**NLP hw1**

# Alejandro Moscoso 332336908

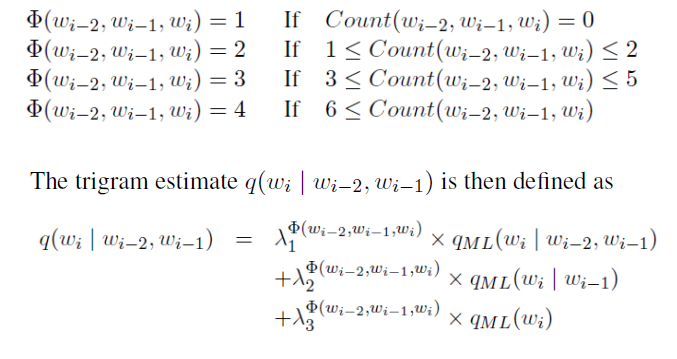
# Alex Finkelshtein : 307573378

**Q1**

We define perplexity as the lectures there for we want to minimize perpexitly with respect to

**Q2**

For every N gram where they can improve our model. Using method from the hw:



Can yield serious problem given that most of the of the lambda took will be given that most count (w1,w2,w3) = 0 in a real problem, there for the influence of having different will be minimum, also are choose taking in to consideration a word that will not influence the probability that they represent.

**Q3**

**Input:** a sequence (sentence), parameters q(s|u,v) and Tag Dictionary (we are given that the set of y’s that correspond to each x is at most K (constant))

**Initialization**: Set ;

. ; =S for (S is all tags)

**Algorithm:**

For k=1,…,n

For

Answer[n-1], Answer[n] =

For k =n-2 : 1

Answer[k] = bp(k+2,answer[k+1], answer[k+2])

**Return**: Answer

The computation complexity is similar to the original Viterbi Algo. The difference is in the inner loops, under the given information about the T(x) size limit, we get that each inner loop is at most K, and also the max operator over at most K elements. Hence we get .

