

# 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 fashion 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, posed 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].

There's a question of efficiency, though. Is it best to have a broad grammar and let the ontology rule out invalid interpretations? Or would it be more reasonable to build the ontological restrictions into the grammar by creating new symbols and rules? I suspect that that's an optimization problem which depends on context.

Here's some question forms:

- Does/Is/Will/Can/Do (A) (R) (B)?
- Does/Is/Will/Can/Do (A)'s (B) (R) (C)?
- How many (A) (R) (B)?
- Which/what (A) (R) (B)?
- Who (R) (B)?
- Of the (A) that (R) (B), do any (R2) (C)?

Relations might be particularly tricky to parse. Dealing with multiple tenses is also an issue. ESL learning tools might serve as good sources of natural language question forms. [2]

## Conclusions

”There was a story from the dawn days of Artificial Intelligence - back when they were just starting out and no one had yet realized the problem would be difficult - about a professor who had delegated one of his grad students to solve the problem of computer vision.” - Eliezer Yudkowsky [3]

I’m begining to feel like that graduate student. It might take a person months just to *learn about* most of the simple question forms one might use in English. And that’s someone who already has full ontological knowledge, robust learning methods and resources, and a whole suite of powerful pattern-matching software in his natural, highly-parallel gray matter.

I think the goal needs to be building a robust, extensible model rather than any practical working example. Something that can grow an learn, ideally with as little expert intervention as possible. Or perhaps the goal here should just be to recognize what obstacles still exist? That could be an entire thesis in and of itself.

Maybe it’s possible to recognize elements in new question forms, figure out their ontological relationships, recognize similar question forms, and guess at a change to the system?

## References

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

- [2] Unknown. Questions - grammar exercises. [http://www.englisch-hilfen.de/en/exercises\\_list/fragen.htm](http://www.englisch-hilfen.de/en/exercises_list/fragen.htm), 2014.
- [3] Eliezer S. Yudkowsky. Hpmor. <http://hpmor.com/>, 2014.