Emily Mulhall
COMP 550
Assignment 2
October 20, 2017

1 Question 1

$$\pi = <3/8,\ 3/8,\ 1/8,\ 1/8>$$

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

$$\mathbf{B} = \begin{bmatrix} 3/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/11 & 7/11 & 1/11 & 1/11 & 1/11 \\ 1/6 & 1/6 & 1/6 & 2/6 & 1/6 \\ 5/13 & 1/13 & 3/13 & 1/13 & 3/13 \end{bmatrix}$$

$Trellis =$

$(3/8) \times (1/7) = 0.0536$	$(0.0536) \times (1/6) \times (3/7) = 0.00383$	0.00128
0.0341	0.000873	0.000873
0.0208	0.001489	0.00298
0.0288	0.0103	0.00206

2 Question 2

The grammar rejects the following sentences: "*Je mangent le poisson", "* Les noirs chats mangent le poisson", "* La poisson mangent les chats". It accepts the following sentences: "Tu regardes la television", "Le chat mange le poisson", "Emily choisit le chat noir".

First, I will prove that the grammar rejects *Je mangent le poisson. The grammar will recognize Je as PRs1 and mangent as Vp3. S will go to PRs1 VPs1, and VPs1 will then go to Vs1 NP. Vs1 will go to regarde, mange, choisis and vends, and because none of these are mangent, the sentence will be rejected.

Next, for "* Les noirs chats mangent le poisson," the grammar will go from S to NPp VPp3. NPp will go to Dtp Npma. Npma will go to A2pm Npm, and A2pm will go to beux, jolis or jeunes, none of which are noirs. Thus, the sentence will be rejected.

Finally, for "* La poisson mangent les chats," the grammar will go from S to NPs VPs3. VPs3 will go to Vs3 NP, and Vs3 will go to regarde, mange, choisit, or vende, and because none of these are mangent the sentence will be rejected.

As for the sentence "Tu regardes la television," going bottom up the grammar will recognize the following constituents: PRs2 Vs2 Dtf Nsf. Dtf Nsf will combine to form NPs which will then be recognized as NP. Vs2 NP will combine to form VPs2. PRs2 and VPs2 will then combine to form S.

For sentence "Le chat mange le poisson," the grammar will recognize the following constituents: Dtm Nsm Vs3 Dtm Nsm. The Dtm Nsm in both cases will combine to form NPs. We now have NPs Vs1 NPs. The end NPs will form NP, and then Vs3 and NP will form VPs3. NPs VPs3 will then form S. Note that mange also would be labelled Vs1, but because there is no rule that says NPs Vs1 that parse will fail.

Finally, for "Emily choisit le chat noir" the grammar will recognize the following constituents: PN Vs3 Dtm Nsm A1sm. Nsm A1sm will combine to form Nsma. Dtm and Nsma will then combine to form NPs. NPs will then be labeled as NP. Vs3 NP will combine to form VPs3. PN Vs3 will then combine to form S.

Because there are so many rules involved in the French grammar attempting to model it with an FSA would quickly get messy as each node would have multiple branches coming from it. There are also plenty of irregularities that would make it complicated to model with an FSA. Modeling the French grammar with a CFG however will lead to slow parsing, as the grammar will have to go through many different parses before arriving at the correct parse. This is due to the fact that each rule has many possible outputs.

My grammar does not account for any plural proper nouns and any proper nouns that take articles. Additionally, it does not account for any verbs that are not present tense and does not account for any prepositional or adverbial phrases.

3 Question 3

	cipher 1	cipher 2	cipher 3
	10.16	14.85	21.51
lm	6.64	6.64	7.14
laplace	97.65	84.16	22.17
lm and laplace		0.0631	0.0830

Overall, the hidden Markov model did worse that I expected it to, particularly on the Caesar cipher, which is really just a mapping of one letter to another. Adding the Laplace smoothing led to the most significant jump in accuracy, particularly for the two more simple ciphers. The most complicated cipher, however, remained around 20 percent accuracy even when using Laplace smoothing. The simple ciphers jumped all the way from around 10 and 15 percent to about 98 and 84 percent, which is quite significant. I did not expect smoothing to have such a significant impact. However, I believe this is due to the fact that our training set (and testing set for that matter) is extremely small.

One of the biggest surprises was how adding in the bigram counts for the data from assignment one led to a decline in accuracy. I expected this to help the most in accuracy, but it did the opposite. However, this could be due to an error in how the transition probabilities were counted and added to the probabilities from the training data.

Overall, I thought that the accuracy for the basic HMM would be much higher for both the first two ciphers. What is particularly interesting is that the basic HMM was most accurate on the most complicated cipher and was worst at the simplest. This seems rather backwards. However, it is unsurprising that the models achieved greatest accuracy with the simplest model overall and lowest accuracy with the most complicated model overall. This is because an HMM is well suited for the simpler models, but is not for a model in which part of the sentence moves before it is encoded.