CSC321 Assignment 3

Yifei Ai

Part 1

a

In my opinion, this architecture won't work well if the word is very long. Since personally I don't think using such 10 hidden units to store the whole information is a good idea, as soon as words get longer, the performance might get down very quickly.

b

The result looks very bad. The only correct answer I get is from the input words with length 2. The model does recognize we need to add 'ay' to the end of the word somehow. It also somehow know the different rules for word start with constant and word start with vowel even though it cannot get 100% accuracy on it. However, it had a bad performance on remembering word, so the inner character in the output is not correct, even some words' length doesn't match the length of input words.
The following for some result

```
roomba --> omcerway
concert --> orcortcay
hello --> erlehay
table --> adletay
ai --> aiway
bc --> mylay
aiyifei --> aigitebway
ac --> amway
io --> ioway
roomb --> oomeway
shfdsa --> assshay
shaft --> aftfhay
```

Part 2

a

When we train the model, we feed the decoder with ground-truth token instead of the output from the decoder. This may cause the problem that the decoder doesn't actually learn to predict character from the output it generates, instead, it just learn to predict the character from the true character, which means the weight for the decoder need to be justified in some sense.

b

First we can use some heuristic function to maintain several generated target sequences.
Second during the training process, we randomly decide using teach-forcing model or we use the previous token generated from decoder.(i.e. randomly decide using y_{t-1} or $\hat{y}_{t-1}$. This will somehow improve the model according to the paper.

Part 5

Non-compound words except very long word has a good performance generally. From the following and the words I tried, some very long words and words with dash have very bad performance, besides, some specific cases i.e. when 'a' and 'j' are together and some of them are very likely make mistakes when the decoder trying to translate.

