, joinYear) -------------- short name S
Registration(<u>studentID</u>, <u>courseID</u>, <u>semester</u>, facultyID, grade)   -- short name R

The primary key in each relation is underlined. Each faculty member and each course belongs to exactly one department. Each student also belongs to one department. However, a student can take courses outside his/her department.

1) **[10 Points]** Write an SQL query that reports for each semester and department the number of course offered. The output schema should look like:

              Semester   DepartmentId   CountCourses
              ----

2) **[10 Points]** For each department, report the year in which the largest number of student joined that department. Note that the "Student" relation has the student's department Id and the join year. The output schema should look like:

              DepartmentId   YearWithMaxEnrollment   EnrollmentCount
              ----

3) **[10 Points]** Give an object-relation design of the schema above. Create the appropriate types and tables and give column the suitable data types.

4) **[10 Points]** Use the SQL-99 syntax (given in class) to query the object-relational model you designed in 3) and report the first name and last name of faculty members who joined in year 2010.

5) **[10 Points]** Write database trigger(s) to enforce the following integrity constraint: The number of students registered in a certain course in a certain semester cannot exceed 50.

Question 2 [50 Points]

Given the attached XML document (recipes.xml), provide the <u>XQuery</u> of the following queries.

1) **[10 Points]** Report the title and nutrition of all recipes.

2) **[10 Points]** For each ingredient, report the recipes (the title only) that uses this ingredient

3) **[10 Points]** Report the title of recipes that have complex ingredients

4) **[10 Points]** For each recipe, report the recipe title, the number of ingredients used, and the number of steps needed in preparation. If there recipe contains 'ComplexIngredient", then you count what's inside them. For example, the output of the first two recipes should look like:

```
<recipe>
        <title>Beef Parmesan with Garlic Angel Hair Pasta</title>
         <numOfIngredient>  11 </numOfIngredient>
         <numOfSteps>  4 </numOfSteps>
</recipe>

<recipe>
        <title>Ricotta Pie</title>
         <numOfIngredient>  11 </numOfIngredient>
         <numOfSteps>  6 </numOfSteps>
</recipe>
```

5) **[10 Points]** Report the recipe with the lowest "fat" number (attribute inside nutrition) and the recipe with the highest "protein" number. Report only the recipe title and the nutrition element. The output should look like:

```
<recipe>
        <title>….</title>
         <nutrition … />
         <Type>  Has the highest protein number </Type>
</recipe>
<recipe>
        <title>….</title>
         <nutrition … />
         <Type>  Has the lowest fat number </Type>
</recipe>
```

Question 3 [40 Points]

Given the following set of transactions, each transaction consists of a set of items. Answer the following questions.

TID	List of Items
T1	I2, I3
T2	I2, I4
T3	I1, I2, I5
T4	I1, I2, I3
T5	I1, I3
T6	I1, I2 ,I3, I5
T7	I2, I3
T8	I1, I3
T9	I3, I5
T10	I1, I2, I4

1) **[20 Points]** Using the Apriori algorithm given in class, show the different scans of the algorithm and in each scan show the candidate itemsets and the frequent itemsets. Assume the minimum support for a frequent itemset = 2 (20%), i.e., above or equal to 20% is frequent.

2) **[10 Points]** Compute the support and confidence of the following association rules

   a. I2 → I3
   b. I3, I5 → I1

3) **[10 Points]** Give two association rules where their confidence is 100%. Report the support of each of these rules as well.

Question 4 [30 Points]

1) **[10 Points]** In distributed databases, the 2 Phase Commit (2PC) algorithms is used to ensure that atomicity of a distributed transactions (that is the transaction is done entirely or not done at all). Describe what will happen in the following case:

   a. Assume that in Phase 1 of the algorithm, all clients sent back to the coordinator a "ready T" message except one client sent "do not commit T" and then this client failed. When this client recovers from the failure, what should it do? Does it need to contact the coordinator again or not?

   b. Assume in Phase 1, the coordinator sends a "prepare T" message and then crashes. What will happen after it recovers from the crash?

2) **[10 Points]** Assume a customer table that has attributes (ID, Name, Age, CountryCode, Salary) where the records are hash-based partitioned over a distributed system based on the ID column. Now given the following query:

> Select *
> From Customer
> Where CountryCode betweem (x, y);

   Describe how this query will execute in parallel? Will we need to re-partition the data or not?

3) **[10 Points]** Assume a customer table that has attributes (ID, Name, Age, CountryCode, Salary) where the records are hash-based partitioned over a distributed system based on the Name column. Now given the following query:

> Select Age, Name, sum(Salary)
> From Customer
> Group By Age, Name;

   Describe how this query will execute in parallel? Will we need to re-partition the data or not?

