CSC321 A3

March 2018

1 Part 1

1.1 Q1

I think it will not perform on long sequences, although the decoder gets to see all information about the input sequence, it receives information through the hidden vector. Whereas the hidden unit size is fixed and relatively small, when input sequence is too long, the single vector will lose information and is not a good representation of input sequence.

1.2 Q2

Qualitatively the results do not look very well, it seems that the model successfully learned to add 'ay' in all cases. Words of length less than or equal to 5 get better translation. It fails to distinguish between vowels such as e and o.

2 Part 2

2.1 Q1

The model will be augmenting mistakes learned from earlier stage. Because at test time, the model won't gain access to the true target, only predict based on the generated token from the model. So mistakes learned from training process will affect the testing process.

2.2 Q2

we are able to address this issue using a curriculum learning approach that alters the training process so that the the model will be able to learn faster from and correct its mistakes at the training time.

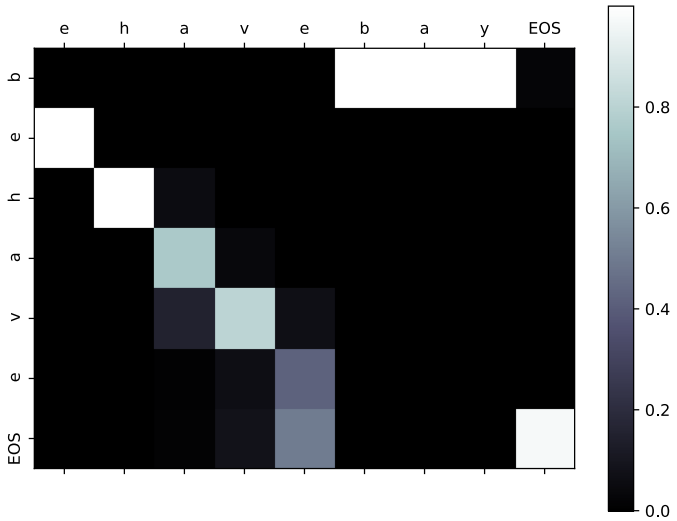

Figure 1: behave – ehavebay

3 Part 5

The model succeeds on words begins with a single consonant, two or more consonants and vowel on short words($\leq$11letters), because the model predicts one word a time.

The model fails on long words, might be due to gradient vanishing on long sentences.

The model also fails on word with '-', because it's relatively hard to predict what's next after a connection line, since '-' is not predictive to what's coming next. Also, due to the fact that the model has a long-term memory. When predicting on words separately the model predicts nicely, but not with a dash implies that the memory on word before the dash affects its predictability on words later.

