Université libre de Bruxelles
Faculté des Sciences Appliquées

September 2, 2014

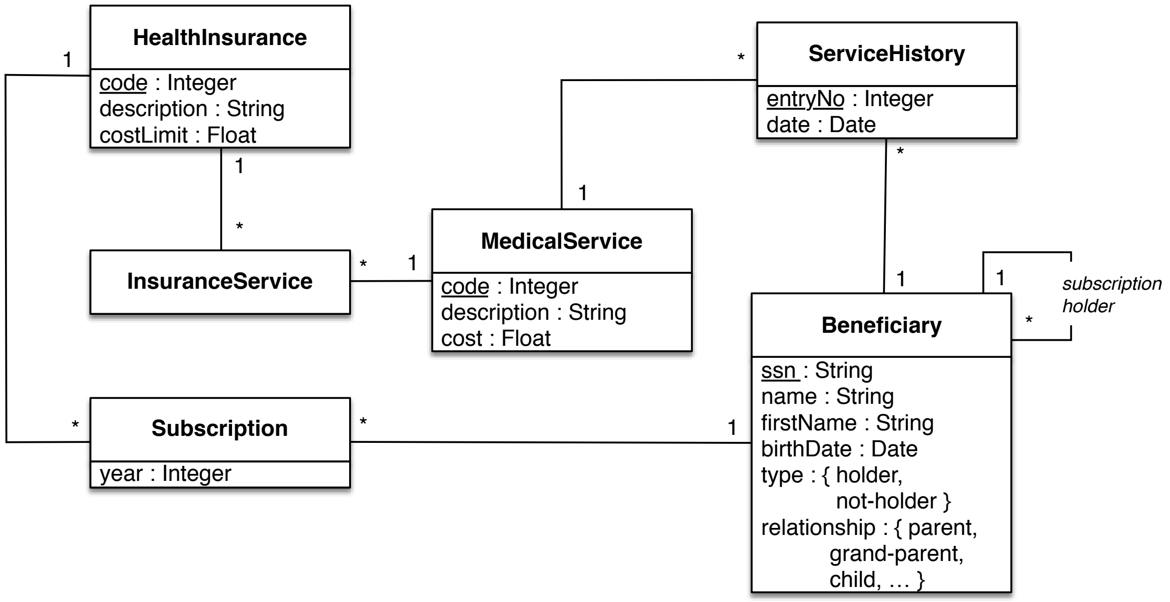
INFO-H-415 – Advanced databases

Second session examination

The exam is divided in four sections. All sub-questions are worth approximately the same amount of points. However, some of these will only take you a minute, some require a bit more thinking, and a couple would require weeks to answer perfectly. Make the best use of your time.

1 Active Databases (5 pt)

Consider the following partial diagram of a health insurance company:

