Using Ontologies to Construct Natural Language Question Grammars

Ernest Kirstein

December 19, 2014

Ontologies have classes, attributes, and relations which may convey some information that may be useful in parsing natural language questions, if only those bits of data can be conveyed in some useful fashon to a NL interpreter.

Classes might be mapped to grammatical symbols. Instances of classes with name attributes or related name class instances might be used to compile some statistical classifier of strings.

I reckon that classes in similar ontological contexts might also fit in similar gramatical contexts.

Ontologically Informed Grammars

Consider the abstract question, "Does (A) have a (B)?". What sorts of things must A and B be? In an ontological context, A and B must be connected by a relation that specifies that the class of A may possess an instance of the class of B.

"Does blue have a dog?" is not a valid question because a color cannot possess an animal. In fact, possed with such a question a person might naturally assume that 'blue' was meant to be 'Blue' (as a nicknam of some yet unknown person).

This is not a totally novel idea, systems which use ontologies that use ontological information to inform parsers have been proposed and created [1].

Consider some parseable grammar G and some ontology O. And then, also consider some bijective mapping M of terminal symbol in G, to some classes in O.

$$M : S \leftrightarrow C$$

And suppose that in the parsing of this grammar, the classification of tokens to those symbols, S, are determined by some statistical models of

the name attributes of instances of those classes C.

Now imagine that we could define some equivalence function, E, of classes, C to other classes C' based on, perhaps, the map of relations between those class? Including indirect relationships? This equivalence function would need to be constructed in such a way that all C'.

References

[1] Smaranda Muresan. Ontology-based semantic interpretation as grammar rule constraints. *Computational Linguistics and Intelligent Text Processing*, 6008:137–149, 2010.