Josh Bishop
Prof. Ritter
due Fri: 3/31 @ 11:59pm

CS 7650 - HW #2

1. Word Embeddings

A.) Compute the CBOW vector representation of the missing word for h=3

$\rightarrow \bar{V}_m = \frac{1}{2h} \sum_{n=1}^{h} V_{w_{m+n}} + V_{w_{m-n}}$

$\rightarrow \bar{V}_3 = \frac{1}{2(3)} \left(\begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 8 \\ -4 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \\ -4 \end{bmatrix} + \begin{bmatrix} -2 \\ -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 4 \\ 0 \end{bmatrix} \right)$

$\rightarrow \boxed{\bar{V}_3 = \begin{bmatrix} 1 & 2 & -\frac{1}{3} \end{bmatrix}}$

B.) Fill in the scores of each word in Table 2 being the missing word in Table 1.

$\rightarrow \log p(w) = \log p(w_m \mid w_{m-h}, w_{m-h+1}, \dots, w_{m+h-1}, w_{m+h})$

$\rightarrow \log p(w) = \log \frac{\exp(u_{w_m} \cdot \bar{V}_m)}{\sum_{j=1}^{V} \exp(u_j \cdot \bar{V}_m)}$

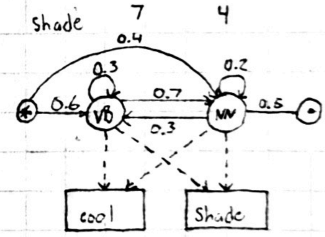
$\rightarrow \log p(w) = u_{w_m} \cdot \bar{V}_m - \log \sum_{j=1}^{V} \exp(u_j \cdot \bar{V}_m)$

Word	Embedding	$u_j \cdot \bar{V}_m$	Unnormalized log p(w)	Normalized p(w)
yellow	[-2, 4, 2]	16/3	-2.848	0.139
pink	[-6, 3, -6]	2	-6.181	0.014
blue	[0, 4, 2]	22/3	-0.848	0.556
orange	[2, 0, 0]	2	-6.181	0.014
white	[1, 3, 2]	19/3	-1.848	0.278

C.) CBOW would predict $\boxed{\text{"blue"}}$ as the missing word

2. Hidden Markov Models and the Viterbi Algorithm

	NN	VB
cool	3	6
shade	7	4



A.) Compute the emission probabilities for each word given each POS tag.
Using Bayes rule, we obtain

$\rightarrow P(cool \mid VB) = P(VB \mid cool) P(cool) / P(VB) = \frac{6}{9} \cdot \frac{9}{20} \div \frac{1}{2} = \boxed{0.6}$

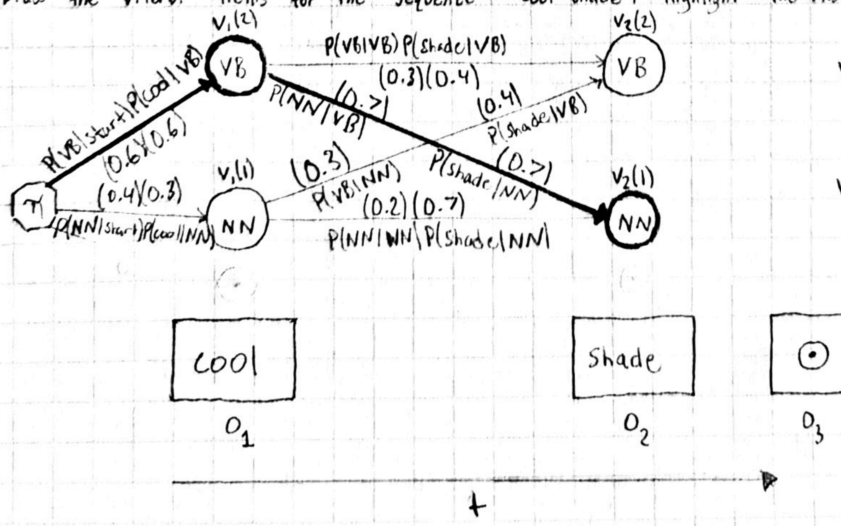
$\rightarrow P(shade \mid VB) = P(VB \mid shade) P(shade) / P(VB) = \frac{4}{11} \cdot \frac{11}{20} \div \frac{1}{2} = \boxed{0.4}$

