CS114B (Spring 2022) Lab 8 Exercise
Structured Perceptrons

March 24, 2022

(You may find the discussion in Chapter 7.5.1 of the Eisenstein book helpful.)

Suppose we are given the following weight matrix $\boldsymbol{\Theta}$:

$\boldsymbol{\Theta}$	$y_i = \ldots$		
	NN	VB	DT
y_{i-1} = <S>	-0.3	-0.7	0.3
y_{i-1} = NN	-0.7	0.3	-0.3
y_{i-1} = VB	-0.3	-0.7	0.3
y_{i-1} = DT	0.3	-0.3	-0.7
x_i = Alice	-0.3	-0.7	0.3
x_i = admired	0.3	-0.3	-0.7
x_i = Dorothy	-0.3	0.3	-0.7
x_i = every	-0.7	-0.3	0.3
x_i = dwarf	0.3	-0.7	-0.3
x_i = cheered	-0.7	0.3	-0.3

(If you think this matrix looks like π, $\mathbf{A}$, and $\mathbf{B}$ stacked on top of each other, you are right! Note that we don't need to account for the unknown word <UNK>, and for simplicity, we will ignore the bias term.)

1. Suppose we are given the following training sentence:

   `Alice/NN admired/VB Dorothy/NN`

   (a) Use the Viterbi algorithm to compute the best tag sequence. As part of your answer, you should fill in the Viterbi trellis below. You should also keep track of backpointers, either using arrows or in a separate table.

	Alice	admired	Dorothy
NN			
VB			
DT			

   (b) Update the weight matrix. Use a constant learning rate $\eta = 1$.

2. Suppose we are given the following testing sentence:

   `Alice cheered`

   Use the Viterbi algorithm to compute the best tag sequence. Again, you should fill in the Viterbi trellis below, and keep track of backpointers.

	Alice	cheered
NN		
VB		
DT		