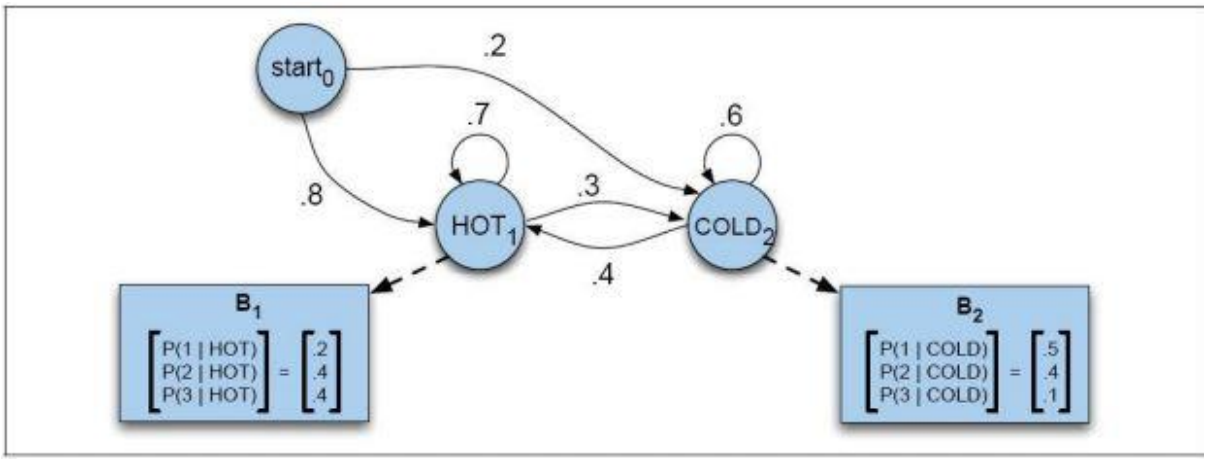
NLP Homework 3

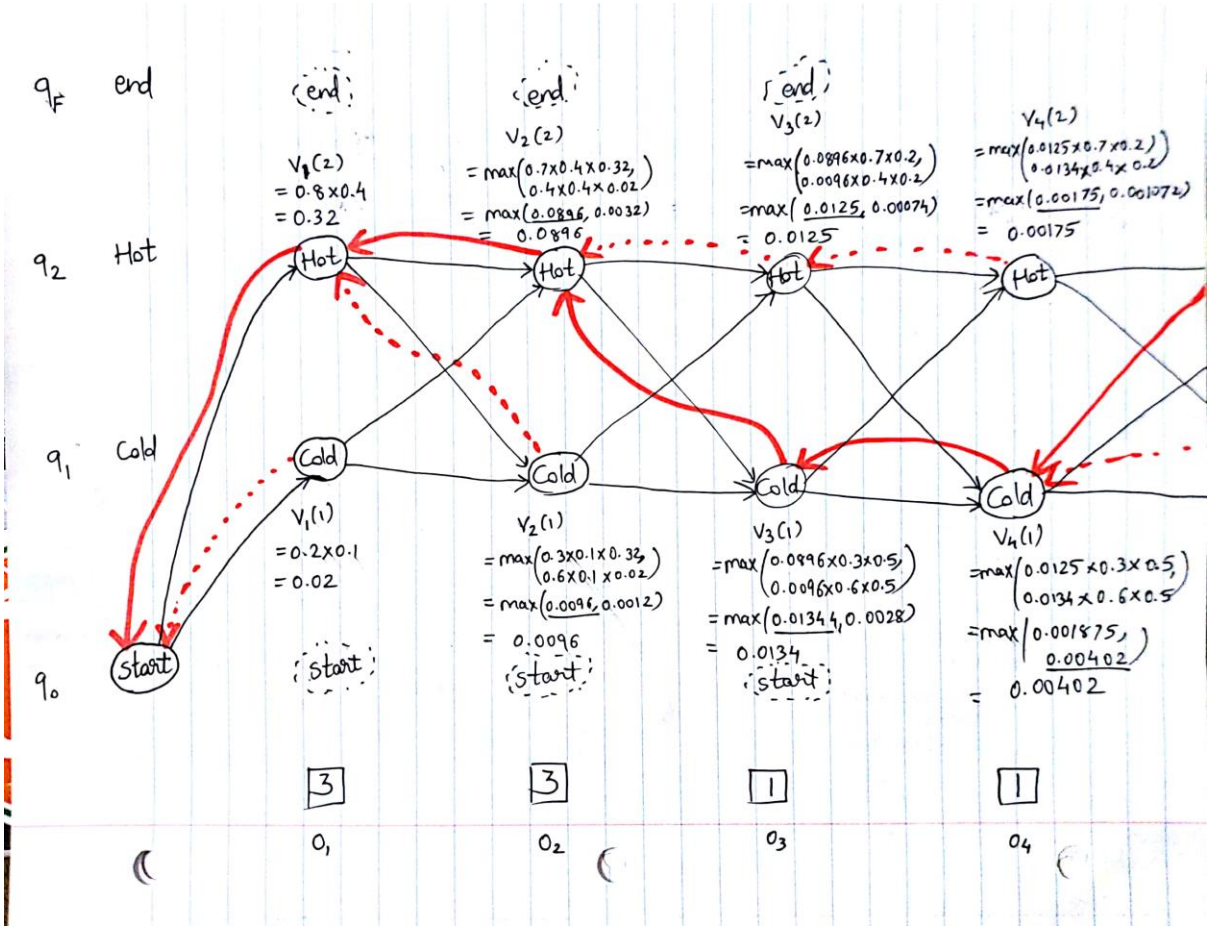
Submitted By: Yash Pradhan

