**Code analysis for Dependency based word embeddings paper code**

* Using conll format file to get tuples.
* Then from these tuples create two separate files – a word vocabulary file which consists of words and their counts and context vocabulary file which consists of contexts and their counts.
* In these files they filter out less frequent words or words below a certain threshold count.
* Then the computation is divided into different threads.
* Use the context file to create a unigram table from which randomly then contexts are chosen for negative sampling based on the number given by user.
* So in each pass we have one (word,context) that is passed and then the consecutive negative samples.