$\rightarrow P(cool \mid NN) = P(NN \mid cool) P(cool) / P(NN) = \frac{3}{9} \cdot \frac{9}{20} \div \frac{1}{2} = \boxed{0.3}$

$\rightarrow P(shade \mid NN) = P(NN \mid shade) P(shade) / P(NN) = \frac{7}{11} \cdot \frac{11}{20} \div \frac{1}{2} = \boxed{0.7}$

	cool	shade
NN	0.3	0.7
VB	0.6	0.4

B.) Draw the Viterbi trellis for the sequence "cool shade". Highlight the most likely.



$v_1(2) = 0.36$

$v_2(2) = \max(v_1(2)(0.12), v_1(1)(0.12))$
$= 0.0432$

$v_1(1) = 0.12$

$v_2(1) = \max(v_1(2)(.49), v_1(1)(.14))$
$= 0.1764$

$V_3 = \max(v_2(1)(0.5), \cancel{v_2(2)(0)})$
$V_3 = 0.0882$

The terminal state ⊡ has no hidden states associated with it, and is only reachable from NN

CS 7650 - NLP HW #2

Josh Bishop
Prof. Ritter
due Fri 3/31 11:59pm

3. Named Entity Recognition

Sentence	Sam	works	at	Berkshire	Hathaway	headquartered	in	Nebraska
Gold Labels	B-PER	O	O	B-ORG	I-ORG	O	O	B-LOC
Sys # 1	O	O	O	B-ORG	O	O	O	B-LOC
Sys # 2	B-PER	O	O	O	O	O	O	B-LOC
Sys # 3	B-PER	O	O	B-ORG	I-ORG	O	O	B-LOC
Sys # 4	B-PER	I-PER	O	B-ORG	I-ORG	O	O	O

For each system compute:

a.) Precision = TP/(TP+FP)

$\text{Precision}_{Sys\#1} = \frac{2}{2+0} = \boxed{1}$

$\text{Precision}_{Sys\#2} = \frac{2}{2+0} = \boxed{1}$

$\text{Precision}_{Sys\#3} = \frac{4}{4+0} = \boxed{1}$

$\text{Precision}_{Sys\#4} = \frac{3}{3+1} = \boxed{\frac{3}{4}}$

b.) Recall = TP/(TP+FN)

$\text{Recall}_{Sys\#1} = \frac{2}{2+2} = \boxed{\frac{1}{2}}$

$\text{Recall}_{Sys\#2} = \frac{2}{2+2} = \boxed{\frac{1}{2}}$

$\text{Recall}_{Sys\#3} = \frac{4}{4+0} = \boxed{1}$

$\text{Recall}_{Sys\#4} = \frac{3}{3+1} = \boxed{\frac{3}{4}}$

c.) F-1 Score = 2 · Recall · Precision / (Recall + Precision)

$\text{F-1}_{Sys\#1} = 2(\frac{1}{2})(1)/(\frac{1}{2}+1) = \boxed{\frac{2}{3}}$

$\text{F-1}_{Sys\#2} = 2\cdot(\frac{1}{2})(1)/(\frac{1}{2}+1) = \boxed{\frac{2}{3}}$

$\text{F-1}_{Sys\#3} = 2\cdot(1)(1)/(1+1) = \boxed{1}$

$\text{F-1}_{Sys\#4} = 2\cdot(\frac{3}{4})(\frac{3}{4})/(\frac{3}{4}+\frac{3}{4}) = \boxed{\frac{3}{4}}$