SAIKS Assignment SS 2020

The assignment consists of performing a set of advanced technical steps that are typical when creating ontology-based applications and documenting the key decisions taken during this process.

Main Steps:

Step 1: Advanced Ontology Engineering
In this step, you will practice advanced ontology engineering techniques. There are two possibilities for choosing the ontology that you will work on.

- Reuse the assignment from the *Introduction to Semantic System (ISS)* course: in this case, start with the ontology that you have developed as part of the ISS assignment. In your report, sum up briefly the application that the ontology aims to support as well as the competency questions that should be covered by the application/ontology.
- Build ontology from scratch: First, you will need to decide on a problem domain that you will consider in your assignment. The choice of the problem domain is free and can be anything from health care (hospital), car manufacturing, family tree to air traffic management. In your report, describe the chosen problem domain and how a hypothetical application in that problem domain could benefit from semantic technologies. What are the main features of the advanced semantic application you envision? What competency questions should be covered by the application/ontology?

In both cases, extend or create the ontology so that it contains at least 30 classes and 25 properties. During ontology engineering, make use of advanced modelling features offered by OWL2 (refer to lecture 2). Additionally, illustrate the use of at least 3 different Ontology Design Patterns (ODPs). These can be selected from those described during the lecture or from the ODP website (ontologydesignpatterns.org). We recommend using Protégé, but any other ontology editor can be used if preferred. In your report, document in detail the OWL2 features and the ODPs that you used.

Step 2: KG creation through OBDA
Install the Protégé Ontop plugin to access a relational database through the ontology created at STEP1. As database, you can:

- use one of the four data sources that were investigated as part of the ISS assignment and create a simple relational database table(s) based on it; otherwise
- procure a relevant relational database, depending on your application domain. For example, you could find relevant data in public (open) data repositories (e.g., kaggle, wikidata, data.gov.at). Alternatively, you could extract relevant data through web-scraping or generate representative synthetic data. The size of the data is not of major importance, but it should contain around hundred relevant entities for your domain.

In a next step, create R2RML mappings (or Ontop/Quest Mapping) to link your ontology to the database. Follow the Ontop Tutorial[1] for detailed information. Apply the mappings, materialize the resulting ontology-based data into a knowledge graph and store it into the GraphDB graph-database (details will be given during the lecture).

Step 3: Verification of Semantic Structures

Ontology verification. Check the quality of the ontology created at STEP1 with the OOPS! ontology pitfall scanner (http://oops.linkeddata.es/) and improve it accordingly. Describe the issues found by the tool and how you addressed them. Note that the number of issues found by OOPS! will have neither negative nor positive influence on your grade.

KG verification. Write 5-10 SHACL rules with varying complexities for validating the knowledge graph resulting from STEP2 and run these following the instructions in the TopBraid SHACL engine[2] or by using SHACL4Protege plugin[3] from Protégé to get validation results.

Step 4: Ontology Alignment

Use an ontology alignment tool for aligning your ontology with another ontology from the same domain. For the alignment result file we recommend the Alignment API format (http://alignapi.gforge.inria.fr/). Shortly summarise the strengths and weaknesses of your alignment result and submit this together with your result file by email. Ontology alignment will be explained in practice during the lecture.

Submission of practical assignment:

- OWL Ontology (STEP1)
- Ontop exercise: database dump, exported R2RML mappings, knowledge graph in GraphDB Repository (STEP2)
- SHACL definitions (.ttl) and SHACL reports (.ttl) (STEP3)
- Alignment result file (STEP4)
- Report including the description of your work for all 4 steps of the assignment and the rationale for key decisions you took.

(Expected) Important Dates:

- **27.03.2020:** formation of groups and selection ISS assignment to extend or topic of the ontology in case the ontology will be built from scratch.
- **24.04.2020:** mid-term presentation.
  5 minutes (2-3 slides) presentation + 5 minutes questions per group. Focus on status of your assignment, next steps, problems you are facing.
- **19.06.2020:** end-term presentation. 10 minutes (10-12 slides) per group and 10 minutes discussion.
- **26.06.2020:** submission of technical outputs and project report in TUWEL.

[1] https://github.com/ontop/ontop/wiki/Tutorials
[2] https://github.com/TopQuadrant/shacl; an example implementation for Java is available at https://github.com/TopQuadrant/shacl/blob/master/src/test/java/org/topbraid/shacl/ValidationExample.java
[3] https://github.com/fekaputra/shacl-plugin