

# Technologies for Information Systems

## Part I (10 points)

prof. L. Tanca – February 10th, 2016

Available time: 25 minutes

<b>Last Name</b> _____
<b>First Name</b> _____
<b>Student ID</b> _____ <b>Signature</b> _____

- 1) Describe concisely the concept of mashup and its utility in data integration, highlighting its distinctive features w.r.t. using other integration techniques.
- 2) Briefly define pervasive data management and the main problems that must be solved.

# Technologies for Information Systems

## Part II (22 points)

prof. L. Tanca – February 10th, 2016

Available time: 1h 50m

Last Name _____	
First Name _____	
Student ID _____	Signature _____

*PoliCo* is a big company composed of many branches in various countries, and whose employees are often on travel for business reasons. Each business trip is performed by an employee and is associated with a certain expenditure item. The trip must be authorized by the employee who is responsible for that expenditure item.

Since business trips represent a very relevant expense for *PoliCo*, the management of the company asked you to design a data warehouse to analyze them.

The following is the schema of the *PoliCo* operational database:

BUSINESSTRIP (SSNEmployee, StartingDate, DurationInDays, FoodCost, LodgingCost, TransportationCost, DestinationCity, ExpenditureItemCode)

EMPLOYEE (SSN, Surname, GivenName, HomeCity, Role, BranchName) // *Role may assume the values "Clerk", "Manager" or "Technician".*

CITY (CityName, Country)

EXPENDITUREITEM (ExpenditureItemCode, Name, Supervisor, YearlyAvailableBudget)  
// *YearlyAvailableBudget represents the total expense which can be associated with the expenditure item in a year.*

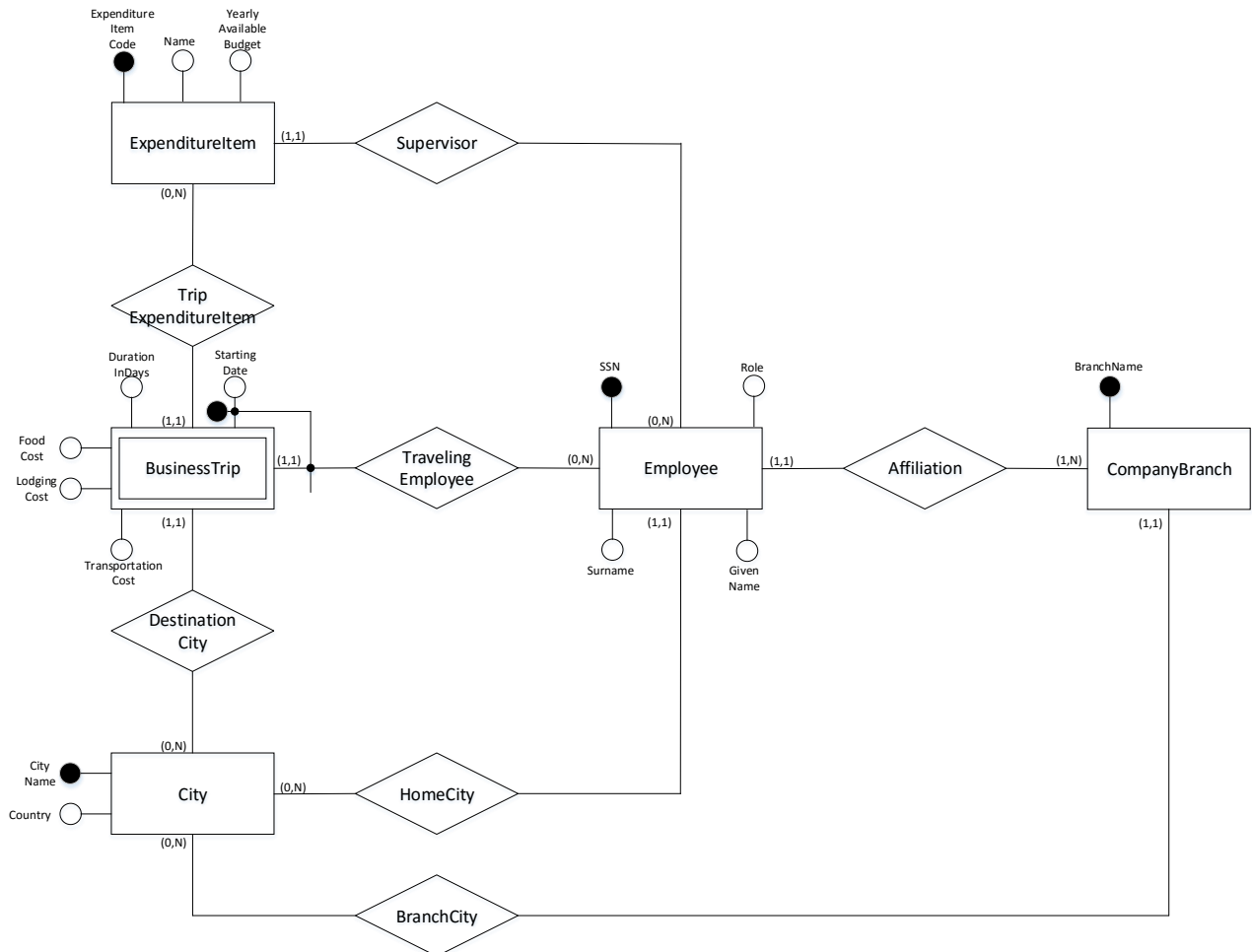
COMPANYBRANCH (BranchName, BranchCity)

