Assignment 4 **LE, Nhat Hung**
COMP 424 - Artificial Intelligence McGill ID: 260793376
Prof. Jackie Chi Kit Cheung Date: April 3, 2019
Winter 2019 Due date: April 10, 2019

Question 1: HMMs for Part-of-Speech (POS) Tagging
In computational linguistics, **part-of-speech tagging** (or **POS tagging**) is the process of marking up words in texts as corresponding to their "part of speech identifications". A simplest form of this task is to identify each word as noun, verb, adjective, adverb, etc. Here we demonstrate the power of HMM by applying it on POS tagging.
You need to design an HMM to make predictions regarding the POS identifications of words in an English sentence. The **observed states** are the words themselves in the given sequence, while the **hidden states** would be the POS tags for the words. The **transition probabilities** would be somewhat like P(verb | noun) that is, what is the probability of the current word having a tag of verb given that the previous tag was a noun. **Emission probabilities** would be like P(that | noun) or P(good | adjective), which denote the probability that we observed the word, say "good", given that the tag is an adjective.
Now suppose you have the following training corpus, where each word is annotated with its POS tag:

That/(**conjunction**) that/(**noun**) is/(**verb**), is/(**verb**).
That/(**conjunction**) that/(**noun**) is/(**verb**) not/(**noun**), is/(**verb**) not/(**verb**).
Is/(**verb**) that/(**verb**) it/(**noun**)?
Is/(**verb**) it/(**noun**) that/(**noun**) good/(**adjective**)?

For each of these following tasks, write down your calculations, or provide the code that you wrote to compute the answer.

(a) Ignoring capitalization and punctuations, there should be a tag set of size 4 and a lexicon of size 5. Write out the corresponding sets of hidden states and observed states. Then define the HMM of this process by giving: 1) the initial probability vector; 2) the transition probability matrix; 3) the emission probability matrix.
All these probabilities should be learned from the training data. To avoid very sparse empirical distributions, we'll apply a technique called **add-one smoothing**: for each entry in a k-class multinomial distribution with N trials, the smoothed version of MLE would be:

$$p_i = \frac{N_i + 1}{N + k}$$

where N_i is the number of observations from class *i*.

Hidden states: {conjunction, noun, verb, adjective} or

$$X = \{C, N, V, A\}$$

Observed states: {that, is, not, it, good}

1) Initial probability vector:

$$\boldsymbol{\pi} = \begin{bmatrix} \frac{2+1}{17+4} & \frac{6+1}{17+4} & \frac{8+1}{17+4} & \frac{1+1}{17+4} \end{bmatrix} = \begin{bmatrix} 1/7 & 1/3 & 3/7 & 2/21 \end{bmatrix}$$

2) Transition probability matrix:

$$\mathrm{P}(C|C) = \mathrm{P}(V|C) = \mathrm{P}(A|C) = \frac{0+1}{2+4} = 1/6$$

$$\mathrm{P}(N|C) = \frac{2+1}{2+4} = 1/2$$

$$\mathrm{P}(C|N) = \frac{0+1}{6+4} = 1/10$$

$$\mathrm{P}(N|N) = \mathrm{P}(A|N) = \frac{1+1}{6+4} = 1/5$$

$$\mathrm{P}(V|N) = \frac{4+1}{6+4} = 1/2$$

$$\mathrm{P}(C|V) = \frac{1+1}{8+4} = 1/6$$

$$\mathrm{P}(N|V) = \frac{3+1}{8+4} = 1/3$$

$$\mathrm{P}(V|V) = \frac{4+1}{8+4} = 5/12$$

$$\mathrm{P}(A|V) = \frac{0+1}{8+4} = 1/12$$

$$\mathrm{P}(C|A) = \mathrm{P}(N|A) = \mathrm{P}(V|A) = \mathrm{P}(A|A) = \frac{0+1}{0+4} = 1/4$$

Therefore,

$$A = \begin{bmatrix} 1/6 & 1/2 & 1/6 & 1/6 \\ 1/10 & 1/5 & 1/2 & 1/5 \\ 1/6 & 1/3 & 5/12 & 1/12 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

