Warren Smith  
11/13/13  
CSC335-01  
Project #2: Proposal

Real-World Problem

When writing a paper or even just looking for a good article to read a person should have an idea of what it is he/she is looking for. It’s hard sometimes to look at an article and know if it contains the information wanted without having to read over the whole thing. When a paper is very long and doesn’t contain an abstract this becomes even more challenging because people don’t want to usual read something unless they know it has to do with what it is they are looking for. Due to this, my proposed project plan is to pre-analyze a file before it being read to help the user determine if it is a suitable for their needs.

Project

To accomplish this, a user interact screen would be presented that prompts the user to type the name of the file in question. It will also ask the user to enter keywords (at the most 3) that will serve as an indicator if the inputted file is worth reading or not. Once that information has been entered the system will use substring search algorithms discussed in class to find each of the keywords in the article. The number of times each keyword has been used in the article will be recorded and a percentage of how much each is present in the paper will be calculated. Based on the percentage the system will determine if the paper is worth reading based on the keywords that the user are most interested in.

* Implemented in Java
* Use principles of brute-force and Rabin-Karp substring search algorithms mixed with counting/percentage algorithms created.

\*The program is useful in helping find files that contain the information the user wants without them being read, review files to make sure certain phrases aren’t present, and other possibilities based on what keywords are entered.

Edited for final submission on 11/30/13

* Final program takes in as many keywords as the user wants now (no longer limited to 3)