

# Final Paper: Language models BSc AI 2013

Eszter Fodor, Sharon Gieske  
(5873320, 6167667)

May 19, 2013

## **Abstract**

This paper describes our implementation of a Markov POS tagger, trained and tested on various sections of the Wall Street Journal corpus. We will describe the research question that we wanted to answer, how we approached the problem and what our results were.

## **Introduction**

The problem that needed to be solved during the four assignments was how to calculate the most probable Part-of-speech (POS) tag sequence using tagging and smoothing on a training corpus. The program that needed to be written would eventually be able to tag the test corpus with POS tags.

Constructing language models is an interesting and important research, because it enables language scientists to examine

## **Approach**

## **Results**

## **Conclusion/Discussion**