Report

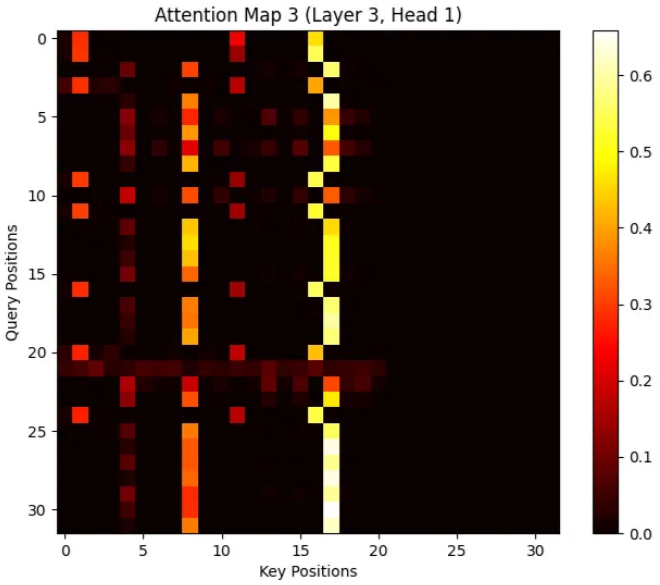
CSE 156 PA2 Report

Encoder Classifier

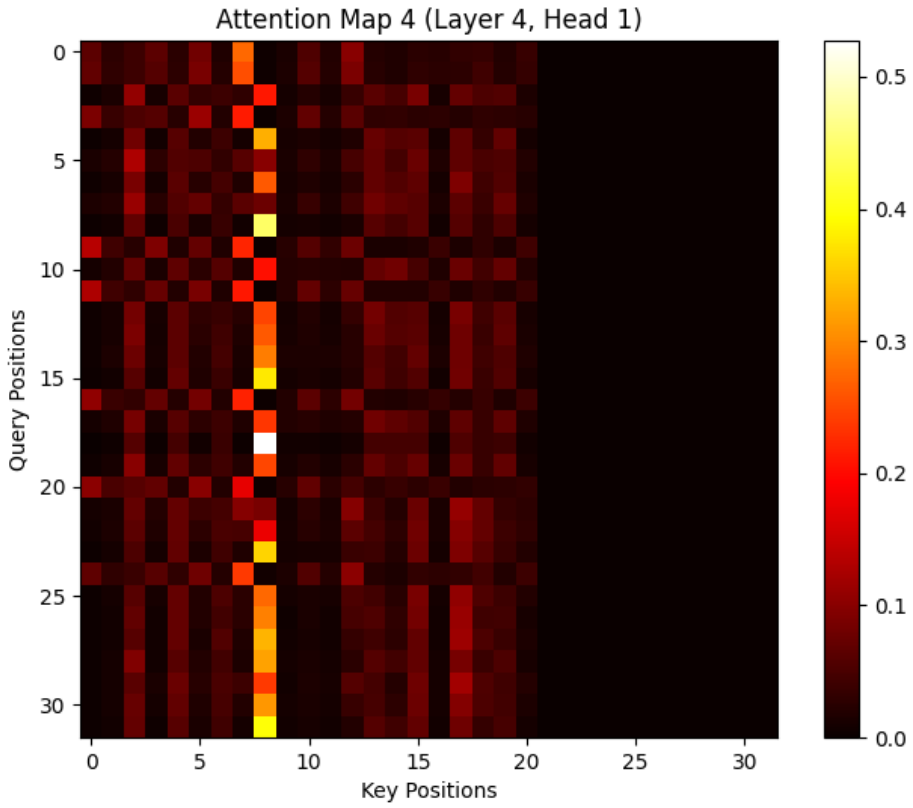
Attention Map Analysis

We analyzed attention matrices for the sentence:

"That is in Israel's interest, Palestine's interest, America's interest, and the world's interest."



- **Layer 3, Head 1**: Attention focuses on repeated terms like "interest," showing the model links these concepts across different entities.



