

Natural Language Processing Homework 4

Katie Chang, Willis Wang

March 22, 2016

README

1 State of the Art Parser

Discussion

Using the Berkeley parser.

(a) Was there anything interesting or surprising about the style of trees produced by the parser?

(b) What were some things that the parser got wrong? What were some hard things that it managed to get right?

(c) Can you stump it with a grammatical sentence that you make up? With a grammatical sentence that you find in an online document?

1.1 Extra Credit

TurboParser does not give us a tree that is rooted at a root node. Instead, the tree separates phrases and seems to use the end words of each phrase to connect each phrase. In Figure 1, there is no tree root, and the nodes VB, NNS, and VBP are in between each distinct phrase.

The Link Grammar Parser does something interesting, as shown in Figure 2: it parses the sentence and identifies links between phrases / words in the sentence. Using this parse, the software creates a constituent tree, which looks much like the trees that we produced in homework 1.

Used overleaf.com to generate LaTeX document.

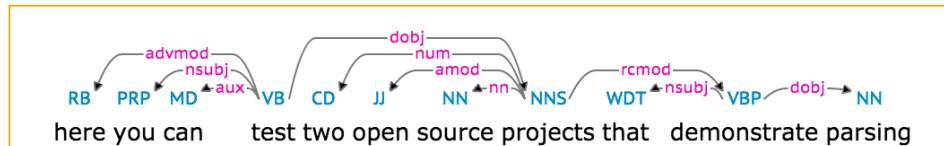


Figure 1: TurboParser Tree Diagram of a sample sentence.

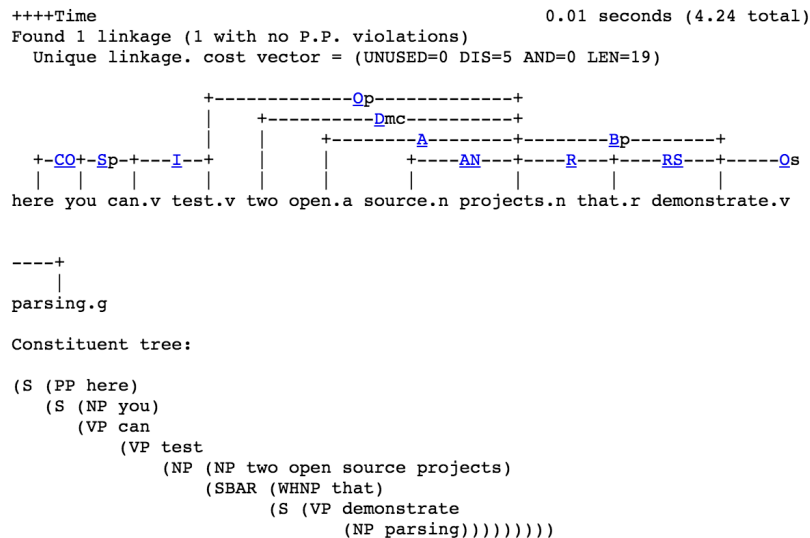


Figure 2: LinkGrammarParser Tree Diagram of a sample sentence.