29 November 2016: Discussion on the non-programming questions in
Assignments

Question from Assignment 3: POS Tagging and learner language

Here are the links to some online demos for POS taggers developed by NLP researchers:

- `http://cogcomp.cs.illinois.edu/page/demo\_view/pos`

- `http://nlp.stanford.edu:8080/parser/` - look at only the tagger output in this demo.

- `http://morphadorner.northwestern.edu/postagger/example/`

- `http://ucrel.lancs.ac.uk/claws/trial.html`

Pick any two taggers from these, and do the following:

1. Create a list of 10 sentences written by English learners at an Elementary proficiency (You can use any sentences from corpus resources you already have. Or pick something from `http://cblle.tufs.ac.jp/llc/icci/search.php?menulang=en`. Or contact me for any other learner corpus resources).

2. Run the taggers on these 10 sentences, and write what you think about the robustness of taggers to such ungrammatical input (and misspelt words). You should write about about both the taggers, and compare their performance. Make sure you have two types of sentences in your chosen list: a) sentences that are incorrect (mis-spelt or ungrammatical), yet, do not affect POS tagging. b) sentences that are incorrect and affect POS tagging.

Note: The taggers may or may not follow the same tagset. Some taggers by nature give fine grained tags, and some do not. Your task is to just notice how accurate they are with learner text and compare the taggers in terms of accuracy, not on the differences between the granularity of tags. Your commentary can be up to 2 pages long. List the sentences you tried with at the end of this document, as an appendix in an additional page.

Question from Assignment 4: Spelling and grammar checking tools

Analyse any two existing spelling and grammar checking tools (e.g., cyWrite, grammarly, languagetool.org etc.) and write a short (1-2 page) summary

of your observations about when the tools work and when they do not, with examples. You don't have to write any programs for this. Access the web interfaces of these tools, test with a few sentences, and draw some conclusions. Your report can be upto 2 pages long. Write a few lines on what tools you chose, and what sentences you tried them with. Then, analyse the performance of each tool, and compare their performances.

Question from Assignment 5: Parsers and learner language

Here is the link to Stanford Parser's online demo. `http://nlp.stanford.edu:8080/parser/`. Now, do the following:

1. Create a list of 10 sentences written by English learners at an Elementary proficiency (You can use any sentences from corpus resources you already have. Or pick something from `http://cblle.tufs.ac.jp/llc/icci/search.php?menulang=en`. Or contact me for any other learner corpus resources). You can use the same 10 sentences from previous assignment.

2. Run the parser on these 10 sentences, and study the phrase structure tree and Universal dependencies representation there (You should also read a little bit on what those tags mean). Now, write what you think about the robustness of both the representations to such ungrammatical input (and misspelt words). You should write about about both phrase structure and dependency representations , and compare their performance. Make sure you have two types of sentences in your chosen list: a) sentences that are incorrect (mis-spelt or ungrammatical), yet, do not affect POS tagging. b) sentences that are incorrect and affect POS tagging.

Note: Your commentary can be up to 2 pages long. List the sentences you tried with at the end of this document.