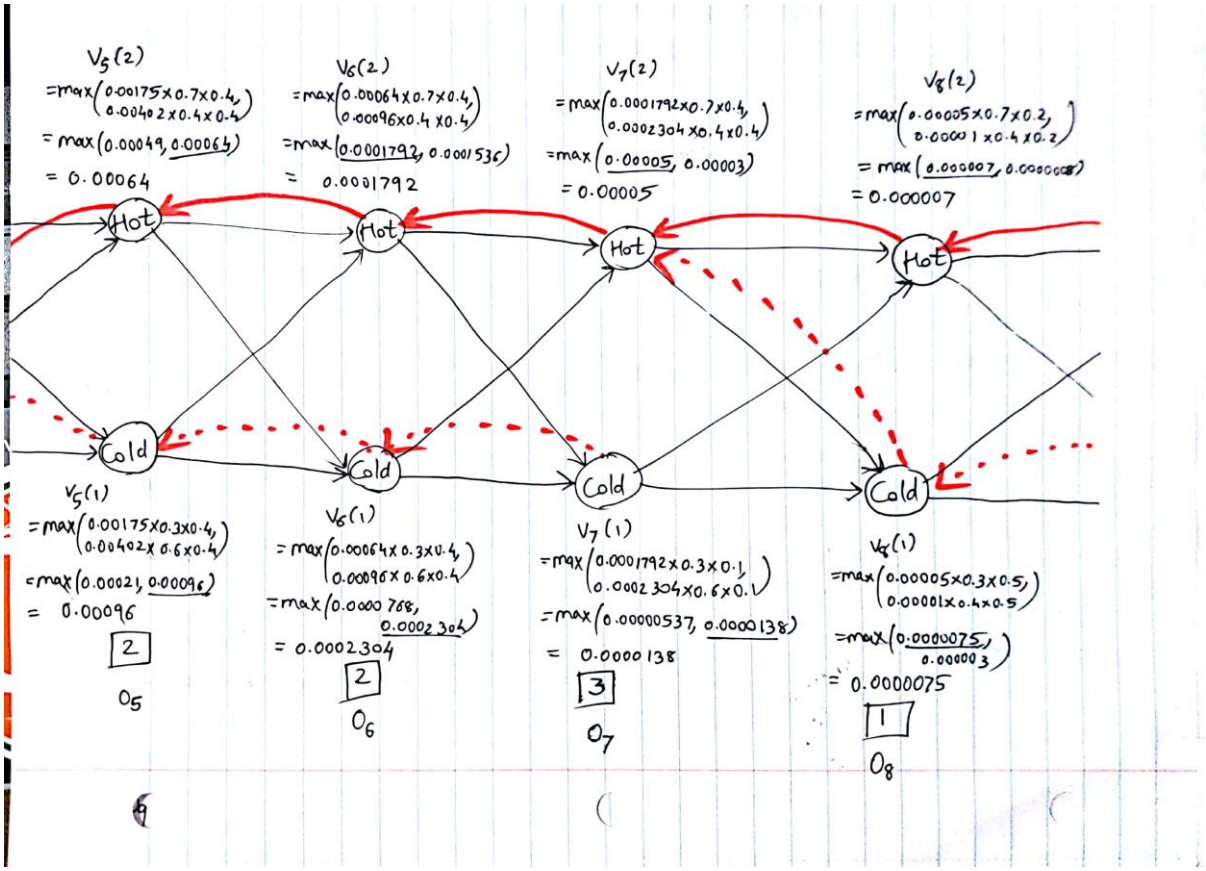
1. HMM Decoding: Viterbi Algorithm

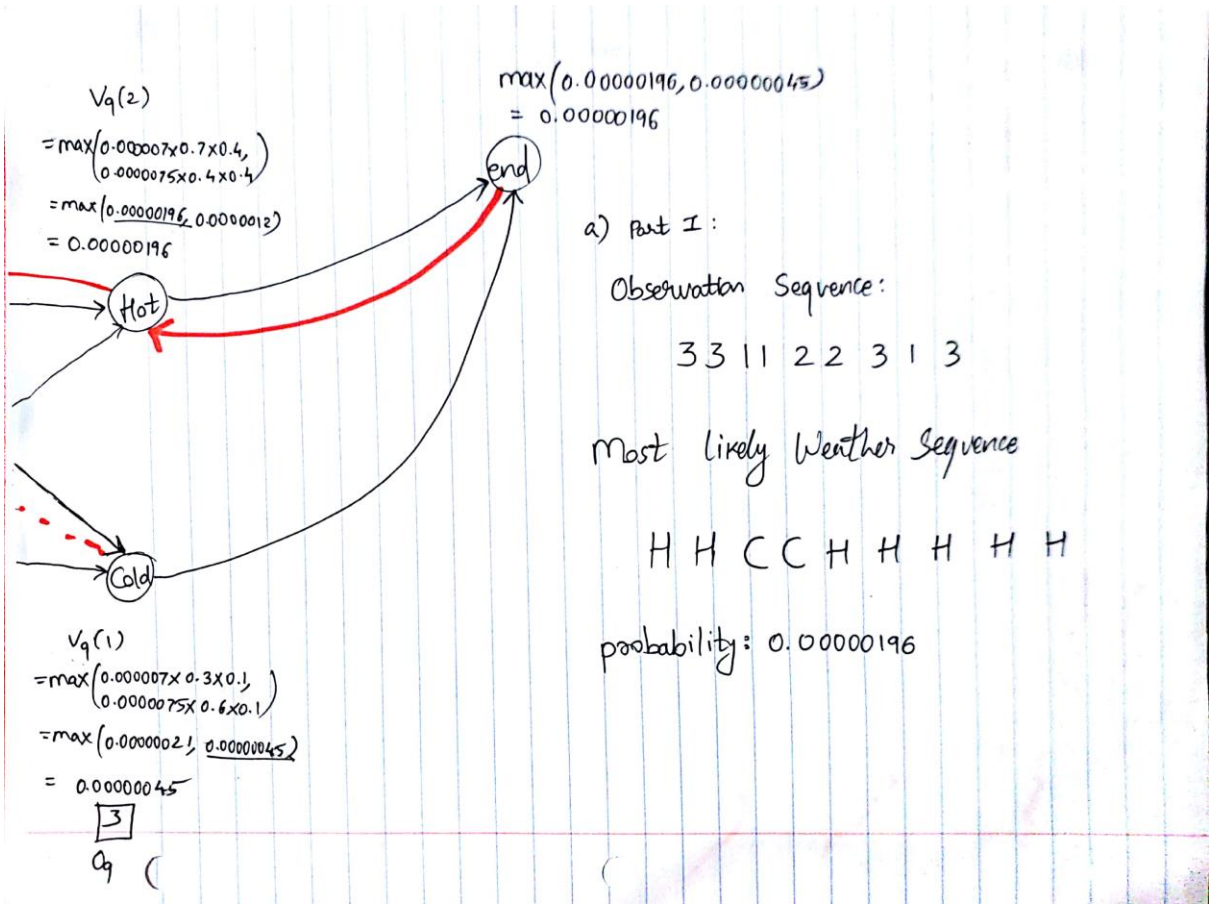


Manually build the Viterbi trellis to compute the most likely weather sequences for each of the two observation sequences