Bonus Question (Optional) [30 Points]

This question is a continuation to the Hadoop project that you delivered before (Refer to the project released on 02/16/2012). Using the same datasets described in the project, write a MapReduce code corresponding to the following queries.

1) Report the transactions whose total number of items above 5 and their total cost above 400

2) For every customer, report the customer Id, the total number of transactions done by this customer, and the sum of "TransTotal" of these transactions.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<?dsd href="recipes.dsd"?>
<collection xmlns="http://recipes.org"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://recipes.org recipes.xsd">
  <description>
        Some recipes.
  </description>

  <recipe>
        <title>Beef Parmesan with Garlic Angel Hair Pasta</title>
        <ingredient name="beef cube steak" amount="1.5" unit="pound"/>
        <ingredient name="onion, sliced into thin rings" amount="1"/>
        <ingredient name="green bell pepper, sliced in rings" amount="1"/>
        <ingredient name="Italian seasoned bread crumbs" amount="1" unit="cup"/>
        <ingredient name="grated Parmesan cheese" amount="0.5" unit="cup"/>
        <ingredient name="olive oil" amount="2" unit="tablespoon"/>
        <ingredient name="spaghetti sauce" amount="1" unit="jar"/>
        <ingredient name="shredded mozzarella cheese" amount="0.5" unit="cup"/>
        <ingredient name="angel hair pasta" amount="12" unit="ounce"/>
        <ingredient name="minced garlic" amount="2" unit="teaspoon"/>
        <ingredient name="butter" amount="0.25" unit="cup"/>
        <preparation>
                <step>
                        Preheat oven to 350 degrees F (175 degrees C).
                </step>
                <step>
                        Cut cube steak into serving size pieces. Coat meat with the bread crumbs
                        and parmesan cheese. Heat olive oil in a large frying pan, and saute 1
                        teaspoon of the garlic for 3 minutes. Quick fry (brown quickly on both sides)
                        meat. Place meat in a casserole baking dish, slightly overlapping edges.
                        Place onion rings and peppers on top of meat, and pour marinara sauce
                        over all.
                </step>
                <step>
                        Bake at 350 degrees F (175 degrees C) for 30 to 45 minutes, depending on
                        the thickness of the meat. Sprinkle mozzarella over meat and leave in the
                        oven till bubbly.
                </step>
                <step>
                        Boil pasta al dente. Drain, and toss in butter and 1 teaspoon garlic. For a
                        stronger garlic taste, season with garlic powder. Top with grated parmesan
                        and parsley for color. Serve meat and sauce atop a mound of pasta!
                </step>
        </preparation>
        <comment>
                Make the meat ahead of time, and refrigerate over night, the acid in the
                tomato sauce will tenderize the meat even more. If you do this, save the
                mozzarella till the last minute.
        </comment>
```

```xml
        <nutrition calories="1167" fat="23" carbohydrates="45" protein="32"/>
</recipe>

<recipe>
    <title>Ricotta Pie</title>
    <ComplexIngredient name="filling">
        <ingredient name="ricotta cheese" amount="3" unit="pound"/>
        <ingredient name="eggs" amount="12"/>
        <ingredient name="white sugar" amount="2" unit="cup"/>
        <ingredient name="vanilla extract" amount="2" unit="teaspoon"/>
        <ingredient name="semisweet chocolate chips" amount="0.25" unit="cup"/>
        <preparation>
            <step>
                Beat the 12 eggs, 2 cups sugar and vanilla extract together. Stir in
                the ricotta cheese and the chocolate chips. Set aside.
            </step>
        </preparation>
    </ComplexIngredient>
    <ComplexIngredient name="dough">
        <ingredient name="flour" amount="4" unit="cup"/>
        <ingredient name="baking powder" amount="5" unit="teaspoon"/>
        <ingredient name="white sugar" amount="1" unit="cup"/>
        <ingredient name="shortening" amount="0.5" unit="cup"/>
        <ingredient name="eggs, lightly beaten" amount="4"/>
        <ingredient name="vanilla extract" amount="1" unit="teaspoon"/>
        <preparation>
            <step>
                Combine the flour, baking powder, and 1 cup of the sugar together. Cut in the
                shortening and mix until the mixture resembles coarse crumbs. Mix in 4 of
                the eggs and 1 teaspoon of the vanilla. Divide dough into 4 balls and chill (if
                needed).
            </step>
        </preparation>
    </ComplexIngredient>
    <ingredient name="milk" amount="*"/>
    <preparation>
        <step>
            Preheat oven to 325 degrees F (165 degrees C). Grease two deep dish pieplates.
        </step>
        <step>
            Roll out 2 of the balls to fit into the pie pans. Do not make the crust too thick
            as it will expand during cooking and get too thick. Do not flute the edges of
            the dough. Roll out the other 2 balls of dough and cut each into 8 narrow
            strips for the top of the crust. Alternately you can use cookie cutters and
            place the cutouts on the top of the pies.
        </step>
        <step>
            Pour the filling evenly into the pie crusts. Top each pie with 8 narrow strips
            of dough or cookie cut-outs. Brush top of pie with milk for shine. Place foil on
            the edge of crust.
        </step>
        <step>
            Bake at 325 degrees F (165 degrees C) for 20 to 30 minutes then remove
```

