Will Childs
CS256
Lab Assignment 6
11 17 2024

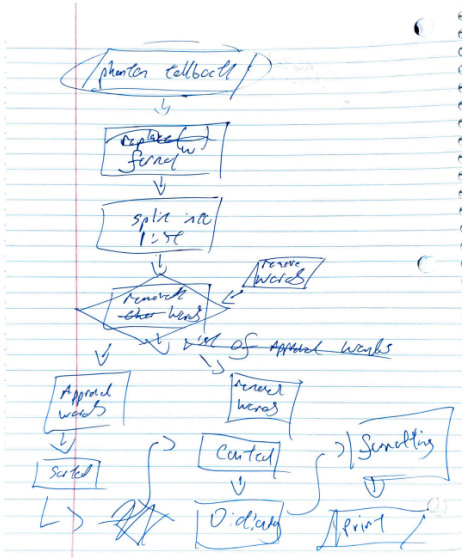
Lab Assignment 6 Design Document

Introduction:
Lab 6 provides a string of the book, *The Phantom Tollbooth*, and we must transmit that string into a format that shows the book's 50 most popular words. This program uses the split command to make the book into a list of words. Then that list is crossed with a few lists to remove unwanted words. That list is formed into a dictionary with the values being the frequency of the word. The dictionary is then formatted to display the 50 most popular words.

Functional Requirements:

1. The only input is the text for *The Phantom Tollbooth*, which is provided via import.
2. The program manipulates the text to print the 50 most popular words and their frequencies.
3. The program prints a formatted version of the dictionary which shows the 50 most popular words and their frequencies.
4. The program only functions if the input is a string

Design Requirements:

1. Variables:
    a. Text of phantom_tollbooth as a string
    b. Text of phantom_tollbooth as a string without select characters, and lowercase letters
    c. Text of phantom_tollbooth as a list
    d. Text of phantom_tollbooth as a list without selected words
    e. List of words to be removed
    f. Text of phantom_tollbooth as a dictionary of words and frequencies
    g. Text of phantom_tollbooth as a dictionary of words and frequencies with the highest 50 words.
2. Operators, functions:
    a. get_text -> takes string from phantom_tollbooth import
    b. replace -> replaces grammatical characters that are unwanted with a space
    c. lower -> converts all characters to lowercase
    d. split -> converts string to list with spaces removed
    e. append -> adds approved words to filtered_words
    f. sorted -> reorganizes list, dict a-z
    g. items -> reorganizes dictionary to output keys and values
3. Data structure:
    a. phantom_tollbooth.py is our imported text
4. Coding concepts:
    a. Strings, list, dictionary, branching
5. Conditions my program needs to handle
    a. Needs to input a string, should work as long as the input is a string ?
6. 

7. Not relevant
8. Output:
    a. Prints keys and values of dictionary
9. Not relevant

Testing Predictions Results:

1. One data set for input showing the first 1000 characters or 100 words dependent on list or string type.
2. Tests were continued to remove words that TagCloud disapproved of
3. Lots of words that were not part of the TagCloud

Reflection and Questions (Lab 6 and on):

1. I had a general idea of the process and worked from the top down, fixing problems as they arrived
2. I started with print functions and continued to edit the string to my liking, then formed a list, and sorted it.
3. How does dictionary sorting work surrounding, tuples and anonymous functions
4. Yes
5. I stuck with the forward motion concept.
6. 100%, I don't know how I could improve this.
7. I understand all of my code completely

Collaboration ***(starting with lab 2– optional)***

I asked copilot about how to sort a dictionary and learned how.