

01/12/2017

BLM2541
Programming Languages
Assignment - 3
(Due 11/12/2017 – 23:59h)

Write a “Word Fixing” program, where the poorly written words within a given text file are detected and the correct alternatives of the given words are saved into a new file. You should exploit the given Turkish Corpus files in order to find the correct alternatives of the given wrong words.

Inputs

- Files consisting of correct word(s)
- Letter Distance for correcting wrong written words: (1,2,3) (It means that the alternative words up to given letter distance should be selected and suggested by your program. For example, if the user gives 2 as Letter Distance, all the alternative words with distance 1 and 2 should be saved to the new file.)

Additional Info about Letter Distance:

If a given word or a group of words is found in the given corpus files, your program should write them into a new text file and add a plus ‘+’ sign after the word. If not found, you would add a negative sign after the word and display all the alternatives based on the Letter Distance.

Letter distance refers to the number of letters which are different from the given word(s). For example, suppose that the word “zeyzin” is given to the program. The letter distance of “zeyzin” to “zeytin” and “benzin” are 1 and 2, respectively.

Important Notes:

Your program should compare the words within the given text file only to the words with the same number of letters. For example, “zeyzin” should be compared only with the words which contain 6 letters.

Do not forget to download the corpus files given as a RAR file (Turkce_Kelimeler_Corpus.rar) in my website.

EXAMPLE

Inputs

Text File Consisting of correct and wrong word(s):

zeyzin
ada
aca
nakarat
aile
vagon
maden filizi
makeni

Letter Distance : 3

Output File

zeyzin	-	zeytin benzin beylik meydan
ada	+	
aca	-	aba aç1 ada aha ama ait adi afi aft ahi dua dađ ece
nakarat	+	
aile	+	
vagon	+	
maden filizi	+	
makeni	-	madeni macera medeni mahana makine

You should also report;

- The hardware specification of your computer (CPU, RAM)
- The compiler
- The development environment

Note: Use reasonable variable and function names.

ATTENTION

Do not forget to read the document “ASSIGNMENT RULES”, located in my website.