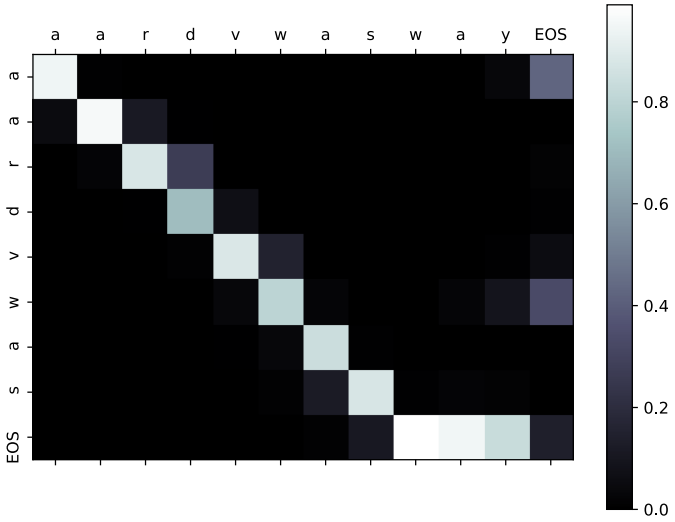

Figure 2: aardvwas – aardvwasway

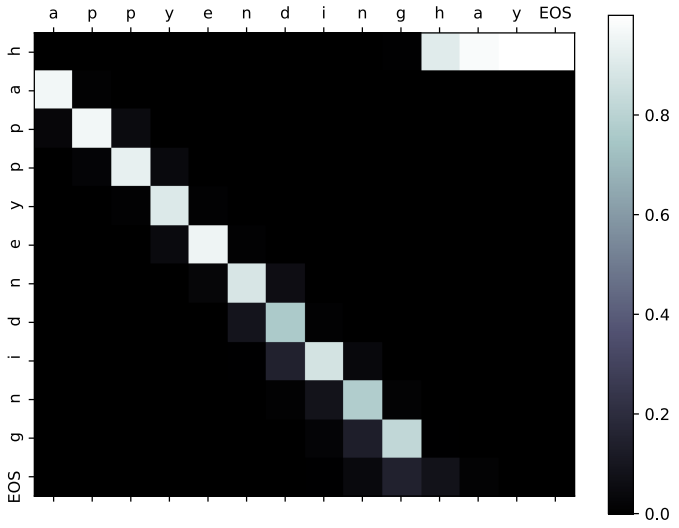

Figure 3: happyending – appyendinghay

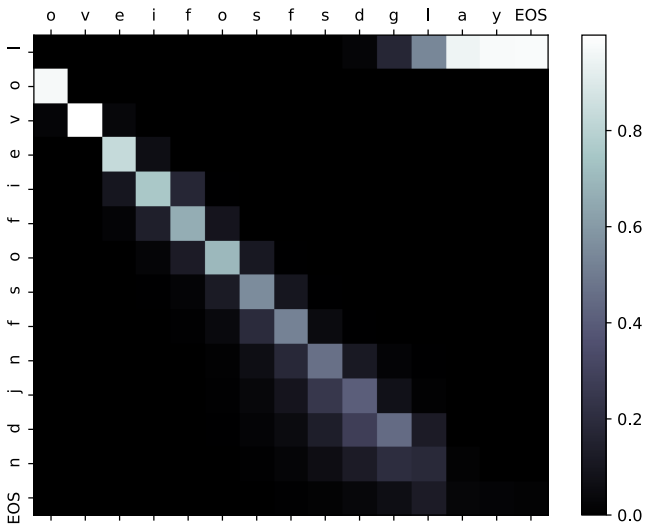

Figure 4: loveifosfnjdn – oveifosfsdglay

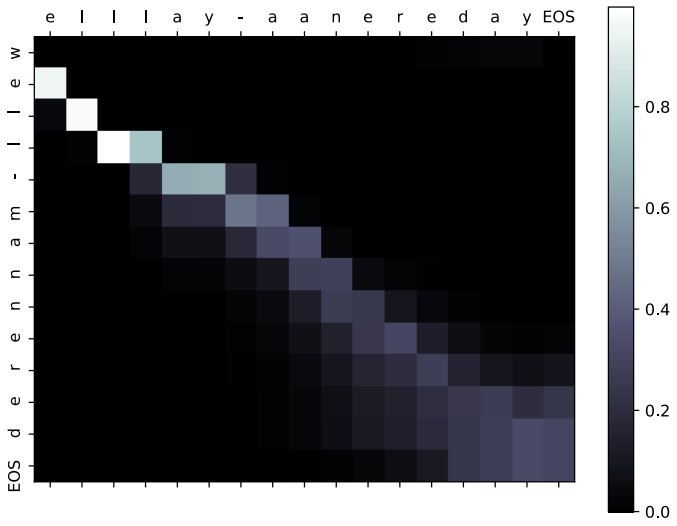

Figure 5: well-mannered – elllay-aanereday

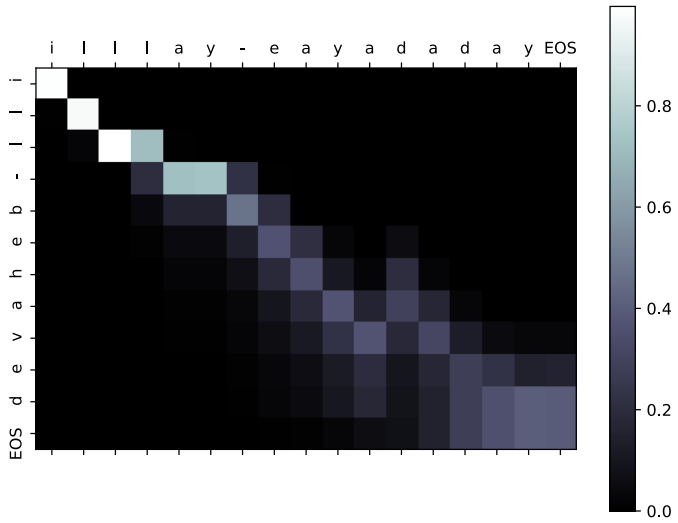

Figure 6: ill-behaved – illlay-eayadaday