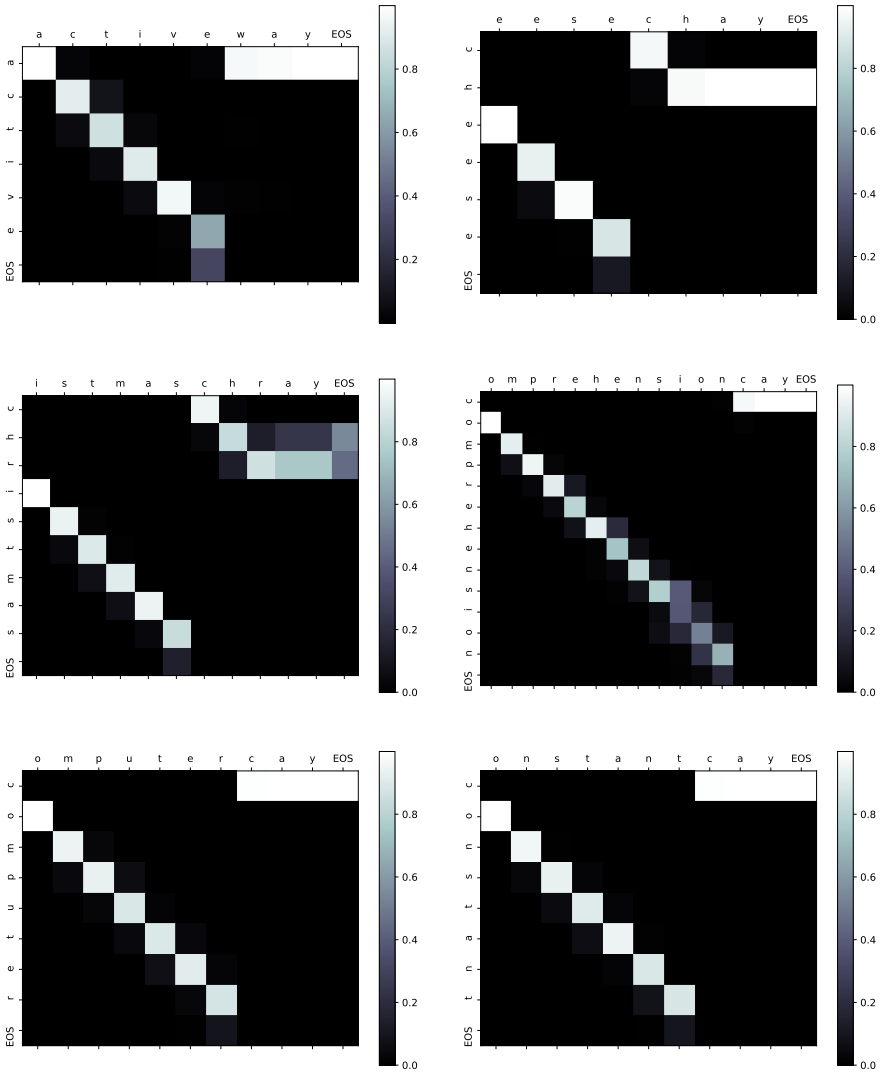

Figure 1: Successful Translation

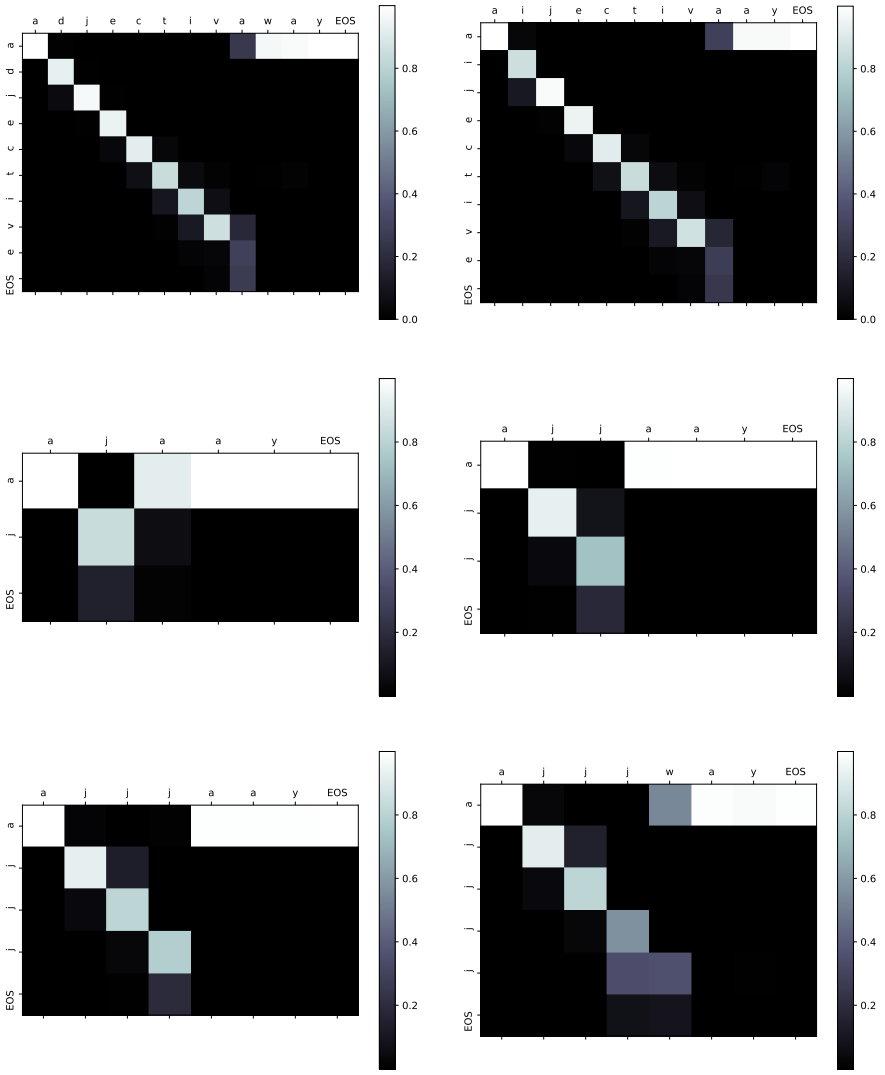

Figure 2: Failed Translation with no dash

