Assignment 4: Hidden Markov Models

L645, Sandra Kuebler

DUE: at beginning of class on Thursday, October 23

1. Forward / Backward Algorithm

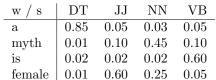
Consider the example of the HMM for POS tagging the sentence "a myth is a female moth". Calculate the following forward accumulators given the following probabilities.

If you are enrolled in B659, write a program to calculate this. Submit your code and an output. Otherwise, you can do the calculation manually.

initial vector: [0.45 0.35 0.15 0.05]

transition matrix:

t / t+1	DT		NN	VB
DT	0.03	0.42	$0.50 \\ 0.65 \\ 0.15$	0.05
JJ	0.01	0.25	0.65	0.09
NN	0.07	0.03	0.15	0.75
VB	0.30	0.25		0.30



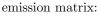
0.13

0.25

0.12

moth





- (a) $\alpha_4(NN)$
- (b) $\alpha_3(VB)$
- (c) $\alpha_1(DT)$
- (d) $\beta_4(NN)$
- (e) $\beta_3(NN)$

Give the probabilities of the accumulators that you need for these calculations and explain how you got to your results. In a program, that translates into having explicit output and comments in the program.

0.20

20 pts.

2. Viterbi Algorithm

Consider the definitions for initialization, induction, and storing the backtrace on slide 6. Modify the formulas so that the HMM uses a trigram model instead of a bigram model. If you need to introduce new subscripts, explain what they are!

10 pts.

3. Additional: Parameter Estimation

Write a program that reestimates the first n rounds of Baum-Welch for the POS tagging example on slide 9 of the state sequence slides for the signal sequence "a female moth is a myth". Use the formulas on page 12 of the parameter estimation slidesfor the reestimation.

additional 20 pts.