Building Geospatial Ontologies from Geographic Database Schemas in Peer Data Management Systems

Danúbia Lima (biaeccentric@gmail.com)
Antonio Mendonça (tony2415@gmail.com)
Ana Carolina Salgado (acs@cin.ufpe.br)
Damires Souza (damires@ifpb.edu.br)
Considered the result of blending the benefits of **P2P networks** with the richer semantics of **databases**.

- They can be used for data exchanging, query answering and information sharing.
Data Management and Ontologies in PDMS

- Data Management is a challenging problem given the **heterogeneity** of the schemas.

- Ontologies have been used as **uniform metadata representation**.

- Due to semantic heterogeneity, ontologies may be also used as a way of providing a **domain reference**.
  
  ✓ A domain ontology may be used as a semantic reference or background knowledge to enhance processes such as ontology matching and query answering.
SPEED System

**SPEED** (Semantic **PEEr-to-Peer** Data Management System) is a PDMS which uses semantics in order to improve its processes.

Legend:
- $I_1$ – Integration Peer
- $I_1D_1$ – Data Peer
- Community Ontology - CMO
- Local Ontology - LO
- Cluster Ontology - CLO
Our goal

- \textbf{GeoMap}: an \textbf{approach} and an \textbf{implemented tool} for automatically building a geospatial peer ontology as a semantic view of data stored in a geographic database

- A set of correspondences between the generated ontology components and the original database schema is automatically generated

- The produced peer ontology will be later used for matching and querying processes in the PDMS
Outline

- Geospatial Data
- GeoMap Approach
- GeoMap Tool
- Conclusions and Future Work
Geospatial Data

- Represented using the **vector model**: points, lines or polygons
- **Heterogeneity** of the sources is even greater: data may have multiple representations, different resolutions and coordinate systems and associated temporal properties
The GeoMap Approach

GeoMap

- GeoSpatial Reference Ontology
- Construct Identifier
- Generator
- Classifier
- Extractor
- Geographic Database
- Peer Ontology
- Correspondences Set
Geospatial Reference Ontology
The GeoMap Tool

- Implemented in JAVA, using the Protégé-OWL API and the Jena framework for ontology manipulation
- It uses geographic databases coded in Oracle DBMS and PostGIS
Experiments

- **Completeness** of the peer ontology
  - Degree to which entities and properties of the peer data source (i.e. the database schema) are not missing in the generated peer ontology

- We have invited some users to produce a manual peer ontology from the geographic database schemas
  - These “gold ontologies” were compared with our produced peer ontologies result

- Our produced peer ontologies are complete in terms of the existing database elements
Conclusions and Future Work

- **GeoMap** accomplishes the extraction of metadata from a geographic database representing them in terms of a peer ontology.

  - It identifies the equivalence **correspondences** between the generated ontology components and the existing database schema entities and properties.

  - A **geospatial reference ontology** is used as a way to provide the semantics of geospatial relationships and types, absent from the set of existing concepts in the OWL model.
The **GeoMap** tool is able to produce the peer ontology in an automatic way, by using the semantics provided by the reference ontology.

- It can use any background knowledge that may support the needed geospatial semantics.
  - For instance, it will be able to use the OGC GeoSparql standard (when it is ready for use).

- It does **not need** the user intervention.
  - This tool will be stated as a service in SPEED system which will be dynamically executed at peer arriving time.
Conclusions and Future Work

- Currently, this version generates ontologies from Oracle databases
- We are right now extending it to also extract metadata from PostGIS
- An important future work concerns identifying correspondences between geospatial peer ontologies
  - We will be able to reformulate and execute geospatial queries among the existing peers in the PDMS
Building Geospatial Ontologies from Geographic Database Schemas in Peer Data Management Systems

Danúbia Lima (biaeccentric@gmail.com)
Antonio Mendonça (tony2415@gmail.com)
Ana Carolina Salgado (acs@cin.ufpe.br)
Damires Souza (damires@ifpb.edu.br)