AI similarities using Linux bash script
Michael Chrisco

1. **Project Description**
   I will be creating basic AI algorithms such as the bag of words model for formal analysis and the term matrix in the linux bash shell. Using such models, I shall demonstrate the speed and space required for AI analysis and applications.

2. **Project Techniques**
   Data structures:
   1) term matrix (string matrix)
   2) K-tree (tree of words)
   3) bag of words (string list)
   4) bash scripting language
   5) documents (a set of strings)

3. **Project Group**
   I will be doing the majority of the project by myself with some notes from my CSCI 150 class (which I will specify with citations from Dr. Cao class).

4. **Project Design**

   1) **Bag_of_words function**
   given a document D, where a document is a collection of strings S.
        separate the words from the strings w in S.
   separate unique words w in S.
   concatenate all unique words in new document document_name_BOW.txt

   2) Create master BOW
   for each document, create a bag of words document.
   Concatenate all documents into Master.txt
   run the bag-of-words function on Master.txt to create a new document MASTER_BOW.txt

   3) Create a matrix on MASTER_BOW.txt for each document.
   word_in_MASTER0|word_in_Document0....n
   word_in_MASTER1|word_in_Document0....n
   word_in_MASTER..|word_in_Document0....n
   word_in_MASTERn-1|word_in_Document0....n

save as new document Document_matrix.txt
4) go over matrixes with different AI algorithms (such as k-tree, simple AI and similarity algorithms) for similarities.
5) show results over the data set for each AI algorithm.

5. **Project Details**
A majority of the program will be made in bash scripting language. The k-tree might be implemented in C++ to save time/space but might not be necessary. The output will show the steps as they are happening. The data set result will be made by using tput commands and formatted output.

6. **Project Timeline**
1st week → create Bag-of-words function, create master script that will control the entire operations
2nd week → create Master-Bag-of-words function. Create matrix function. Integrate with master script.
3rd week → integrate output for matrix with AI simple scripts. Create k-tree script.
4rth week → presentation and paper.