Git hub Link: https://github.com/VidukIn/Python-Abbreviation

Program Overview

Objective: Read a text file which contains names and form three letter abbreviations following the specified rules.

Rules:

1. Abbreviations must only consist of uppercase letters
2. Apostrophes and any other non-letter characters should be avoided
3. Each abbreviation should start with first letter of the word and two other letter pairs from the word (Assumption: Each word will start with a letter)
4. Abbreviation names for certain text file should be distinct. No two names should have same abbreviation.
5. When forming abbreviations:
   a. First letter has score 0
   b. If letter is the last letter of the word it has a score of 5, but if that letter is E score will be 20
   c. If the letter is first nor last, it will have score as follows:
      i. If its second letter 1+value of that letter from values.txt file
      ii. If its third letter 2+value of that letter from values.txt file
      iii. For any other position 3+value of that letter from values.txt file

Implementation

Rule number one and two will be handled inside the generate_abbreviation (Function 03) function of the program. First line of that program will convert names into upper case and then it will remove apostrophe's and non-letter characters.

Function 01 - read_letter_scores

This function will read letter scores file and stores them in a dictionary (letter_scores). It will Iterate through each line of the file, splitting the letter and its score, converting the score to an integer, and storing them as key-value pairs.

calculate_score(word, letter, letter_scores)

This function will implement the rule 05. It will first get the letter score dictionary created in function 01. Then it will check the letters of word and their positions to assign scores. As per above rules it will check whether if it is the first letter of the word (0 index). It gets score 0. Then it will consider the last letter (len(word)-1). Inside that it will check for another condition. If the letter is E it will assign 20 and

otherwise it will be 5. Finally it will consider 3 rules inside if else to assign values for 2nd letter, 3rd letter and other letter positions.

generate_abbreviations(name, letter_scores, used_abbreviations)

This function will iterate through each word in given names. It will start forming abbreviations with first letter. It will generate all 3 letter abbreviations possible for each word as step 01. Then it will call calculate score function and calculate score for each abbreviation formed. Then it will store abbreviation, score and the respective name in the abbreviation set created at the start of this function.

find_best_abbreviation(name, letter_scores, used_abbreviations)

Identifies the best abbreviation for a given name by considering their scores and uniqueness. Selects the abbreviation with the lowest scores that have not been used for other names.

**The program maintains a dictionary "used_abbreviations" to keep track of abbreviations used for each name to ensure uniqueness. It also handles cases where no valid abbreviations can be generated for a name.

This will satisfy the rule 04.