3) Emission probability matrix:

$$B = \begin{array}{c} \\ \text{C} \\ \text{N} \\ \text{V} \\ \text{A} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{that} & \text{is} & \text{not} & \text{it} & \text{good} \end{array} \\ \begin{bmatrix} \frac{2+1}{2+5} & \frac{0+1}{2+5} & \frac{0+1}{2+5} & \frac{0+1}{2+5} & \frac{0+1}{2+5} \\ \frac{3+1}{6+5} & \frac{0+1}{6+5} & \frac{1+1}{6+5} & \frac{2+1}{6+5} & \frac{0+1}{6+5} \\ \frac{1+1}{8+5} & \frac{6+1}{8+5} & \frac{1+1}{8+5} & \frac{0+1}{8+5} & \frac{0+1}{8+5} \\ \frac{0+1}{1+5} & \frac{0+1}{1+5} & \frac{0+1}{1+5} & \frac{0+1}{1+5} & \frac{1+1}{1+5} \end{bmatrix} \end{array}$$

$$B = \begin{bmatrix} 3/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 4/11 & 1/11 & 2/11 & 3/11 & 1/11 \\ 2/13 & 7/13 & 2/13 & 1/13 & 1/13 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/3 \end{bmatrix}$$

(b) What is the probability of the observed sentence “Not that good”?

This is a likelihood problem. We will use the forward algorithm.

Let

$$\theta = (\pi, A, B)$$

Then,

$$\mathrm{P}(\text{"Not that good"}|\theta) = \sum_{i=1}^{|X|} \mathrm{P}(\text{"Not that good"}, X_{t=3} = x_i|\theta) = \sum_{i=1}^{|X|} \alpha_i(3)$$

We will compute this using a trellis

$$\begin{bmatrix} \alpha_1(1) & \alpha_1(2) & \alpha_1(3) \\ \alpha_2(1) & \alpha_2(2) & \alpha_2(3) \\ \alpha_3(1) & \alpha_3(2) & \alpha_3(3) \\ \alpha_4(1) & \alpha_4(2) & \alpha_4(3) \end{bmatrix} = \begin{bmatrix} \pi_1 b_1(\text{"not"}) & \sum_{i=1}^{|X|} \alpha_i(1) a_{i1} b_1(\text{"that"}) & \alpha_1(3) \\ \pi_2 b_2(\text{"not"}) & \sum_{i=1}^{|X|} \alpha_i(1) a_{i2} b_2(\text{"that"}) & \alpha_2(3) \\ \pi_3 b_3(\text{"not"}) & \sum_{i=1}^{|X|} \alpha_i(1) a_{i3} b_3(\text{"that"}) & \alpha_3(3) \\ \pi_4 b_4(\text{"not"}) & \sum_{i=1}^{|X|} \alpha_i(1) a_{i4} b_4(\text{"that"}) & \alpha_4(3) \end{bmatrix}$$

$$= \begin{bmatrix} 1/49 & 30799/2942940 & \sum_{i=1}^{|X|} \alpha_i(2) a_{i1} b_1(\text{"good"}) \\ 2/33 & 60883/3468465 & \sum_{i=1}^{|X|} \alpha_i(2) a_{i2} b_2(\text{"good"}) \\ 6/91 & 16433/1639638 & \sum_{i=1}^{|X|} \alpha_i(2) a_{i3} b_3(\text{"good"}) \\ 1/63 & 31513/7567560 & \sum_{i=1}^{|X|} \alpha_i(2) a_{i4} b_4(\text{"good"}) \end{bmatrix}$$

$$= \begin{bmatrix} 1/49 & 30799/2942940 & 34851227/39278439200 \\ 2/33 & 60883/3468465 & 1988499113/1666528063200 \\ 6/91 & 16433/1639638 & 158955547/131302211040 \\ 1/63 & 31513/7567560 & 1080385973/454507653600 \end{bmatrix}$$

Therefore,

$$\mathrm{P}(\text{"Not that good"}|\theta) = \sum_{i=1}^{|X|} \alpha_i(3)$$

$$= \frac{34851227}{39278439200} + \frac{1988499113}{1666528063200} + \frac{158955547}{131302211040} + \frac{1080385973}{454507653600}$$

$$= \frac{1289394674759}{227481080626800} \approx 0.006$$

(c) What is the probability of that, if there's another word after "good" in the previous sequence "Not that good", the 4th word has POS tag as a noun?

