

Assignment #6: Dependency parsing using machine learning techniques

Elliot Jalgard
ada10eja@student.lu.se

October 15, 2014

1 Introduction

The objective of this assignment was to use the trained Weka models from the previous assignment to train a classifier and compare our results to the gold standard file.

2 Execution and Results

We started out by complementing the `parse()` method in `Parser.java`. We created two different versions of the method, one for unlabeled models and one for labeled. We then added two classes, `Guide2.java` and `Guide6.java` and ran the program for each type of model we trained in assignment five.

The tables below show the results of the `eval.pl` script. The additional parameters used for the last tables are the third postag on the stack and in the queue.

2 parameters, unlabeled	Score	2 parameters, labeled	Score
Labeled attachment	0.84 %	Labeled attachment	60.05 %
Unlabeled attachment	72.85 %	Unlabeled attachment	72.93 %
Label accuracy	0.84 %	Label accuracy	63.19 %

4 parameters, unlabeled	Score	4 parameters, labeled	Score
Labeled attachment	0.46 %	Labeled attachment	68.13 %
Unlabeled attachment	81.08 %	Unlabeled attachment	80.50 %
Label accuracy	0.46 %	Label accuracy	70.76 %

6 parameters, unlabeled	Score	6 parameters, labeled	Score
Labeled attachment	0.46 %	Labeled attachment	68.19 %
Unlabeled attachment	81.40 %	Unlabeled attachment	80.72 %
Label accuracy	0.46 %	Label accuracy	71.00 %

3 Discussion

The jump in accuracy between two and four parameters is large, while the jump between four and six parameters is barely noticable. In order to increase the score, we need to be more creative with the parameters and not only choose the next postag on the stack and in the queue.