```xml
                foil. Continue to bake for another 25 or 30 minutes or until a knife inserted in
                the center comes out clean.
            </step>
        </preparation>
        <nutrition calories="349" fat="18" carbohydrates="64" protein="18"/>
</recipe>

<recipe>
        <title>Linguine Pescadoro</title>
        <ingredient name="linguini pasta" amount="16" unit="ounce"/>
        <ComplexIngredient name="sauce">
            <ingredient name="olive oil" amount="4" unit="tablespoon"/>
            <ingredient name="minced cloves of garlic" amount="2"/>
            <ingredient name="Italian seasoning" amount="0.5" unit="teaspoon"/>
            <ingredient name="dried thyme" amount="0.25" unit="teaspoon"/>
            <ingredient name="crushed red pepper flakes" amount="0.25" unit="teaspoon"/>
            <ingredient name="crushed tomatoes" amount="1" unit="can"/>
            <ingredient name="black olives, drained" amount="6" unit="ounce"/>
            <ingredient name="whole baby clams" amount="10" unit="ounce"/>
            <ingredient name="minced clams, with juice" amount="6.5" unit="ounce"/>
            <ingredient name="small salad shrimp" amount="0.25" unit="pound"/>
            <ingredient name="scallops" amount="0.25" unit="pound"/>
            <ingredient name="lemon zest" amount="2.5" unit="teaspoon"/>
            <ingredient name="salt" amount="*"/>
            <ingredient name="ground black pepper" amount="*"/>
            <preparation>
                <step>
                    In a heavy saucepan over medium heat saute garlic in olive oil
                    until garlic softens.
                </step>
                <step>
                    Add Italian seasoning, thyme, crushed red pepper flakes,
                    crushed tomatoes, black olives, and the juice from both cans of clams.
                    Simmer for 15 minutes.
                </step>
                <step>
                    Mix in canned clams, shrimp, scallops, lemon zest, and salt and pepper to taste.
                </step>
                <step>
                    Simmer for an additional 15 minutes or until shrimp and clams are cooked.
                </step>
            </preparation>
        </ComplexIngredient>
        <preparation>
            <step>
                In a large pot of boiling salted water cook linguini until al dente. Drain.
            </step>
            <step>
                Toss cooked and drained linguine pasta over seafood sauce. Serve warm.
            </step>
        </preparation>
        <nutrition calories="532" fat="12" carbohydrates="59" protein="29"/>
</recipe>
```

```xml
<recipe>
    <title>Zuppa Inglese</title>
    <ingredient name="egg yolks" amount="4"/>
    <ingredient name="milk" amount="2.5" unit="cup"/>
    <ingredient name="olive oil" amount="2" unit="tablespoon"/>
    <ingredient name="Savoiardi biscuits" amount="21"/>
    <ingredient name="sugar" amount="0.75" unit="cup"/>
    <ingredient name="Alchermes liquor" amount="1" unit="cup"/>
    <ingredient name="lemon zest" amount="*"/>
    <ingredient name="flour" amount="0.5" unit="cup"/>
    <ingredient name="fresh whipping cream" amount="*"/>
    <preparation>
        <step>
            Warm up the milk in a nonstick sauce pan
        </step>
        <step>
            In a large bowl beat the egg yolks with the sugar, add the flour
            and combine the ingredients until well mixed.
        </step>
        <step>
            Add the milk, a little bit at the time to the egg mixture, mixing well.
        </step>
        <step>
            Put the mixture into the sauce pan and cook it on the stove at a medium
            low heat. Mix the cream continuously with a wooden spoon. When it starts
            to thicken remove it from the heat and pour it on a large plate to cool off.
        </step>
        <step>
            Stir the cream now and then so that the top doesnt harden.
        </step>
        <step>
            Dip quickly both sides of the lady fingers in the liquor. Layer them one
            at the time in a glass bowl large enough to contain 7 biscuits.
        </step>
        <step>
            Spread 1/3 of the cream and repeat the layer with lady fingers. Finish with the cream.
        </step>
    </preparation>
    <comment>
        Refrigerate for at least 4 hours better yet overnight. Before serving decorate
        the zuppa inglese with whipped cream.
    </comment>
    <nutrition calories="612" fat="49" carbohydrates="45" protein="4" alcohol="2"/>
</recipe>
</collection>
```