Sequence 1: 331122313

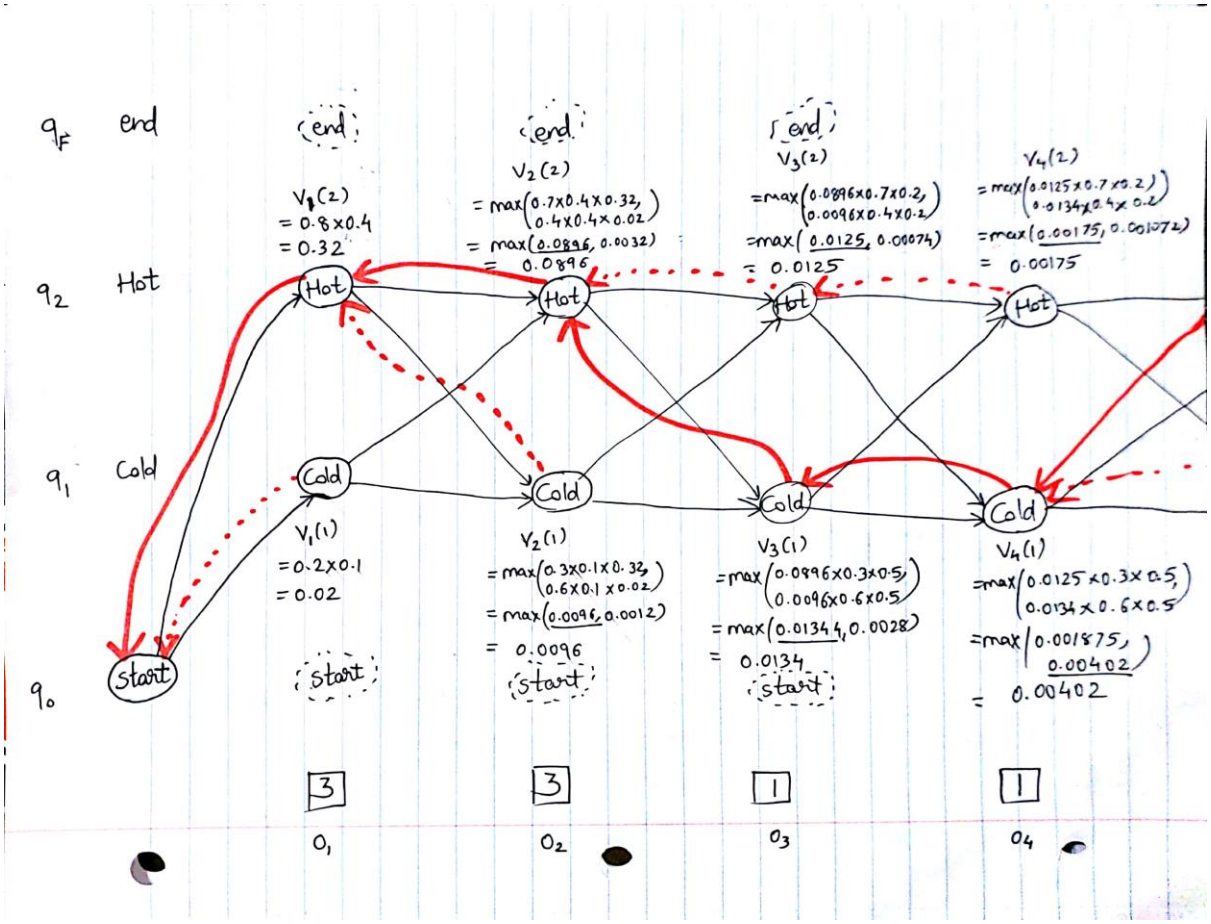