Want

$$\mathrm{P}(X_4 = x_2 = N|E_{1:3} = \text{"Not that good"})$$

This is a **prediction problem**, the algorithm for which is

$$\begin{aligned} &\text{for } m = 0 : k-1, \\ &\mathrm{P}(X_{t+m+1}|E_{1:t}) = \sum_{j=1}^{|X|} \mathrm{P}(X_{t+m+1}|X_{t+m} = x_j)\mathrm{P}(X_{t+m} = x_j|E_{1:t}) \\ &\text{Save computation for next iteration} \end{aligned}$$

where the term

$$\mathrm{P}(X_{t+m} = x_j|E_{1:t})$$

is obtained from solving a **filtering problem** (for the algorithm's first iteration).

The filtering problem is

$$\mathrm{P}(X_t = x_i | E_{1:t}) = \mathrm{P}(X_t = x_i | E_{1:t}, \theta) = \frac{\mathrm{P}(X_t = x_i, E_{1:t}|\theta)}{\mathrm{P}(E_{1:t}|\theta)} = \frac{\alpha_i(t)}{\mathrm{P}(E_{1:t}|\theta)}$$

where the denominator is the result of the **likelihood problem** from part (b).

Therefore,

$$\mathrm{P}(X_4 = N | E_{1:t} = \text{"Not that good"})$$
$$= \sum_{j=1}^{|X|} \mathrm{P}(X_4 = x_2 | X_3 = x_j)\mathrm{P}(X_3 = x_j | E_{1:t} = \text{"Not that good"})$$
$$= \left(\frac{1289394674759}{227481080626800}\right)^{-1} \sum_{j=1}^{|X|} a_{j2}\alpha_j(3)$$
$$= \frac{15287463585371}{51575786990360} \approx 0.296$$

(d) What is the most likely sequence of POS tags for the sentence in (b)?

Want to find

$$X^*_{1:3} = \mathrm{argmax}_{X_{1:3}} \mathrm{P}(X_{1:t} | E_{1:3} = \text{"Not that good"})$$

This is a **decoding problem**. We will therefore use the Viterbi algorithm (trellis like in forward algorithm, but cell values are maxes instead of sums):

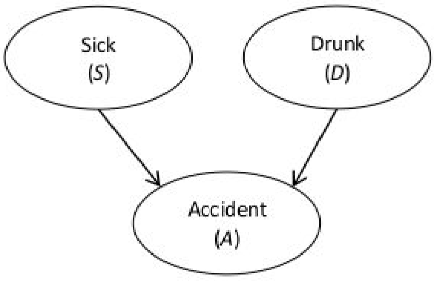
$$\begin{bmatrix} \delta_1(1) & \delta_1(2) & \delta_1(3) \\ \delta_2(1) & \delta_2(2) & \delta_2(3) \\ \delta_3(1) & \delta_3(2) & \delta_3(3) \\ \delta_4(1) & \delta_4(2) & \delta_4(3) \end{bmatrix} = \begin{bmatrix} \pi_1 b_1(\text{"not"}) & \max_i \delta_i(1) a_{i1} b_1(\text{"that"}) & \delta_1(3) \\ \pi_2 b_2(\text{"not"}) & \max_i \delta_i(1) a_{i2} b_2(\text{"that"}) & \delta_2(3) \\ \pi_3 b_3(\text{"not"}) & \max_i \delta_i(1) a_{i3} b_3(\text{"that"}) & \delta_3(3) \\ \pi_4 b_4(\text{"not"}) & \max_i \delta_i(1) a_{i4} b_4(\text{"that"}) & \delta_4(3) \end{bmatrix}$$
$$= \begin{bmatrix} 1/49 & 3/637 & \max_i \delta_i(2) a_{i1} b_1(\text{"good"}) \\ 2/33 & 8/1001 & \max_i \delta_i(2) a_{i2} b_2(\text{"good"}) \\ 6/91 & 2/429 & \max_i \delta_i(2) a_{i3} b_3(\text{"good"}) \\ 1/63 & 1/495 & \max_i \delta_i(2) a_{i4} b_4(\text{"good"}) \end{bmatrix}$$
$$= \begin{bmatrix} 1/49 & 3/637 & 1/8918 \\ 2/33 & 8/1001 & 3/14014 \\ 6/91 & 2/429 & 4/13013 \\ 1/63 & 1/495 & 1/3822 \end{bmatrix}$$

We will now start backtracking to find $X^*_{1:3}$:
The max value in column 3 is $\delta_4(3) = 1/3822$. Therefore, $X^*_3 = x_4 = A = \textit{adjective}$.
$\delta_4(3)$ was maximized by choosing $\delta_1(2)$. Therefore, $X^*_2 = x_1 = C = \textit{conjunction}$.
$\delta_1(2)$ was maximized by choosing $\delta_3(1)$. Therefore, $X^*_1 = x_3 = V = \textit{verb}$.

