15-400 Milestone Report V

Sonya Anopa

April 1, 2016

0.1 Milestone Progress

I have been trying to add relations / roles, but because I realized I needed to do some additional work before adding them, they have not all been added. Specifically, I have added the 'amod', 'appos', and 'compound' relations, but no other relations (to note, my goal was to include all the relations I talked about previously by this milestone, which therefore has not been achieved).

First, I wanted to make sure the heads of dependency trees I was adding as concepts to Scone were correct. So, I took a detour from adding relations by adding an additional dependency parser (a neural dependency parser) to double-check the results from the previous one. I also introduced a small heuristic for checking whether I should include any parsers' results while creating concepts. So, if the head of the parse (the word I'm trying to add as a concept and connect to the synset) does not agree between the two parsers, I try to take the more reasonable one, as I discovered that in a number of cases when the parsers disagree, one of them is correct.

After that, I added the adjective ('amod'), 'compound' relations, and 'appos'. The 'compound' relations simply add any compound words into the concept to be created - i.e. the words "phone book" are truly used as one word, so I wouldn't want to create a concept only with the work "book." Adding adjectives was pretty straightforward, but then I discovered another pathway where I paused on adding relations. When adding the concept from the head of the dependency parse tree to Scone, I want to make it as specific of a node as possible. That is, I don't want to simply say it is a "thing," which is the highest node in Scone hierarchy. I also cannot simply find the concept in WordNet because there are several senses per word, so I need to choose one of them. So, I have started doing some word sense disambiguation to try to do so. This direction of work is also useful for the 'appos' relations, which create an eq-link ("equivalent") between two concepts, as I need to be able to connect the second concept to something as well.

0.2 Looking Ahead

I am going to continue adding the roles/relations not added yet and working on some of the less well-defined tasks like word sense disambiguation (since it's not a solved problem in NLP as is, I'm not going to go too far in). Hopefully I will finish adding the relations before the next milestone and start making the tool more general, as it is currently very specific to WordNet and more hard-coded.

0.3 Revisions to Future Milestones

The next milestone originally was to have a limited import of WordNet to check for correctness. As I have been doing that all along, I don't think it is necessary as a milestone, which is useful given that I haven't entirely finished my previous milestone. Instead, my milestone will be (as above) to finish adding relations / roles. The last milestone suggested I finish importing WordNet, which is also something that should not take long. Therefore, the last milestone will instead be to generalize the tool to accept definitions rather than WordNet specifically.

0.4 Resources Needed

I have all the resources I need so far.