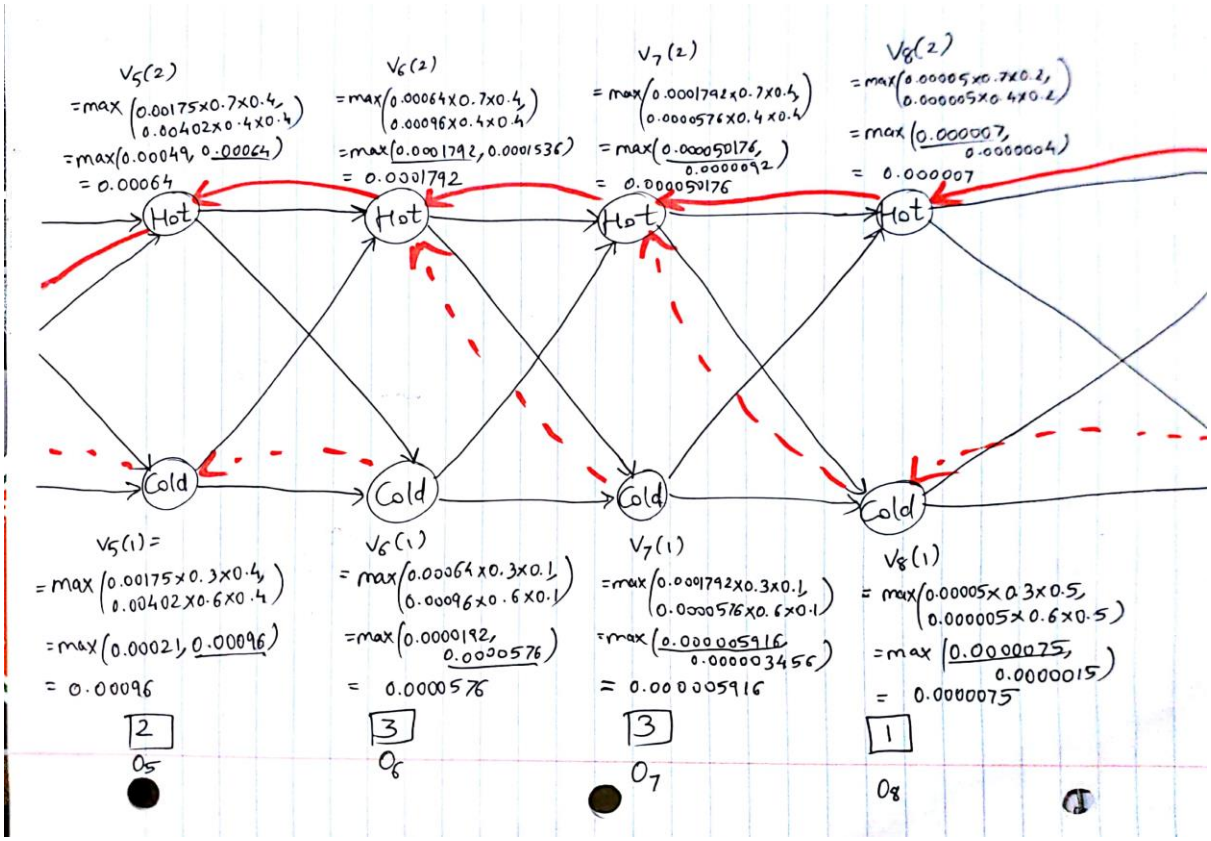


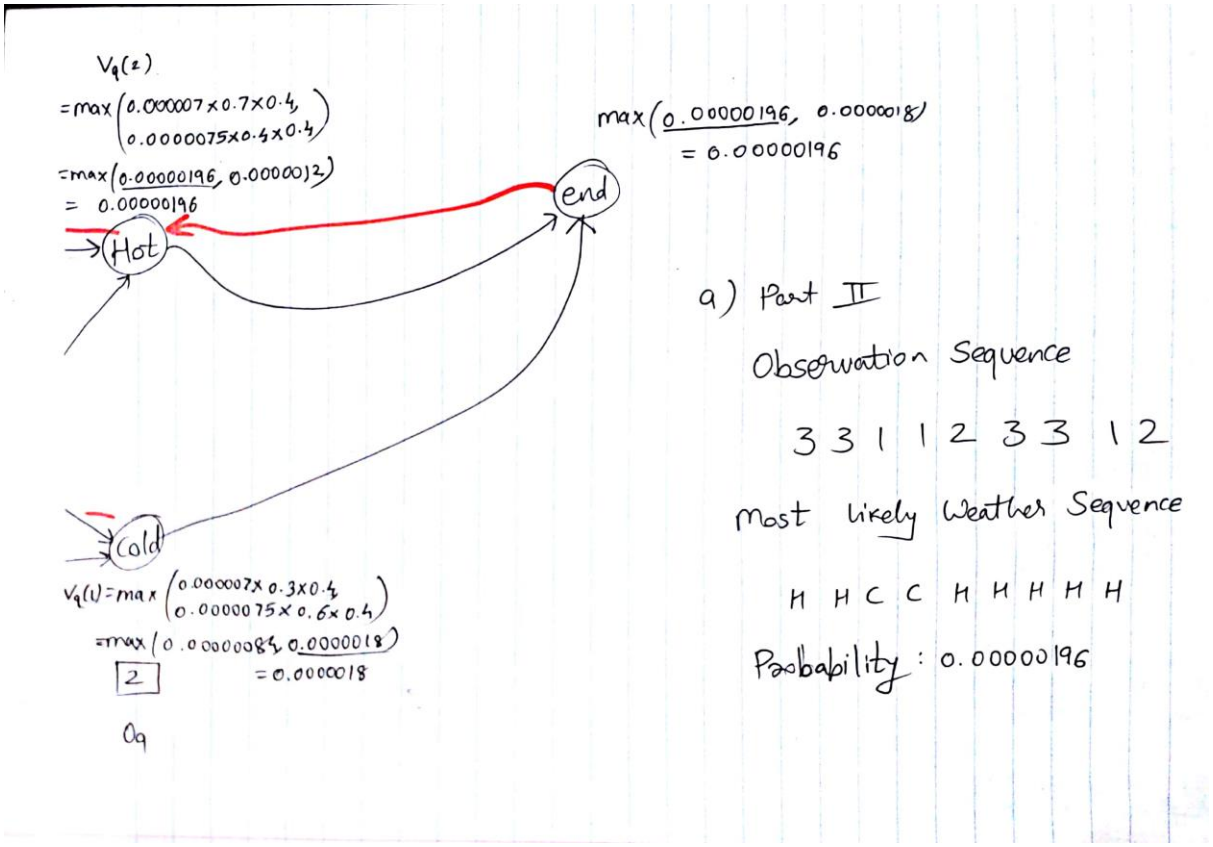


The solid red line, helps us trace back to find out the most likely sequence.

