Dr.Erdem KAYA ekaya@metu.edu.tr Mehmet Taha ŞAHİN staha@metu.edu.tr

Assignment 03 METU CENG310 Spring 2021 Data Structures and Algorithms with Python

Start Date: June 3rd, 2021 Due Date: June 16th, 2021

1 Spell Checker

In this assignment, you are expected to implement a spell checker class (SpellChecker) as described in P-10.55 in the textbook.

The SpellChecker class has to use either ChainHashMap or ProbeHashMap class both of whose implementations are provided in the text book (Please ignore the "For an extra challenge, consider phonetic substitutions as well." part.). In this hash table data structure you should have the most common 10.000 words of English (https://github.com/first20hours/google-10000-english), so that it could be used as a reference against which input words are compared.

The SpellChecker class will have a public method called check which will take a file name string as its sole parameter and create an output file with a list of spell checking recommendations. At each line of the output file, there has to be one of the input word followed by

- 1. The word 'OK' in case of an exact match (please make a case insensitive comparison after removing punctuation letters), or
- 2. At most two recommended words in case of no matches (please try only by inserting a single character in between two adjacent characters in a word or by dropping only one of the letters of a word if a word is at least of length 3), or
- 3. The phrase 'No Recommendation' when there is no recommendation for the input word.

An example input file content:

```
Murphys Law Murphy
The chance of the bead falling with the butered side down is directly propotional to the cost of th carpet.
```

An example output file content:

```
1 Murphys --> murphy
2 Law --> OK
3 Murphy --> OK
4 The --> OK
5 chance --> OK
6 of --> OK
7 the --> OK
8 bead --> bread, beads
g falling --> OK
10 with --> OK
11 the --> OK
12 butered --> No Recommendation
13 side --> OK
14 down --> OK
15 is --> OK
16 directly --> OK
17 propotional --> No Recommendation
18 to --> OK
19 the --> OK
20 cost --> OK
21 of --> OK
22 th --> OK
23 carpet. --> OK
```

2 Delivery Instructions

Please hand in your module as a single file named as spellchecker.py over ODTUClass by 11:59pm on due date. An Assignment-03 page will be generated soon after the start date of this assignment. Should you have any questions pertaining to this assignment, please ask them in advance (rather than on the due date) for your own convenience. Whatever IDE you use, you have to make sure that your module could be run on a Python interpreter:

```
sc = SpellChecker()
sc.check('Assignment3Input.txt')
```