Therefore, the most likely sequence of POS tags for the sentence "not that good" is

$$X^*_{1:3} = \text{verb, conjunction, adjective}$$

Question 2: Utility



S (F)	S (T)
0.75	0.25

D (F)	D (T)
0.85	0.15

S	D	A (T)
F	F	0.05
F	T	0.1
T	F	0.1
T	T	0.75

Consider the Bayes Net shown here, with all Bernoulli variables, which models driving accidents. Having an accident has a utility of -100 if the car was insured, but has a utility of -500 if the car was not insured (or the insurance claim is denied). Having insurance when there is no accident has a utility of -20, and not having insurance and an accident has a utility of 0.

Use the principle of Maximum Expected Utility and Value of Information to answer the following questions. For parts a)-c) assume the insurance company pays for 100% of the cases.

a) Given no information on whether the driver will be driving in a sick or drunk state, how much should they pay to insure the car?

From the problem statement we get

	A(T)	A(F)
I(T) (insurance)	-100	-20
I(F) (no insurance)	-500	0

$$\begin{aligned}
\mathrm{EU}(I(T)) &= -100\mathrm{P}(A(T)) - 20\mathrm{P}(A(F)) \\
&= -100\sum_{S,D}\mathrm{P}(A(T),S,D) - 20\sum_{S,D}\mathrm{P}(A(F),S,D) \\
&= -100\sum_{S,D}\mathrm{P}(A(T)|S,D)\mathrm{P}(S)\mathrm{P}(D) - 20\sum_{S,D}\mathrm{P}(A(F)|S,D)\mathrm{P}(S)\mathrm{P}(D) \\
&= -27.4
\end{aligned}$$

Similarly,

$$\mathrm{EU}(I(F)) = -500\mathrm{P}(A(T)) - 0\mathrm{P}(A(F)) = -46.25$$

Therefore, the amount “they” should “pay to insure the car” is (in quotes because this question makes no sense)

$$\mathrm{EU}(I(T)) - \mathrm{EU}(I(F)) = 18.85$$

b) How much should they pay for insurance if you know for certain that they will drive both sick (S=True) and **not drunk** (D=False)?

$$\begin{aligned}\mathrm{EU}(I(T)) &= -100\mathrm{P}(A(T), S(T), D(F)) - 20\mathrm{P}(A(F), S(T), D(F)) \\ &= -100(0.1) - 20(0.9) \\ &= -28\end{aligned}$$

$$\begin{aligned}\mathrm{EU}(I(T)) &= -500\mathrm{P}(A(T), S(T), D(F)) - 0\mathrm{P}(A(F), S(T), D(F)) \\ &= -50\end{aligned}$$

$$\mathrm{EU}(I(T)) - \mathrm{EU}(I(F)) = 22$$

c) How much should they pay for insurance if you know that they will drive drunk (D=True)?

$$\begin{aligned}\mathrm{EU}(I(T)) &= -100\mathrm{P}(D(T))\sum_S \mathrm{P}(A(T)|S, D(T))\mathrm{P}(S) \\ &-20\mathrm{P}(D(T))\sum_S \mathrm{P}(A(F)|S, D(T))\mathrm{P}(S) \\ &= -6.15\end{aligned}$$

$$\mathrm{EU}(I(F)) = -315/16 \approx -19.7$$

$$\mathrm{EU}(I(T)) - \mathrm{EU}(I(F)) = 1083/80 \approx 13.5$$

d) A company is offering cheaper insurance but has a reputation of rejecting 10% of insurance claims. How much should they charge for this insurance, to make it competitive with the insurance offered by the more reliable company? (Hint: Set the cost of the new insurance to have the same MEU as the other insurance.)