Sequence 2: 331123312







The solid red line, helps us trace back to find out the most likely sequence.

2. Maximum Entropy Modelling

$$f_1(c,x) = \begin{cases} 1 \text{ if } word_i = \text{"race"} \ \& \ c = \text{NN} \\ 0 \text{ otherwise} \end{cases}$$

$$f_2(c,x) = \begin{cases} 1 \text{ if } t_{i-1} = \text{TO } c = \text{VB} \\ 0 \text{ otherwise} \end{cases}$$

$$f_3(c,x) = \begin{cases} 1 \text{ if } t_{i-1} = \text{DT } c = \text{NN} \\ 0 \text{ otherwise} \end{cases}$$

$$f_4(c,x) = \begin{cases} 1 \text{ if is\_lower\_case}(word_i) = \text{"race"} \ \& \ c = \text{VB} \\ 0 \text{ otherwise} \end{cases}$$

$$f_5(c,x) = \begin{cases} 1 \text{ if } word_i = \text{"race"} \ \& \ c = \text{VB} \\ 0 \text{ otherwise} \end{cases}$$

$$f_6(c,x) = \begin{cases} 1 \text{ if } t_{i-1} = \text{TO} \ \& \ c = \text{NN} \\ 0 \text{ otherwise} \end{cases}$$

		Weights					
		f1	f2	f3	f4	f5	f6
Tags	VB	0	0.75	0	0.10	0.15	0
	NN	0.3	0	0.9	0	0	-0.2

Sentences:

a. Secretariat/NNP is/VBZ expected/VBN to/TO race/?? tomorrow/NN

b. the/DT race/?? for/IN outer/JJ space/NN

2. Maximum Entropy Modeling.

Compute best tag for word "race"?

		f1	f2	f3	f4	f5	f6
Tags	VB	0	0.75	0	0.10	0.15	0
	NN	0.3	0	0.9	0	0	−0.2

a. Secretariat/NNP is/VBZ expected/VBN to/TO race/?? tomorrow/NN.

		f1	f2	f3	f4	f5	f6
VB	f	0	1	0	1	1	0
NN	f	1	0	0	0	0	1

$$P(VB|race) = \frac{e^{0.75}\, e^{0.10}\, e^{0.15}}{e^{0.75}\, e^{0.10}\, e^{0.15} + e^{0.3}\, e^{-0.2}} = \frac{2.71828}{2.71828 + 1.10517}$$

$$= 0.7109$$

$$P(NN|race) = \frac{e^{0.3}\, e^{-0.2}}{e^{0.75}\, e^{0.10}\, e^{0.15} + e^{0.3}\, e^{-0.2}} = \frac{1.10517}{2.71828 + 1.10517}$$

$$= 0.289$$

Ans: Best tag for race is VB.

b. the/DT race/?? for/IN outer/JJ space/NN.

		f1	f2	f3	f4	f5	f6
VB	f	0	0	0	1	1	0
NN	f	1	0	1	0	0	0

$$P(VB|\text{race}) = \frac{e^{0.10}\, e^{0.15}}{e^{0.10}\, e^{0.15} + e^{0.3}\, e^{0.9}}$$

$$= \frac{1.2840}{1.2840 + 3.3201}$$

$$= 0.27888$$

$$P(NN|\text{race}) = \frac{e^{0.3}\, e^{0.9}}{e^{0.10}\, e^{0.15} + e^{0.3}\, e^{0.9}}$$

$$= \frac{3.3201}{1.2840 + 3.3201}$$

$$= 0.7211$$

The Best tag for race is NN.

Answers:

a. Secretariat/NNP is/VBZ expected/VBN to/TO **race/VB** tomorrow/NN

b. the/DT **race/NN** for/IN outer/JJ space/NN