**Zachary Jones  
1/18/13  
Lab 2**

Subproblems

* Processing all files in a directory
* Parsing an individual file
* Storing the parsed sentences in a data structure of some form
* Do we hold all files in memory at once or only the two we’re looking at?
* Need hashing function for our hash table, which takes a set of words and produces an integer to use as the index of the storage array
* How to store and read a hash table to a file to avoid holding them all in memory at the same time.

Data Structures

* Hash table to store word sequence permutations
* Hash table could even just store Boolean representing whether that word sequence appears or not
* Data structure for storing the relationships between files (how many phrases in common)

Function structure

* int main()
* HashTable \* parseFile(char \* fileName)
* I’m just not sure yet.