

Computational Linguistics Final
Viterbi Problem
Jean Mark Gawron
San Diego State University
Linguistics 581
May 5, 2014

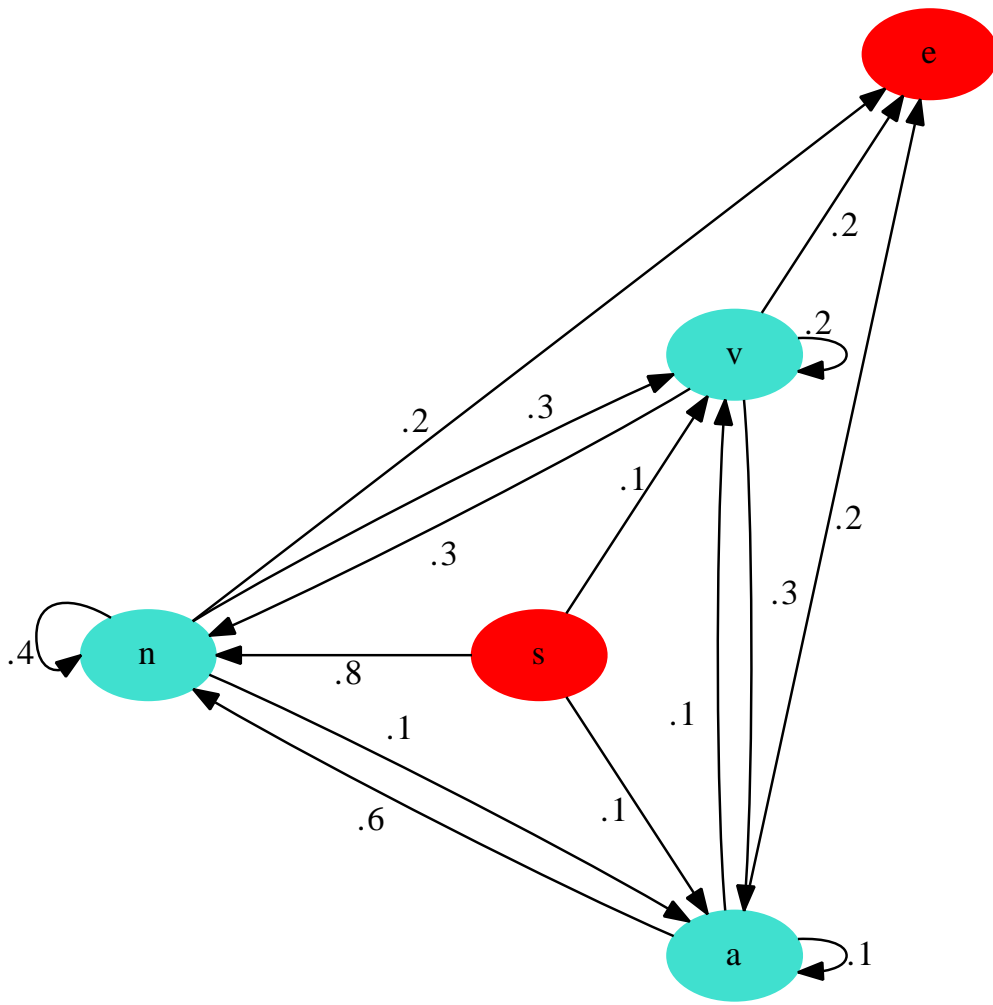
1 Viterbi Algorithm

Figure 1 shows another toy HMM to be used for tagging.

1. The numbers on the transition are transition probabilities. The .5 on the transition from state *s* to state *n* means you can make that transition with a probability of .5, that is, $P(n | s) = .5$.
2. Best path for this problem means highest probability.
3. The difference between this HMM and the one in the homework problem is that there is both a start and an end state. The difference between this HMM and the one on the midterm is that there are observation probabilities for the words “start” and “end” and these are always inserted before and after any input string. So, all observations start at the *s* state with the word *start* and end at state *e* with the word *end*. Tagging the first word in a sentence as a noun means transitioning from the start state “s” to the “n” state.
4. The observation probabilities of our HMM are as follows:

w v	w a	w n	w s	w e
$\begin{bmatrix} \text{dog} & .25 \\ \text{fly} & .25 \\ \text{flies} & .5 \\ \text{start} & 0 \\ \text{end} & 0 \end{bmatrix}$	$\begin{bmatrix} \text{dog} & .7 \\ \text{fly} & .15 \\ \text{flies} & .15 \\ \text{start} & 0 \\ \text{end} & 0 \end{bmatrix}$	$\begin{bmatrix} \text{dog} & .3 \\ \text{fly} & .5 \\ \text{flies} & .2 \\ \text{start} & 0 \\ \text{end} & 0 \end{bmatrix}$	$\begin{bmatrix} \text{dog} & 0 \\ \text{fly} & 0 \\ \text{flies} & 0 \\ \text{start} & 1.0 \\ \text{end} & 0 \end{bmatrix}$	$\begin{bmatrix} \text{dog} & 0 \\ \text{fly} & 0 \\ \text{flies} & 0 \\ \text{start} & 0 \\ \text{end} & 1.0 \end{bmatrix}$

To help you complete the problem, here are Viterbi and backtrace table templates appropriate for the given input. The Viterbi values for $t=0$ the backtrace values for $t=1$ have generously been filled in.



Viterbi FSA

Figure 1: FSA for Viterbi Problem

Viterbi

e	0.0				
v	0.0				
n	0.0				
a	0.0				
s	1.0				
	0	1	2	3	4
	start	dog	flies	fly	end

Backtrace

e					
v		s			
n		s			
a		s			
s					
	0	1	2	3	4
	start	dog	flies	fly	end

Your task in this problem is to use the Viterbi algorithm to find the **best** path through the HMM in Figure 1 for the following input:

dog fly flies

You need to fill out the Viterbi table with appropriate Viterbi numbers. You also need to fill out the backtrace table with the states transitioned from on best paths. Finally, you will need to interpret your final input by stating what the best path through the HMM is, given this input.