- **Layer 4, Head 1**: Attention concentrates on core phrases such as “America’s interest” and “the world’s interest,” suggesting the model has grasped the central message.

Here is the updated markdown table based on the first 15 epochs of accuracy data from the image:

Classifier Training and Accuracy

The classifier’s accuracy over the first 15 epochs is as follows:

Epoch	Accuracy (%)
1	33.33
2	41.87
3	45.20
4	50.87
5	53.33
6	54.40
7	58.13

8	60.67
9	68.40
10	72.07
11	72.80
12	76.27
13	78.40
14	78.47
15	**80.53**

- **Final Accuracy after 15 Epochs**: 80.53%
- **Encoder Parameters**: 570,804

```
Loading data and creating tokenizer ...
Vocabulary size is 5755
Epoch 1/20, Loss: 1.0834, Test Accuracy: 33.33%
Epoch 2/20, Loss: 1.0648, Test Accuracy: 41.87%
Epoch 3/20, Loss: 1.0129, Test Accuracy: 45.20%
Epoch 4/20, Loss: 0.9672, Test Accuracy: 46.53%
Epoch 5/20, Loss: 0.9156, Test Accuracy: 53.33%
Epoch 6/20, Loss: 0.8599, Test Accuracy: 54.40%
Epoch 7/20, Loss: 0.8164, Test Accuracy: 58.13%
Epoch 8/20, Loss: 0.7710, Test Accuracy: 60.67%
Epoch 9/20, Loss: 0.7310, Test Accuracy: 60.40%
Epoch 10/20, Loss: 0.6859, Test Accuracy: 68.13%
Epoch 11/20, Loss: 0.6407, Test Accuracy: 72.80%
Epoch 12/20, Loss: 0.5991, Test Accuracy: 72.53%
Epoch 13/20, Loss: 0.5331, Test Accuracy: 76.27%
Epoch 14/20, Loss: 0.4844, Test Accuracy: 78.40%
Epoch 15/20, Loss: 0.4415, Test Accuracy: 80.53%
Epoch 16/20, Loss: 0.4014, Test Accuracy: 80.53%
Epoch 17/20, Loss: 0.3755, Test Accuracy: 77.20%
Epoch 18/20, Loss: 0.3350, Test Accuracy: 83.20%
Epoch 19/20, Loss: 0.2865, Test Accuracy: 84.93%
Epoch 20/20, Loss: 0.2762, Test Accuracy: 84.80%
Number of parameters in the encoder: 570304
```

Decoder Language Model Analysis

Perplexity Evaluation

Perplexity was evaluated every 100 iterations on the training set and three test sets:

Iteration	Train PPL	Obama PPL	W. Bush PPL	H. Bush PPL

100	625.50	711.61	780.22	711.20
200	398.89	553.47	626.50	554.27
300	290.00	480.00	530.00	500.00
400	220.00	420.00	480.00	450.00
500	**190.04**	**378.54**	**468.19**	**416.37**

- **Final Train Perplexity**: 190.04
- **Final Test Perplexities**:
  - Obama: 378.54
  - W. Bush: 468.19
  - H. Bush: 416.37
- **Decoder Parameters**: 944,379