Question 4: Bandits

Consider the following 6-armed bandit problem. The initial value estimates of the arms are given by $Q = \{1, 2, 2, 1, 0, 3\}$, and the actions are represented by $A = \{1, 2, 3, 4, 5, 6\}$. Suppose we observe that each lever is played in turn: (from lever 1 to lever 6, and then start from lever 1 again):

$$A_t = ((t-1) \bmod 6) + 1 \qquad (1)$$

We also observe that the rewards R_t seem to fit the following function:

$$R_t = 2\cos\left[\frac{\pi}{6}(t-1)\right] \qquad (2)$$

So, the first two action-reward pairs are $A_1 = 1, R_1 = 2,$ and $A_2 = 2, R_2 = \sqrt{3}$.

a) Show the estimated Q values from $t = 1$ to $t = 12$ of the trajectory using the average of the observed rewards, where available. Do not consider the initial estimates as samples.

t

1: $Q_1(1) = R_1 = 2$

2: $Q_1(2) = R_2 = \sqrt{3}$

3: $Q_1(3) = R_3 = 1$

4: $Q_1(4) = R_4 = 0$

5: $Q_1(5) = R_5 = -1$

6: $Q_1(6) = R_6 = -\sqrt{3}$

7: $Q_2(1) = Q_1(1) + (R_7 - Q_1(1))/2 = 2 + (-2 - 2)/2 = 0$

8: $Q_2(2) = Q_1(2) + (R_8 - Q_1(2))/2 = \sqrt{3} + (-\sqrt{3} - \sqrt{3})/2 = 0$

9: $Q_2(3) = Q_1(3) + (R_9 - Q_1(3))/2 = 1 + (-1 - 1)/2 = 0$

10: $Q_2(4) = Q_1(4) + (R_{10} - Q_1(4))/2 = 0 + (0 - 0)/2 = 0$

11: $Q_2(5) = Q_1(5) + (R_{11} - Q_1(5))/2 = -1 + (1 - (-1))/2 = 0$

12: $Q_2(6) = Q_1(6) + (R_{12} - Q_1(6))/2 = -\sqrt{3} + (\sqrt{3} - (-\sqrt{3}))/2 = 0$

Therefore, at $t = 12$

$$Q = \{0, 0, 0, 0, 0, 0\}$$

b) It turns out the player was following an ε-greedy strategy, which just happened to coincide with the scheme described above in (1) for the first 12 time steps. For each time step t from 1 to 12, report whether it can be concluded with certainty that a random action was selected.

Assume we take the initial estimates as samples.

t	Chosen arm	Arm(s) with highest Q value	Certain that random arm was selected	Updated Q values
1:	1	6	yes	$\{1.5, 2, 2, 1, 0, 3\}$
2:	2	6	yes	$\{1.5, 1 + \sqrt{3}/2 \approx 1.87, 2, 1, 0, 3\}$
3:	3	6	yes	$\{1.5, 1.87, 1.5, 1, 0, 3\}$
4:	4	6	yes	$\{1.5, 1.87, 1.5, 0.5, 0, 3\}$
5:	5	6	yes	$\{1.5, 1.87, 1.5, 0.5, -0.5, 3\}$
6:	6	6	no	$\{1.5, 1.87, 1.5, 0.5, -0.5, (3 - \sqrt{3})/2 \approx 0.63\}$

7:	1	2	yes	$\{1/3 \approx 0.33, 1.87, 1.5, 0.5, -0.5, 0.63\}$
8:	2	2	no	$\{0.33, 2/3 \approx 0.67, 1.5, 0.5, -0.5, 0.63\}$
9:	3	3	no	$\{0.33, 0.67, 2/3 \approx 0.67, 0.5, -0.5, 0.63\}$
10:	4	2 and 3	yes	$\{0.33, 0.67, 0.67, 1/3 \approx 0.33, -0.5, 0.63\}$
11:	5	2 and 3	yes	$\{0.33, 0.67, 0.67, 0.33, 0, 0.63\}$
12:	6	2 and 3	yes	$\{0.33, 0.67, 0.67, 0.33, 0, 1\}$

c) Suppose now we continue to visit the levers iteratively as in (1), and that the observed rewards continue to fit the pattern established by (2). Is there a limiting expected reward $Q^*(a)$ for each action $a \in A$ as t approaches infinity? Justify your answer.

For all actions a, the limiting expected reward $Q^*(a)$ is 0.

An action is performed once every 6 iterations, in other words, at $t, t+6, t+(2)(6), ..., t+6k$. We can prove that for each action, its rewards are consecutive opposite values of each other.

$$R_{t+6} = 2\cos\left[\frac{\pi}{6}((t+6)-1)\right] = 2\cos\left[\frac{\pi}{6}(t-1)+\pi\right] = -2\cos\left[\frac{\pi}{6}(t-1)\right]$$
$$= -R_t$$

Therefore, as t approaches infinity, the average reward or Q value of any action approaches 0.