CIS 526 : Homework 2
Decoding

Anwesha Das

February 20, 2015

1 MODEL IMPLEMENTED

The code given only implemented monotone decoding, i.e., did not consider different permutations of target phrases.

So the next step was to implement reordering. Instead of considering phrases from just the first few of the input sentences, it would consider the whole sentence. By varying the stack size and the number of possible translations options for every possible input phrase, it gave good results but not enough to beat the baseline.

The other things that I tried were introducing a distance based penalty and future cost.
We calculate the distance between the index of the first word of the phrase being translated and the index of the last word of the last phrase translated. Usually in French to English translations, there is not much of a realignment. Hence the distance jumped over during reordering should not be much. Thus we put a penalty on the longer jumps. I found penalizing all jumps **above a distance of 6** gave the best results.
We penalize by taking the absolute value of the distance and raising it to a constant value α. α lies between 0 and 1, where 0 is the harshest penalty and 1 is most relaxed. Another observation I made was that using a value of $\alpha = 0.9$ gave the best result.
Using a distance based penalty gave me a model score of -1323.8577.

Implementing reordering may lead to a situation where a part of the sentence with the lowest probability may get translated first, but translating the rest of the sentence becomes more expensive due to higher costs. Hence we estimate the probability of translating the rest of the sentence and consider it too, so that we do not end up picking the more expensive option.

Using the algorithm given in the book, I implemented the future cost. **Future cost along with reordering** alone gave me my best score -1249.20.

The following section summarizes the different model scores.

2 RESULTS SUMMARY

Here s represents the stack size, k is the number of translations to be considered, a is the α and d is distance above which we penalize reordering probability.

MODEL	s	k	a	d	Score
Given Code (Monotone)	10	20	-	-	-1353.54
Reordering	3000	100	-	-	-1339.12
Reordering with Penalty	200	20	0.9	6	-1323.85
Reordering with Future Cost	2000	200	-	-	-1249.20