Dependency Parsing

Language Technology 1
WS 2014

Günter Neumann

Exercise 1

Parse the sentence:

$_0$ He $_1$ said$_2$ he $_3$ will $_4$ now $_5$ consider $_6$ those $_7$ offers $_8$. $_9$

Assume the gold standard: <i,j> means there is a link from word with index i to j, and word i is the modifier and word j is the head. The link/edge is not labelled.

{<1,2>, <2,0>, <3,4>, <4,2>, <5,4>, <6,4>, <7,8>, <8,6>, <(9,2>}

Choose any algorithm of your choice (might as well be your own)

Specify how different configurations look like in each step

Specify a feature model (i.e. feature templates you would like to use) and

instantiate it with concrete values for each configuration

- assume that you have word forms and POS tags,
  but no morphological information available

Discuss the advantages and disadvantages of your algorithm (runtime complexity, number of configurations, determinism, incrementality etc.)

Exercise 2

1. Goto http://mdparser.sb.dfki.de/

2. Download DFKI's MDParser.

3. Install the MDParser as specified in the README.txt file

4. Run the mdpTrainTest.jar command as specified in the README.txt file with the training and test data for English available in the folder resources/input/ (which you will have after installation of the MDParser).

5. Run the MDParser on some text using the trained model.

6. Write a small (half page) report about the "readability" of the README.txt and your experience when solving this exercise 2 (what was clear, not so clear, difficult etc.? how can the documentation be improved? Etc.)