1. (3 points) Perform the reverse engineering of the given logical schema into a conceptual schema (Entity-Relationship model).
2. With respect to the produced ER diagram, discover the fact(s) that are useful specifically for answering the queries reported below. For each of these facts:
  - a. (3 points) Produce the attribute tree (with pruning and grafting).
  - b. (3 points) Produce the conceptual schema (fact schema).
  - c. (2 points) Identify the measure(s) and produce the glossary.

3. (3 points) Produce a logical schema consistent with the conceptual schema.
4. Write in SQL the following queries against the designed logical schema:
  - a. (1.5 points) Find the **average duration in days** of the business trips performed by **technicians** grouped **by departure month** and **city** of the **company branch** to which the employee belongs.
  - b. (1.5 points) Considering only the trips related to the **Italian branches** of *PoliCo*, find the total expense for each starting date, company branch and home country of the traveling employee. Include in the answer also the aggregations computed using only one or two of the three attributes.
  - c. (2.5 points) Select SSN, given name and surname of the supervisor(s) having authorized the greatest number of trips with departure in the second quarter of 2015 and performed by employees residing in Milan.
  - d. (2.5 points) Select for each expenditure item the Italian city/cities which has/have been the destination of the greatest number of trips with departure in 2014.

# SOLUTION

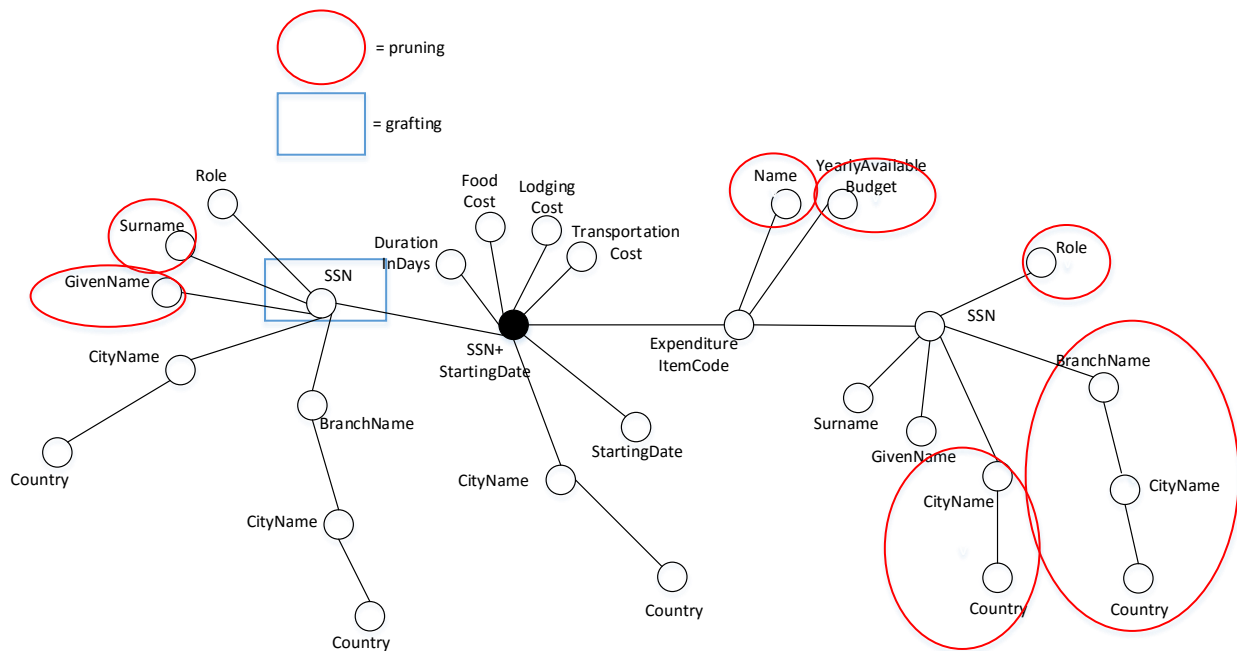
## 1. Reverse engineering



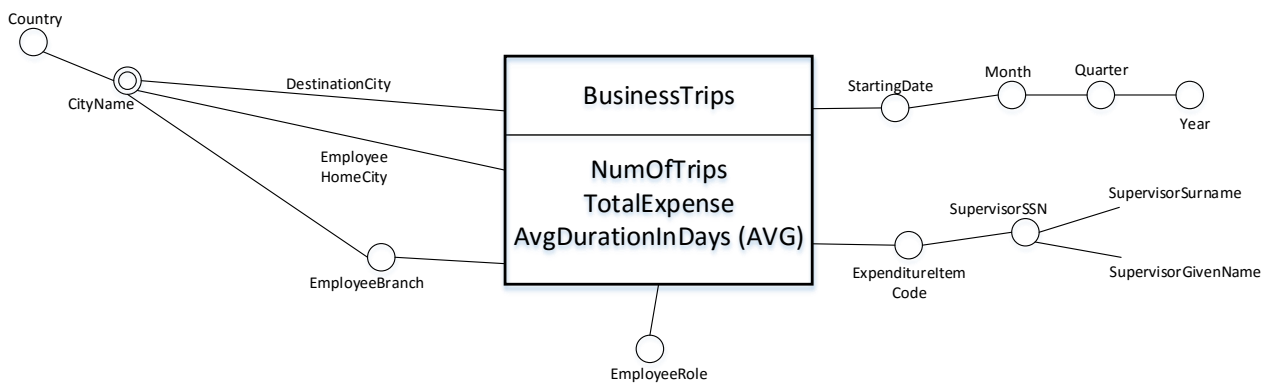
## 2. Conceptual design

Fact: BusinessTrips (BusinessTrip entity)

### 2a) Attribute tree



## 2b) Fact schema



## 2c) Glossary

NumOfTrips

```
SELECT B.StartingDate, B.DestinationCity, B.ExpenditureItemCode, E.HomeCity, E.Role, E.BranchName,
       COUNT(*)
FROM BusinessTrip AS B, Employee AS E
WHERE B.SSNEmployee=E.SSN
GROUP BY B.StartingDate, B.DestinationCity, B.ExpenditureItemCode, E.HomeCity, E.Role, E.BranchName
```

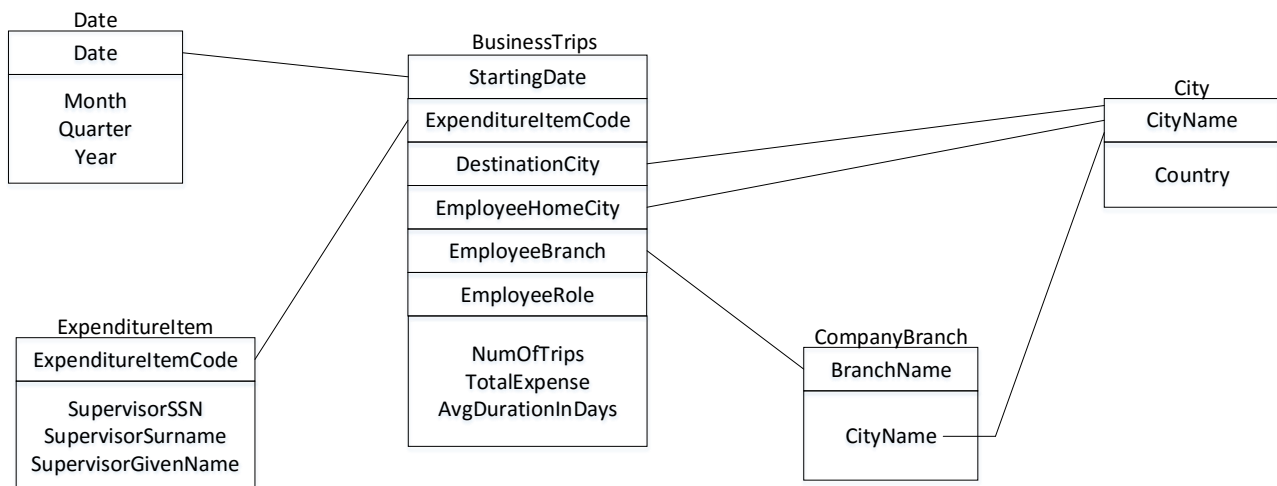
#### TotalExpense

```
SELECT B.StartingDate, B.DestinationCity, B.ExpenditureItemCode, E.HomeCity, E.Role, E.BranchName,
       SUM(B.FoodCost+B.LodgingCost+B.TransportationCost)
FROM BusinessTrip AS B, Employee AS E
WHERE B.SSNEmployee=E.SSN
GROUP BY B.StartingDate, B.DestinationCity, B.ExpenditureItemCode, E.HomeCity, E.Role, E.BranchName
```

