CSC 261 Lab 8
File: Analysis.pdf
Summary: Our inference process and results.
Our aim is to attribute a set of corrupted letters to their original typists, without having an original letter to compare to. We do this by training a transition model for the language, where each character has a certain probability of being followed by each other character. Additionally, we train typo models for the typists, which give the probability that, given the true character, the typist types each possible character.

First, the data we gathered, by training on 2 letters and testing on a third, attributed letter. We do this in order to test the accuracy of our model. The table gives the precise values and the graph is a visual aid, although it only has the data from this lab. The differences between labs are discussed in the paragraph following.

Table 1: All Data

Letter Tested	1	8	16	4	9	18
Letter Trained On						
1,4		71.1	102.8		64.9	-15.9
8,4	171.4		120.4		169.8	114.4
16, 4	236.9	58.4			228.2	446.1
1,9		-239.9	-46.4	-134.4		-335.3
8, 9	131.6		81.2	46.9		-38.2
16,9	78.9	164.8		61.8		229.5
1,18		-172.1	-78.1	-126.8	-189.5	
8, 18	72.0	40.5		19.1	-63.3	
16, 18	-12.9		100.5	-9.7	12.9	

The average evidence that Typist 1 wrote letters 1, 8, and 16 (which they did in fact write) was 49.0 dB. The average evidence that Typist 1 wrote letters 4, 9, and 18 (which they did not) was 26.7 dB. Of the 18 tests for the letters truly written by Typist 1, 13 reported positive evidence that Typist 1 had written them, and 5 reported negative evidence. Of the 18 tests for the letters truly written by Typist 2, 10 reported positive evidence that Typist 1 had written them, and 8 reported negative evidence. Altogether, this means of the 36 letters, we correctly identifed 21 of them. This is about a 58% accuracy rate, or if we apply Laplace's rule of succession, our best guess for the true accuracy rate is about 57.9%. If you only include results where the absolute value of the evidence is greater than 50 (somewhat arbitrarily chosen by us as "strong evidence"), $\frac{17}{27}$ are correct, an accuracy rate of about 63%. This is a little better, which you would expect if the evidence was a good predictor of true typist, but not dramatically so. With the old data, we had an average absolute evidence of about 383 — now, we have an average absolute evidence of 119 instead. We think this is partly caused by the transition model enforcing some convergence of our model for the state of each typist, even if the typo models are different. In addition, in the last lab, the more training we smaller the more the absolute evidence became - maybe adding additional pieces to our model does this in this case, as well.

When model was trained for both typists on all six letters available with attribution and correct text, the program was able to correctly determine the typist with 67% percent accuracy. Thus, assuming that this result can be generalized, the predictions for the six new attributed letters lacking accompanying text would be approximately 67% accurate as well. We could also consider applying Laplaces rule of succession to our prediction of accuracy - its not clear we should have a uniform prior, but we could do so, in which case we will instead get a prediction of $5/8 = 62.5\%$. The confidence in the predictions varies depending on the results for particular letters. Positive numbers indicate that the prediction is for typist 1, and negative numbers indicate that the prediction is for typist 2. A large positive number indicates that the confidence in the prediction should be high. Accordingly, a negative number farther away from 0 indicates that the confidence for the prediction for typist 2 should be higher. Notably, the evidence contracts with more data, which will affect our expectations of how strong evidence we will see to draw a conclusion.

The model trained for both typists on all six letters available with attribution and correct text predicts that unknown letter 5 was written by typist 1 with a Jayness evidence of 9.15. This indicates that the confidence in the prediction should not be very high as it is close to 0. For letter 6, the model predicts that the letter was written by typist 2 with a Jayness evidence of -10.91. The Jayness evidence is close to 0 which means again that the confidence in the prediction should not be very high. The Jayness evidence for letter 20 is -124.99 so we should have more confidence in this prediction compared to the previous predictions made about the unknown letters. For letter 21, the Jayness evidence was -50.15 which is a lower Jayness evidence compared to letter 20, but is still relatively high compared to the first 2 unknown letters. Therefore, we should have a relatively high confidence in this prediction by the program. Letter 23 has a Jaynes evidence of 1.99 which is very close to 0 which indicates that the prediction that this letter was written by typist 1 should not hold much confidence. Letter 24 has a Jayness evidence of -30.17 which indicates that we should have decent confidence that this letter was written by typist 2. Overall, the previous tests on the model indicate that the program has an overall accuracy of 67% so we can use that knowledge along with the Jayness evidence to determine how much confidence we should have on a given prediction.