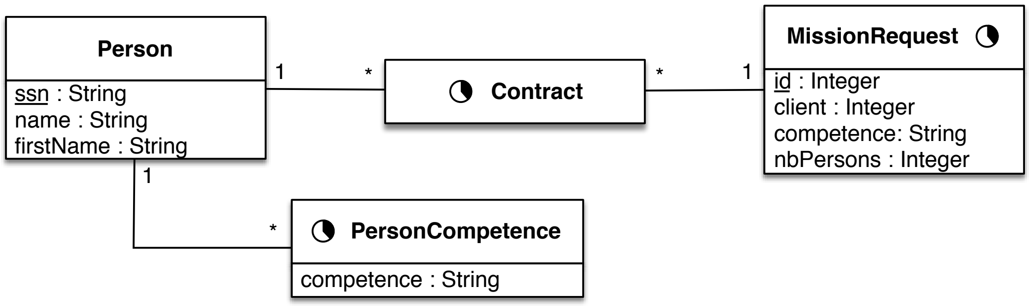


Questions

1. Define a **relational schema** corresponding to the conceptual schema above.
2. Write the SQL Server code ensuring the following **integrity constraints**. Whenever multiple triggers are needed to enforce a single integrity constraint, list all of them, but write the code in full for only one of them. Throughout the entire question you should provide at least one example for each of ON INSERT, ON UPDATE, and ON DELETE triggers.
   (a) Beneficiaries who have their `type` set to "holder" must have the `relationship` and the reference to the holder set to `null`.
   (b) For beneficiaries who are "not-holder", the reference to the holder must be set to another beneficiary who is "holder" and the `relationship` may not be empty.
   (c) To be able to add a record to the **`ServiceHistory`** at a certain `date`, the following constraints must be met:
      - The beneficiary must have subscribed an insurance that covers the service and the `date` must be in the corresponding subscription's year.
      - The total cost of all services already taken by the beneficiary over the year of the `date` for a given subscription may not exceed the insurance's `costLimit`.

2 Temporal Databases (5 pt)

Consider the following partial conceptual diagram of a temporary employment agency.



Reminder: The "empty clocks" denotes things that occur at one point in time, while the "full clocks" denotes things that occur over an interval of time.

Questions

1. Define a **relational schema** corresponding to the conceptual schema above.
2. Write the SQL Server code ensuring the following **integrity constraints**:

   (a) To be able to sign a contract, a person must have previously acquired the required competence.

   (b) The lifecycle of a contract must lie within the one of the referencing mission request.

3. Write the following **queries** in SQL:

   (a) Give the name of the people that are currently under contract and that have no contract signed thereafter.

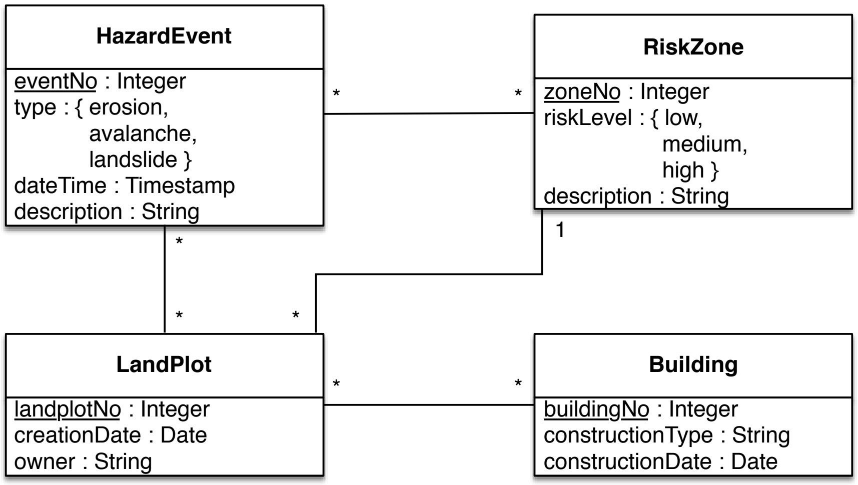
   (b) Give the contract history of all persons, independently from the agency's client.

   (c) For each person, give the number of days that he/she worked on a contract in 2013.

   (d) Give the name of the person who has worked on the greatest number of contracts simultaneously.

3 Object Databases (5 pt)

A risk assessment application keeps track of hazard events such as erosion, landslides and avalanches. For this a set of risk zones are defined that are related to land plots and their buildings. The schema of such application is as follows.



Here is part of the associated types in Oracle:

```
CREATE TYPE THazardEvent;
CREATE TYPE TRiskZone;
CREATE TYPE TLandPlot;
CREATE TYPE TBuilding;

CREATE TYPE TSetRefRiskZones AS TABLE OF REF TRiskZone;
CREATE TYPE TSetRefLandPlots AS TABLE OF REF TLandPlot;

CREATE OR REPLACE THazardEvent AS OBJECT(
    eventNo INTEGER,
    type ENUM('erosion', 'avalanche', 'landslide'),
    dateTime TIMESTAMP,
    description VARCHAR2(100),
    riskzones TSetRefRiskZones,
    landplots TSetRefLandPlots
);
```

Questions

1. Write the **type definitions** for `TRiskZone`, `TLandPlot`, and `TBuilding`.

2. Write the following **Oracle queries**:

   (a) The list of landplots in a high risk zone where more than 25 hazards have happened.