#### AvgDurationInDays

```
SELECT B.StartingDate, B.DestinationCity, B.ExpenditureItemCode, E.HomeCity, E.Role, E.BranchName,
       AVG(B.DurationInDays)
FROM BusinessTrip AS B, Employee AS E
WHERE B.SSNEmployee=E.SSN
GROUP BY B.StartingDate, B.DestinationCity, B.ExpenditureItemCode, E.HomeCity, E.Role, E.BranchName
```

### 3. Logical design



### 4. Query answering

4a) Find the average duration in days of the business trips performed by technicians grouped by departure month and city of the company branch to which the employee belongs.

```
SELECT D.Month, C.CityName, SUM(B.NumOfTrips*B.AvgDurationInDays)/SUM(B.NumOfTrips)
FROM BusinessTrips AS B, Date AS D, CompanyBranch AS C
WHERE B.StartingDateDate=D.Date AND B.EmployeeBranch=C.BranchName AND
      B.EmployeeRole='Technician'
GROUP BY D.Month, C.CityName
```

**4b) Considering only the trips related to the Italian branches of *PoliCo*, find the total expense for each starting date, company branch and home country of the traveling employee. Include in the answer also the aggregations computed using only one or two of the three attributes.**

```
SELECT B.StartingDate, B.EmployeeBranch, C-E.Country, SUM(B.TotalExpense)
FROM BusinessTrips AS B, City AS C-E, CompanyBranch AS C, City AS C-B
WHERE B.EmployeeHomeCity=C-E.CityName AND B.EmployeeBranch=C.BranchName AND
      C.CityName=C-B.CityName AND C-B.Country='Italy'
GROUP BY B.StartingDate, B.EmployeeBranch, C-E.Country WITH CUBE
```

**4c) Select SSN, given name and surname of the supervisor(s) having authorized the greatest number of trips with departure in the second quarter of 2015 and performed by employees residing in Milan.**

```
CREATE VIEW AuthorizedTrips (SupervisorSSN, SupervisorGivenName, SupervisorSurname, NumOfTrips) AS
(
    SELECT E.SuperiorSSN, E.SuperiorGivenName, E.SuperiorSurname, SUM(B.NumOfTrips)
    FROM BusinessTrips AS B, ExpenditureItem AS E, Date AS D
    WHERE B.ExpenditureItemCode=E.ExpenditureItemCode AND B.StartingDate=D.Date AND
          D.Quarter='2-2015' AND B.EmployeeHomeCity='Milan'
    GROUP BY E.SuperiorSSN, E.SuperiorGivenName, E.SuperiorSurname
)

SELECT SupervisorSSN, SupervisorGivenName, SupervisorSurname
FROM AuthorizedTrips
WHERE NumOfTrips = (
    SELECT MAX(NumOfTrips)
    FROM AuthorizedTrips
)
```

**4d) Select for each expenditure item the Italian city/cities which has/have been the destination of the greatest number of trips with departure in 2014.**

```
CREATE VIEW ExplItemCityTrips (ExpenditureItemCode, CityName, NumOfTrips) AS (
    SELECT B.ExpenditureItemCode, B.DestinationCity, SUM(B.NumOfTrips)
    FROM BusinessTrips AS B, City AS C, Date AS D
    WHERE B.DestinationCity=C.CityName AND B.StartingDate=D.Date AND C.Country='Italy' AND
          D.Year='2014'
    GROUP BY B.ExpenditureItemCode, B.DestinationCity
)

SELECT E.ExpenditureItemCode, E.CityName
FROM ExplItemCityTrips AS E
WHERE E.NumOfTrips = (
    SELECT MAX(E2.NumOfTrips)
    FROM ExplItemCityTrips AS E2
    WHERE E2.ExpenditureItemCode=E.ExpenditureItemCode )
```