Natural language queries for enhanced knowledge access

Danica Damljanović, Valentin Tablan
University of Sheffield

Objectives

• Developing a tool for quering the knowledge store using Natural Language (NL) queries; this knowledge store includes ontologies and the knowledge base created based on them.

• Users have the freedom to enter queries of any length and form.

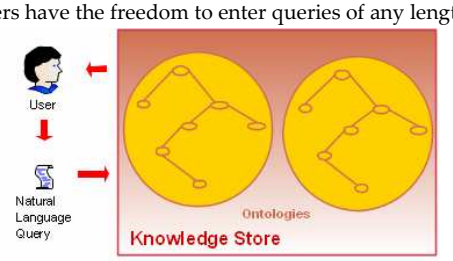

Figure 1: Querying the knowledge store using human language

Motivation

• Existing query languages (e.g., SeRQL) are complex.

• Writing queries:

  • requires using the exact syntax that is not easy to learn,

  • is error-prone task and

  • requires understanding of Semantic Web technologies.

Inspiration

CLOnE Query Language (QL) was inspired by CLOnE (Controlled Language for Ontology Editing): CLOnE provides users to edit ontologies using natural language (Tablan et al., 2006). Similarly, CLOnE QL enables users to query the existing knowledge by transforming input NL queries into SeRQL queries using GATE resources for natural language processing.

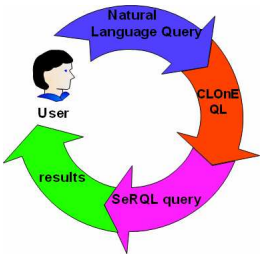

Figure 2: The process of querying the knowledge store using NL

Our Approach

CLOnE QL analyses the input query using GATE pipeline of Processing Resources which are executed in the sequence as shown on Figure 3. The output of CLOnE QL analyser is a set of SeRQL queries. This queries are than executed to return the results to the user.

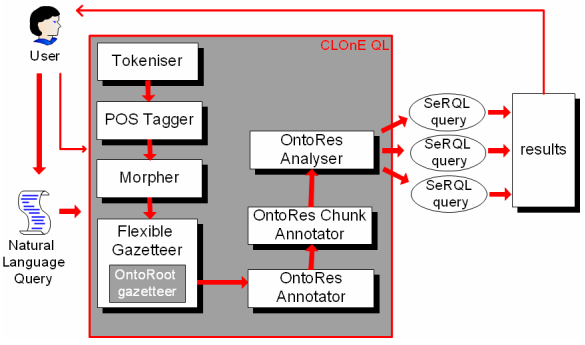

Figure 3: CLOnE Query Language Application

Generating SeRQL queries from NL

Figure 4 illustrates an example query shown in GATE GUI. It depicts key ontology-aware annotations used in the process of SeRQL query construction.

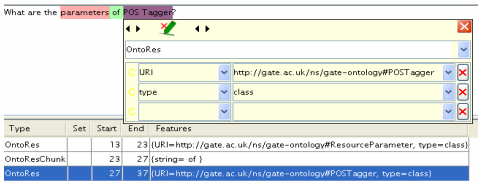

Figure 4: Example query in GATE GUI

Figure 5 illustrates the process of analysing the query from figure 4. As word 'parameters' was annotated as a class with name ResourceParameter, and 'POS Tagger' was recognized as a class with name "POSTagger", relevant **relations** between these two resources are found inside the ontology. Their name is compared with the text between these two resources (in this example: 'of') and the SeRQL queries are constructed. The priority is given to the SeRQL query that includes property which name is the most similar to the text between found resources.

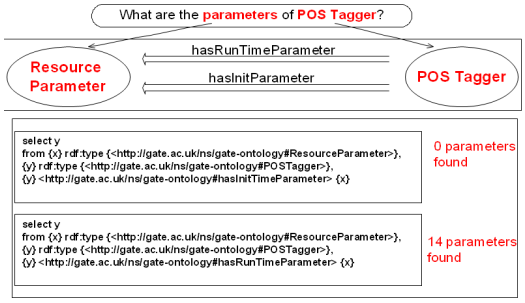

Figure 5: From natural language to SeRQL query

An Example

Figure 6 depicts the screenshot from the Web-service based application we implemented to run the CLOnE QL.

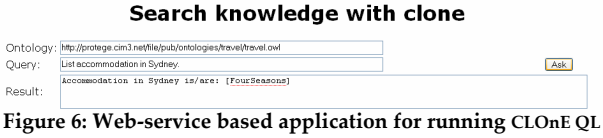

Figure 6: Web-service based application for running CLOnE QL

Related work

•KIM (Popov et al., 2003) provides semantic search interface that hides complex query syntax, but requires that the user is familiar with the underlying ontology.

•AquaLog (Lopez, Pasin, Motta, 2005) uses a controlled language for querying the ontology: this portable system is coupled with a learning mechanism but heavily relies on techniques from natural language processing; it requires syntactically correct sentences and cannot process concept-based queries such as 'accommodation Rome'.

•SemSearch (Lei et al., 2006) system accepts concept-based queries such as 'news:PhD Students' but doesn't consider properties, and there could be more than one relation between two types of concepts; it doesn't accept natural language questions such as 'how many PhD Students are registered at the University of Sheffield?'.

Acknowledgements

The research for this poster was conducted as part of the European Union Sixth Framework Program projects TAO (FP6-026460).

References

Lei, Y., Uren, S., V.,Motta, E. (2006). SemSearch: A Search Engine for the Semantic Web. In Staab, S., Studer, R. (Eds.): EKAW 2006. Springer-Verlag Berlin Heidelberg 2006, pp. 238-245.

Lopez, V.,Pasin, M., Motta, E.(2005). AquaLog: An Ontology-Portable Question Answering System for the Semantic Web. ESWC 2005: 546-562

Popov, B., Kiryakov, A., Kirilov, A., Manov, D., Ognyanoff, D. & Goranov, M.: KIM – Semantic Annotation Platform. In the Proceedings of the 2nd International Semantic Web Conference (ISWC2003), Florida, USA, October 2003, pp. 834-849.

Tablan, V., Polajnar, T., Cunningham, H., Bontcheva, K.: User-friendly ontology authoring using a controlled language. Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC), Genoa, Italy, May 2006.