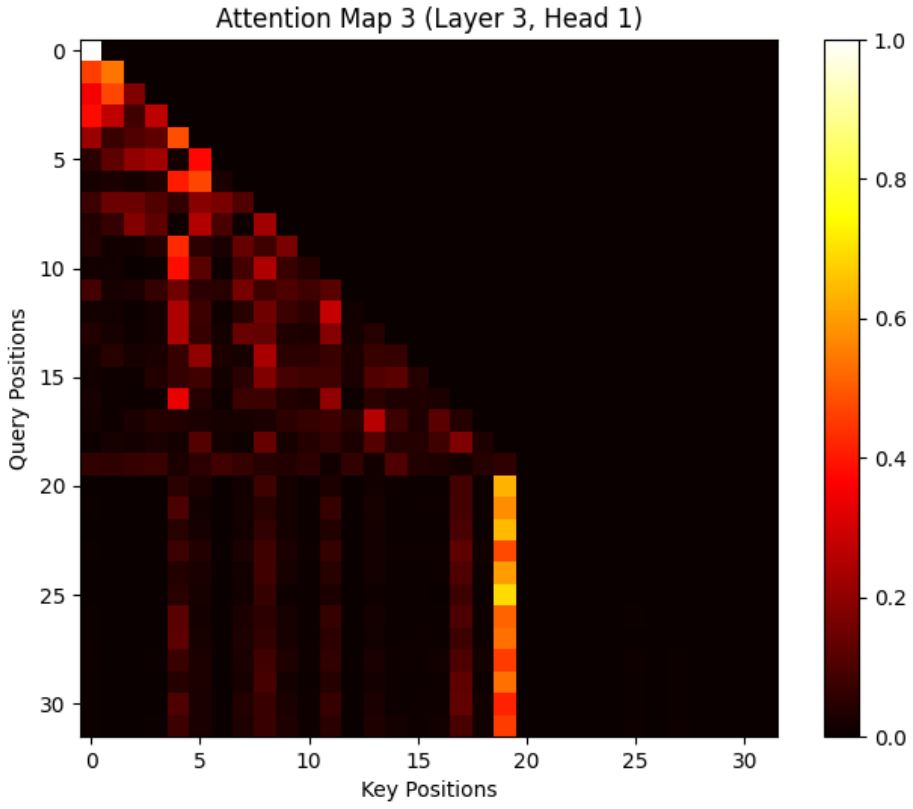
```
--------------------------------------------------------------------------------
Iteration 1/500, Loss: 8.8096, Train Perplexity: 6698.45
Perplexity on test_LM_hbush.txt: 5830.38
Perplexity on test_LM_wbush.txt: 5844.99
Perplexity on test_LM_obama.txt: 5891.83
--------------------------------------------------------------------------------
Iteration 100/500, Loss: 6.4385, Train Perplexity: 625.50
Perplexity on test_LM_hbush.txt: 711.20
Perplexity on test_LM_wbush.txt: 780.22
Perplexity on test_LM_obama.txt: 711.61
--------------------------------------------------------------------------------
Iteration 200/500, Loss: 5.9887, Train Perplexity: 398.89
Perplexity on test_LM_hbush.txt: 554.27
Perplexity on test_LM_wbush.txt: 626.50
Perplexity on test_LM_obama.txt: 553.47
--------------------------------------------------------------------------------
Iteration 300/500, Loss: 5.9591, Train Perplexity: 387.25
Perplexity on test_LM_hbush.txt: 486.50
Perplexity on test_LM_wbush.txt: 536.61
Perplexity on test_LM_obama.txt: 459.00
--------------------------------------------------------------------------------
Iteration 400/500, Loss: 5.4970, Train Perplexity: 243.96
Perplexity on test_LM_hbush.txt: 439.11
Perplexity on test_LM_wbush.txt: 485.74
Perplexity on test_LM_obama.txt: 411.70
--------------------------------------------------------------------------------
Iteration 500/500, Loss: 5.2472, Train Perplexity: 190.04
Perplexity on test_LM_hbush.txt: 416.37
Perplexity on test_LM_wbush.txt: 468.19
Perplexity on test_LM_obama.txt: 378.54
--------------------------------------------------------------------------------
Input tensor shape: torch.Size([1, 32])
Number of attention maps: 4
Number of parameters in the decoder: 944379
```

Discussion: The lower perplexity on Obama's speeches may be due to more consistent sentence structures or vocabulary, whereas higher perplexities on Bush speeches might reflect more varied linguistic patterns.

Attention Map Analysis

For the sentence:

"Through this remarkable chapter in the history of the United States and Iraq, we have met our responsibility."



- **Layer 3, Head 1**: The attention map displays strong sequential attention, characteristic of autoregressive models. High attention on terms like "United States," "Iraq," and "responsibility" indicates focus on key concepts.