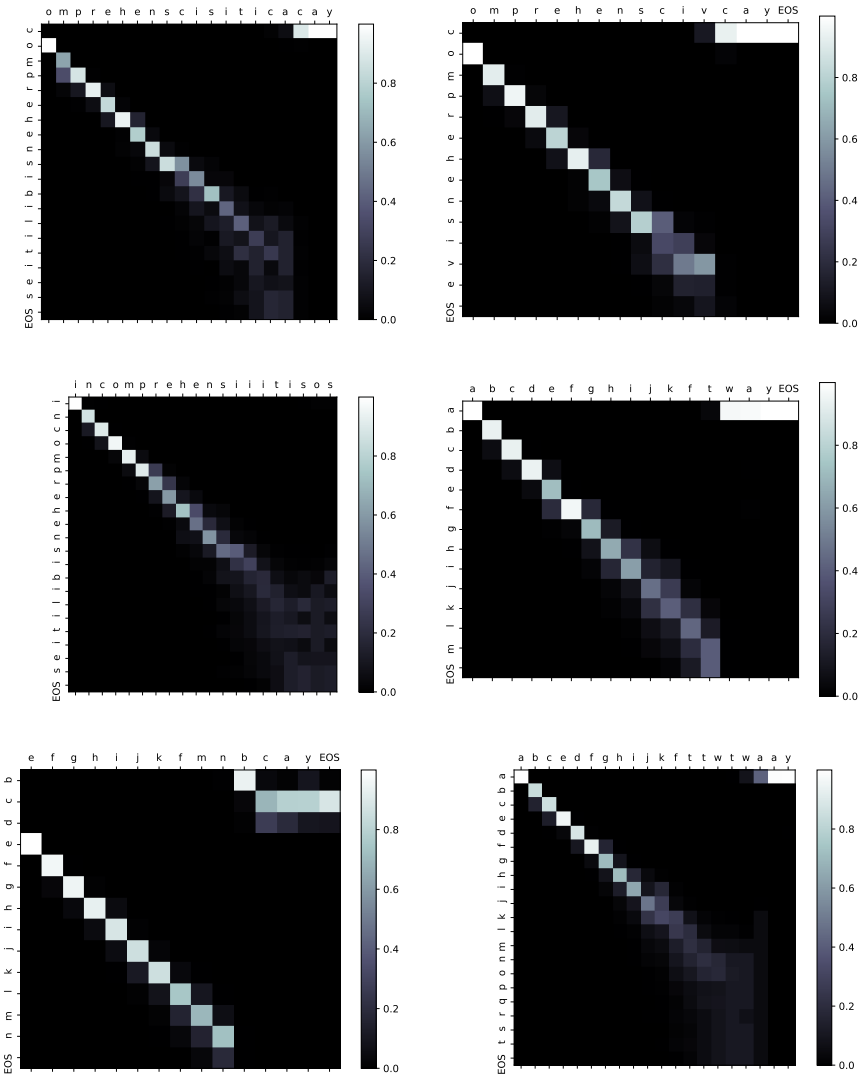

Figure 3: Failed Translation with long word

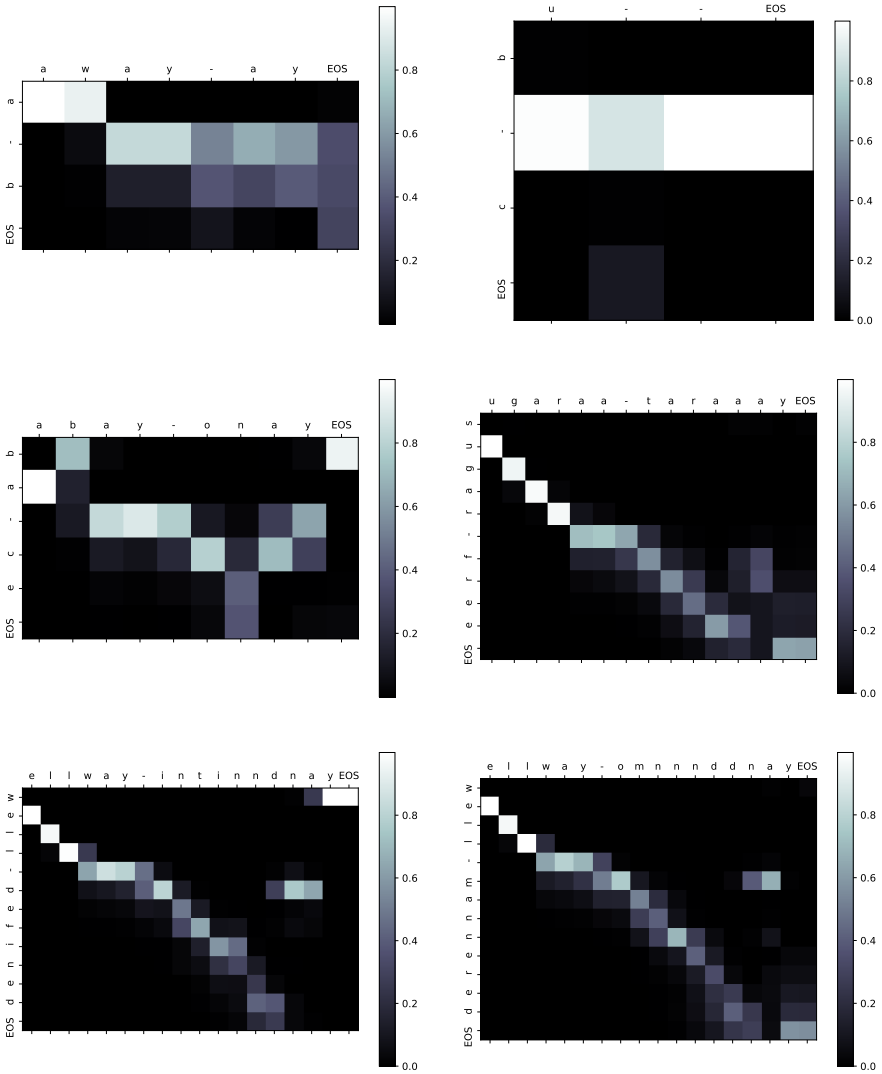

Figure 4: Failed Translation with words with dash