

NLP Assignment 2 Report

Authors & contributors:

Petrus Kiendys, Amin Abu-Sneineh,
Oskar Berntorp, Azddin Bezberkan

Carrying out the experiments

Choose two (reasonable) sentences from the training set, and draw the tree structure by hand

Answer: We chose the following two sentences:

- “Individuell beskattning av arbetsinkomster” (line 1 from the files containing bracketing formatted phrase structure trees)
- “Kommunalskatteavdraget slopas” (line 17 from the files containing bracketing formatted phrase structure trees)

dependency trees

[insert diagram of sent1_deptree]

[insert diagram of sent2_deptree]

[cfg-pos] style trees

[insert diagram of sent1_cfg-pos]

[insert diagram of sent2_cfg-pos]

[cfg-dep] style trees

[insert diagram of sent1_cfg-dep]

[insert diagram of sent2_cfg-dep]

Does this algorithm do a good job of creating phrase-structure trees from a linguistic point of view?

Answer: It does not. On the following image we can see sentence 2 being transformed from a dependency tree into a phrase-structure tree:

[insert diagram of sent2_deptree]

[insert diagram of sent2_pstree (via algorithm 1)]

As we can see, by comparing the tree produced by algorithm 1 to the phrase structure trees of the bracketing formatted files, the algorithm is too simplistic and does a poor job of creating proper phrase structure trees.

It's important to note that phrase-structure trees adhere to constituency relations and dependency trees adhere to dependency relations. As such it is hard to transform a dependency tree to a phrase structure tree with simplicity.

Furthermore, the algorithm only specifies how nodes should be manipulated in the input dependency tree and does not specify how the dependency label that describe the relationship between each node should be presented on the phrase-structure tree. Simply adding a "P" suffix to a non-terminal node does not properly solve this issue as the label contains a word/terminal that was transformed from the dependency tree rather than a having a label consisting of a phrase (such as VB, NP, etc.).

[fill in Q&A...]