

Homework 4*Handed Out: 11/06/2024**Due: 11/20/2024*

- **Your Name:** *Alan Wu*
- **Your PennKey:** *alanlwu*
- **Your PennID:** *41855518*

1 Multiple Choice & Written Questions

1. (a) If we always try to sample the word with the greatest probability, then we first will sample the word with the greatest probability, then sample the next word with the greatest joint probability.

For the first word, we have to select the word where $P(w_1|\text{Bob})$ is highest, which is going to be the word loves, with a probability of 0.5.

For the second word, we have to select the word where $P(w_2|\text{Bob, loves})$ is highest, which is the word cookie, with a probability of 0.4.

Therefore, we are left with the sentence "Bob loves cookie".

- (b) We need to estimate the log likelihood of the sentences "Bob loves cookie" and "Bob hates cookie".

We know the formula for the log likelihood estimate is going to be the sum of the log joint probabilities of the words in the sentence.

Bob loves cookie

$$\begin{aligned}\ln P(\text{Bob loves cookie}) &= \ln P(\text{loves}|\text{Bob}) + \ln P(\text{cookie}|\text{Bob, loves}) \\ &= \ln 0.50 + \ln 0.40 \\ &= -1.609\end{aligned}$$

The log likelihood estimate of the sentence "Bob loves cookie" is -1.609.

Bob hates cookie

$$\begin{aligned}
\ln P(\text{Bob hates cookie}) &= \ln P(\text{hates}|\text{Bob}) + \ln P(\text{cookie}|\text{Bob, hates}) \\
&= \ln 0.40 + \ln 0.20 \\
&= -2.526
\end{aligned}$$

The log likelihood estimate of the sentence "Bob hates cookie" is -2.526.

(c)

- (d) We have to compute the log likelihood of the remaining sentences that we didn't compute earlier: "Bob loves Bob" and "Bob hates cherry"

We already know that the log likelihood of the sentence "Bob loves cookie" is -1.609 and the log likelihood of the sentence "Bob hates cookie" is -2.526. The higher log likelihood indicates higher chance of occurring.

Bob loves Bob

$$\begin{aligned}
\ln P(\text{Bob loves Bob}) &= \ln P(\text{loves}|\text{Bob}) + \ln P(\text{Bob}|\text{Bob, loves}) \\
&= \ln 0.50 + \ln 0.25 \\
&= -2.079
\end{aligned}$$

The log likelihood estimate of the sentence "Bob loves Bob" is -2.079.

Bob hates cherry

$$\begin{aligned}
\ln P(\text{Bob hates cherry}) &= \ln P(\text{hates}|\text{Bob}) + \ln P(\text{cherry}|\text{Bob, hates}) \\
&= \ln 0.40 + \ln 0.70 \\
&= -1.272
\end{aligned}$$

The log likelihood of the sentence "Bob hates cherry" is -1.272.

The top two likelihood of beam search are "Bob loves cookie" and "Bob hates cherry".

2. (a) In order to compute the attention scores, we need to take the dot product of the hidden state of the input word and the hidden state value of the output word.

We are given the hidden state for the output word to be $s_1 = [0.5, 0.2, 0.4, 0.1]$

$$E^1 = [s_1^T h_1, s_1^T h_2, s_1^T h_3, s_1^T h_4]$$

Let us compute each term in the vector for h_1, h_2, h_3, h_4

h_1 :

$$\begin{aligned}s_1^T h_1 &= 0.7 \cdot 0.5 + 0.2 \cdot 0.2 + 0.3 \cdot 0.4 + 0.1 \cdot 0.1 \\ &= 0.35 + 0.04 + 0.12 + 0.01 \\ &= 0.52\end{aligned}$$

h_2 :

$$\begin{aligned}s_1^T h_2 &= 0.2 \cdot 0.5 + 0.7 \cdot 0.2 + 0.3 \cdot 0.4 + 0.1 \cdot 0.1 \\ &= 0.1 + 0.14 + 0.12 + 0.01 \\ &= 0.37\end{aligned}$$

h_3 :

$$\begin{aligned}s_1^T h_3 &= 0.0 \cdot 0.5 + 0.6 \cdot 0.2 + 0.4 \cdot 0.4 + 0.3 \cdot 0.1 \\ &= 0.0 + 0.12 + 0.16 + 0.03 \\ &= 0.31\end{aligned}$$

h_4 :

$$\begin{aligned}s_1^T h_4 &= 0.1 \cdot 0.5 + 0.1 \cdot 0.2 + 0.0 \cdot 0.4 + 0.9 \cdot 0.1 \\ &= 0.05 + 0.02 + 0.0 + 0.09 \\ &= 0.16\end{aligned}$$

Therefore, the dot-product attention scores $E^1 = [0.52, 0.37, 0.31, 0.16]$

- (b) To derive the attention distribution α^1 , we need to apply the softmax function to the attention scores E^1 .