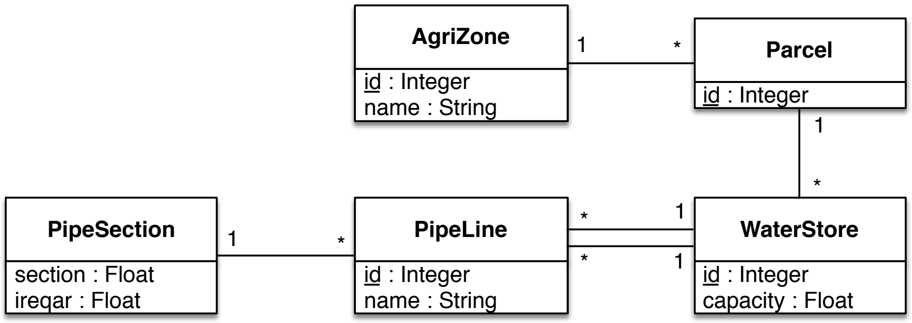
   (b) Give the hazard events that happened in 2013 ordered by event type with the number of land plots and buildings they affected. Notice, as indicated by the schema, that a building may be located in more than one land plot.

   (c) For each land plot owner give the total number of land plots they have in each risk level

4 Spatial Databases (5 pt)

Consider a large agricultural *zone* in a country that suffers from high climate variability. The government has planned to build a range of small-scale *water stores* that could help farmers overcome dry spells that would otherwise cause their crops to fail. A very simple *network of pipelines* (also to be built) is aimed at reliably feeding the parcels with water: each pipeline links two water stores and traverses one ore more parcels. (Note that the pipelines must not be straight)

The following preliminary conceptual diagram has already been elaborated



Note that each pipeline has a given section that quantifies the volume of water it can transport per hour. Based on a geological study, an equivalent of area (expressed in m^2) that can be irrigated with each pipe section has been determined (ireqar).

Questions

1. Based on the above diagram, design a spatial database that would keep track of the parcels, the water stores, and the pipeline network. Practically, you are asked to provide the set of SQL statements that is required to generate the corresponding PostGIS database.

2. Beyond the aim of managing the water network in the future, the database also should serve to validate its initial design. Assuming that the database has been populated with data, you are further asked to provide the SQL statements that give the following information:

   (a) give the parcels that are not traversed by any pipeline;

   (b) identify the pipelines that are feeding more than five parcels with water;

   (c) identify the pipelines that are longer than 5km and that come closer to a third water store than 500m.

   (d) list the parcels that are not sufficiently irrigated. (Hint: use the sum of the ireqar field of all pipelines that traverse a parcel.)

You might need some of the following PostGIS functions:

- `AddGeometryColumn(table_name, column_name, srid, type, dimension)`, where type can be any of POINT, LINE, MULTILINE, MULTIPOLYGON, etc.

- `ST_Envelope(geom)` returns a geometry representing the bounding box of the supplied geometry: ((MINX, MINY), (MINX, MAXY), (MAXX, MAXY), (MAXX, MINY), (MINX, MINY)).

- `ST_Intersects(geom1, geom2)` returns true if the Geometries/Geography spatially intersect in 2D.

- `ST_Intersections(geom1, geom2)` returns a geometry that represents the shared portion of geom1 and geom2.

- `ST_DWithin(geom1, geom2, dist)` returns true if the geometries are within the specified distance of one another.

- `ST_Area(geom)` returns the area of the geometry in the units of the srid.

- `ST_Length(geom)` returns the 2D length of the geometry if it is a linestring or multilinestring.