AR2SPARQL: An Arabic Natural Language Interface for the Semantic Web

Abstract

With the growing interest in supporting the Arabic language on the Semantic Web (SW), there is an emerging need to enable Arab users to query ontologies and RDF stores without being challenged with the formal logic of the SW. In the domain of English language, several efforts provided Natural Language (NL) interfaces to enable ordinary users to query ontologies using NL queries. However, none of these efforts were designed to support the Arabic language which has different morphological and semantic structures.

As a step towards supporting Arabic Question Answering (QA) on the SW, this work presents AR2SPARQL, a NL interface that takes questions expressed in Arabic and returns answers drawn from an ontology-based knowledge base. The core of AR2SPARQL is the approach we propose to translate Arabic questions into triples which are matched against RDF data to retrieve an answer. The system uses both linguistic and semantic features to resolve ambiguity when matching words to the ontology content. To overcome the limited support for Arabic Natural Language Processing (NLP), the system does not make intensive use of sophisticated
linguistic methods. Instead, it relies more on the knowledge defined in the ontology and the grammar rules we define to capture the structures of Arabic questions and to construct an adequate RDF representations. AR2SPARQL has been tested with two different datasets and results have shown that it achieves a good retrieval performance in terms of precision and recall.

References

approaches to semitic languages: Association for Computational Linguistics.


22. Kaufmann, E. and A. Bernstein. 2007. How useful are natural language interfaces to the semantic web for casual end-users?: Springer.


Semantic Web: Research and Applications Springer. p. 473-487.

Index Terms

Computer Science

Information Sciences

Keywords

Semantic Web, SPARQL, Natural Language Interface, Arabic Question Answering