Let us compute the exponential of all the values in E^1 and sum them up.

$$\begin{aligned}\text{sum} &= e^{0.52} + e^{0.37} + e^{0.31} + e^{0.16} \\ &= 5.667\end{aligned}$$

Now, we can compute the attention distribution α^1 by dividing each value in E^1 by the sum.

$$\begin{aligned}
\alpha^1 &= \left[\frac{e^{0.52}}{5.667}, \frac{e^{0.37}}{5.667}, \frac{e^{0.31}}{5.667}, \frac{e^{0.16}}{5.667} \right] \\
&= \left[\frac{1.682}{5.667}, \frac{1.448}{5.667}, \frac{1.363}{5.667}, \frac{1.174}{5.667} \right] \\
&= [0.297, 0.256, 0.241, 0.207]
\end{aligned}$$

The attention distribution α^1 is $[0.297, 0.256, 0.241, 0.207]$

- (c) To compute the attention output a^1 , we need to take the dot product of the attention distribution α^1 and the hidden states of the input words.

Let us compute each term in the attention output for h_1, h_2, h_3, h_4

h_1 :

$$\begin{aligned}
\alpha^{1T}h_1 &= 0.7 \cdot 0.297 + 0.2 \cdot 0.256 + 0.3 \cdot 0.241 + 0.1 \cdot 0.207 \\
&= 0.2079 + 0.0512 + 0.0723 + 0.0207 \\
&= 0.3521
\end{aligned}$$

h_2 :

$$\begin{aligned}
\alpha^{1T}h_2 &= 0.2 \cdot 0.297 + 0.7 \cdot 0.256 + 0.3 \cdot 0.241 + 0.1 \cdot 0.207 \\
&= 0.0594 + 0.1792 + 0.0723 + 0.0207 \\
&= 0.3316
\end{aligned}$$

h_3 :

$$\begin{aligned}
\alpha^{1T}h_3 &= 0.0 \cdot 0.297 + 0.6 \cdot 0.256 + 0.4 \cdot 0.241 + 0.3 \cdot 0.207 \\
&= 0.0 + 0.1536 + 0.0964 + 0.0621 \\
&= 0.3121
\end{aligned}$$

h_4 :

$$\begin{aligned}
\alpha^{1T}h_4 &= 0.1 \cdot 0.297 + 0.1 \cdot 0.255 + 0.0 \cdot 0.241 + 0.9 \cdot 0.207 \\
&= 0.0297 + 0.0256 + 0.0 + 0.1863 \\
&= 0.2415
\end{aligned}$$

Therefore, the attention output $a^1 = [0.3519, 0.3309, 0.3115, 0.2415]$

3. (a) The optimal action would be to move to the left to E towards A.

When we compare the reward of the two actions, we notice that the discount factor is 1.0, meaning that the reward is not discounted. We can compare the reward of moving towards A and moving towards H weighted the same. This compares 100 vs 30. Therefore, the best action to take with discount factor 1 is move to the left.

- (b) The optimal action would be to move to the right towards H.

In this case, we are comparing the discounted reward of moving towards A and moving towards H (the terminal states)

A is 5 steps away from F, and H is 2 steps away from F. The discount factor is 0.5, thus

$$\text{Discounted Reward of moving towards A} = 100 \cdot 0.5^5 = 3.125$$

$$\text{Discounted Reward of moving towards H} = 30 \cdot 0.5^2 = 7.5$$

Since we want to maximize the discounted reward, we will move towards H and go to G.

- (c) We can compute the discount factor for both directions of being equally likely by using the previous discounted rewards.

We know that starting at F, A is 5 steps away and H is 2 steps away. The discount factor is 0.5. We develop the following expression:

$$\begin{aligned} 30 \cdot \gamma^2 &= 100 \cdot \gamma^5 \\ \gamma^3 &= 0.3 \\ \gamma &= 0.3^{1/3} \\ &= 0.67 \end{aligned}$$

Therefore, the discount factor that makes moving towards A and moving towards H equally likely is 0.67.

- (d) bruh
 - (e) bruh
 - (f) bruh
4. (a)
- (b)
- (c)
5. (a)

2 Coding Questions Written Answers