

Part A step 2

Maurits Offerhaus 10400036

Amir Alnomani 10437797

Joram Wessels 10631542

February 2016

introduction

Compiles using Python version 3.5 with the following command:

```
assignment2.py -corpus [path] -n [value] -conditional-prob-file [path]  
-sequence-prob-file [path] -scored-permutations [path]
```

1

We used our code from the first step.

These are the 10 most common bigrams in the corpus with their frequencies:

```
python assignment2.py -corpus [path] -n [value]
```

BiGram	Frequency
of the	2507
to be	2232
in the	1917
I am	1365
of her	1264
to the	1142
it was	1010
had been	995
she had	978
to her	964

2

The probabilities of the ngrams in "conditional prob":

I do not	0.7297297297297297
in the world	0.11163275952008346
I am sure	0.2681318681318681
she could not	0.3531468531468531

3

4