Core techniques of
QA Systems over KBs
a Survey

Guillermo Echegoyen Blanco

Summary

Here goes the Summary

- A Question Answering System should be able to: *Understand a Natural Language Question so as to be able to answer based on some pre-known data.*

- Typically involves accepting a question and generating a SparQL query capable of extracting the information which answers the user question.

- Question Analysis
- Phrase Mapping
- Disambiguation
- Query Construction
- Distributed Knowledge

Analyze syntactic features to extract meaningful information:

- Type of question (is it a Which, What... question).
- Multilinguality (is it in English, French...).
- Correspondance to KB entities/classes.
- Tokens in the sentence and it's relations.
- Useless words in the sentence.

Techniques based on:

- Recognizing Named Entities
- Segmenting with *POS** Tags
- Identifying dependencies using parsers

POS Tag: Part-Of-Speech Tag

Identify Named Entities and map to resource in KB

- **NER Tools**: Tools from NLP, **Standford NER Tool**. Domain specific, low precision 51% (He et al. 2014)

- **N-Gram**: Map n-grams to KB entities. Adv: Each NE can be recognized in the KB. Disadv: Dissambiguation explodes (too much candidates). (SINA: Shekarpour et al. 2015, CASIA: He et al. 2014)

- **Entity Linking Tools**: **DBpedia Spotlight** (Daiber et al. 2013) and **AIDA** (Yosef et al. 2011). Recognize NE and find the underlying KB resource, dissambiguating on the way. Adv: All-in-one. Disadv: Limited service, KB dependant.

Identify which phrase correspond to instances, properties, classes. . . and which is irrelevant.

J. Daiber et al. "Improving efficiency and accuracy in multilingual entity extraction". en. In: *Proceedings of the 9th international conference on semantic systems, ACM.* 2013.

S. He et al. "CASIA@ V2: a MLN-based question answering system over linked data". nl. In: *Proceedings of QALD-4.* 2014.

S. Shekarpour et al. *Sina: semantic interpretation of user queries for question answering on interlinked data. Web Semant Sci Serv Agents World Wide Web 30(Supplement C):39–51.* en. 2015. DOI: doi:10.1016/j.websem.2014.06.002.

M.A. Yosef et al. "Aida: An online tool for accurate disambiguation of named entities in text and tables". en. In: *Proceedings of the VLDB.* 2011.