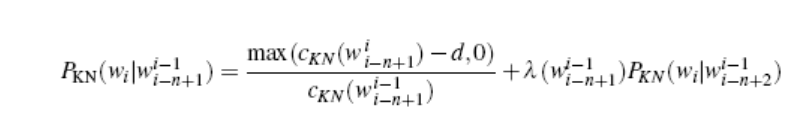
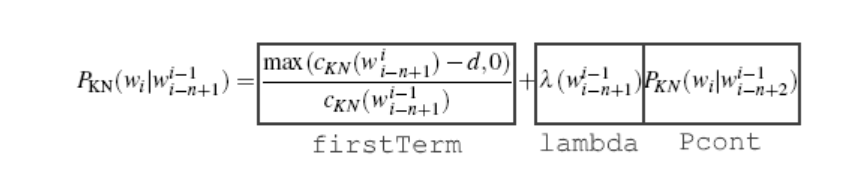
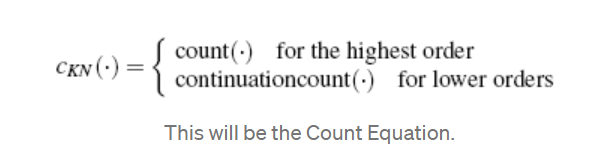
<https://medium.com/@dennyc/a-simple-numerical-example-for-kneser-ney-smoothing-nlp-4600addf38b8>



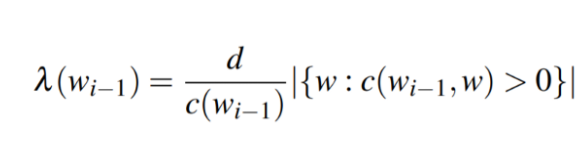


1. First Term



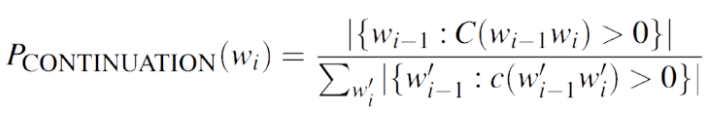
* Nominator is simple count of the full n-gram (4gram is count of the 4 words)
* D is zero at highest order
* Denominator is count of string (n-gram minus final word)
* For lower ngrams: continuation count ->
  + Nominator is count of possible word types before the ngram we consider at the time
  + Denominator is count of possible word types before the string we consider at the time

1. Lambda



* Zero for highest order n-gram
* Denominator is count of string
* Nominator is the number of different words that follow the string
* Lambda is a function of the string, not the ngram, therefore same for all final words

1. Continuation Probability



* Numerator is number of different string types that precede a final word
* Denominator is length of n-gram table