Translating competency questions to simple SPARQL-OWL

Dawid Wiśniewski*, Jedrzej Potoniec*, Agnieszka Ławrynowicz*

*Faculty of Computing, Poznan University of Technology
ul. Piotrowo 3, 60-965 Poznan, Poland
{dwisniewski, jpotoniec, alawrynowicz}@cs.put.poznan.pl

Abstract

Competency questions can be defined as a set of questions that evaluate the quality of a given ontology. They represent functional requirements stated in natural language that the complete ontology should be able to answer to. In order to verify whether a knowledge base has an answer for a competency question a manual translation from natural language to a query language must be performed. The SPARQL-OWL query language is an extension of SPARQL that is able to utilize OWL Direct Semantics entailment regime. One of the main reasons to use SPARQL-OWL rather than simple SPARQL is the fact that the SPARQL-OWL enables the users to take advantage of general, terminological knowledge of a given ontology while looking for an answer. Moreover, using SPARQL-OWL one can utilize the open world assumption. We propose a method that creates SPARQL-OWL queries from a set of competency questions given an ontology. The method, after preprocessing phase consisting of tokenization, pos-tagging and dependency tree analysis, extracts entities and relations from the question and groups them, so the entities being arguments for relations are stored together with the relation for further steps. Next, for each relation and its arguments we score possible translations from the ontology using three coefficients utilizing different aspects of ontology information: word-embedding based phrase similarity, property domain/range restrictions and presence in ontology axioms. The three coefficients are aggregated to a confidence score for each possible translation candidate. The candidate with the highest confidence score is held for further step. Then, the SPARQL-OWL query is constructed if the system is confident enough about the translation. If the confidence score is too low, the user is informed about inability to create a query. To evaluate our method, we have created a benchmark from publicly available competency questions. For every question in the set, a SPARQL-OWL query was proposed. Then, the benchmark was limited to contain only simple SPARQL-OWL queries (with single someValuesFrom or hasValue expressions). We decided to start from simple cases, because most of the complex cases consist of some composition of simple expressions with someValuesFrom or hasValue – solving simple cases first should be a good base for further work. From the prepared benchmark – 7 out of 13 cases were translated correctly. The proposed method shows that it is possible to construct simple SPARQL-OWL queries from some competency questions, indicating it is worth to explore further extensions that may lead to higher completion of OWL axioms coverage.