**Spell Checker for User Input [30%] + [10% bonus]**

Natural language processing technology is still booming around the world, and it is widely used in many fields, which greatly facilitates our daily life. There are several examples for your consideration. First, [deep learning translation](https://www.deepl.com/translator) can accurately translate one language into another language, such as between Chinese and English. Second, Microsoft Word Document has the function to prompt multiple similar words for the misspelled word for the user to modify. Third, the software [Grammarly: Free Online Writing Assistant](https://www.grammarly.com/) can check spelling and grammar by scanning user’s inputs, and giving several recommended sentences.

Now, suppose we have a word dictionary kept in a document. To solve the common problem that the user inputs some misspelled words in the console, we want to write a spell checker to find these misspelled words by comparing each word in the user inputs with those in the word dictionary. Then, the spell checker should give some recommended words from the word dictionary as the references for user correction. Next, the spell checker should prompt the user to select one of the recommended words, or give up correcting the identified misspelled word. Finally, the spell checker should save the corrected user inputs in a document. For details, please refer to the instructions and samples below.

You are required to write a program, called spellChecker, to check user inputs, show recommended words, prompt the user to correct the inputs, and save the corrected user inputs in the document based on the following instructions:

* For the **paragraph** that the user inputs to console, your program can assume as follows:
  + The whole input is one **single** paragraph.
  + The **whole** paragraph must be provided before the **enter** key is pressed and read to a **cstring** from the console.
  + Assume all the words in the paragraph are in **lower case**.
  + Every word in the paragraph consists of at most **30** characters, and the inputs will not be jammed with empty lines.
  + The number of words in the paragraph is at most **1000** words.
* The given word dictionary document is a text file, containing a list of **109582** English words, named "[**wordsEn.txt**](https://canvas.cityu.edu.hk/courses/45715/files?preview=8642172)", and the output text file is called "**output.txt**". Your program should place it in the project directory of your VS C++ project.
  + Hint: Exercise 3 in Tutorial 8 illustrates a method to read and check against this word dictionary.
* To simplify the program, we assume each word in the paragraph either is correct or contains exactly **one misspelled character**. Moreover, we assume that if the word is correct, it must exist in the word dictionary.
* The spell checker detects each incorrect spelling mistake by searching each word in the paragraph whether it is in the word dictionary.
  + If the word cannot be found, take it as a misspelled word;
  + Otherwise, take it as a correct one.
* For each misspelled word, suggests a bulleted sequence of words by selecting from the word dictionary.
  + Each word in the selected words should contain one character different from the misspelled word.
  + Show **at most 10** suggestions for the misspelled word.
  + If the number of selected words is 0, meaning that there is no valid suggested correct word for the current misspelled word, show "**No suggest.**" in the console.
* For each misspelled word, prompt the user to choose one of the bullets or give up the modification.
  + If the user chooses the displayed index number of a suggested word (note that each suggested word should have different index numbers), the program replaces the misspelled word with the selected one.
  + If the user chooses to give up, the program continues to handle the next misspelled word.
* Finally, the program writes the corrected paragraph into the "output.txt" file.

Hints:

* The number of words in the word dictionary is very large and cannot be stored in one array. You can divide the word dictionary into several parts, and read one part at one time. After one part has been looked through, you can read another part of the word dictionary into the same array.
* Each word in the paragraph either is correct or contains exactly one misspelled character, where one misspelled character could be **one more** or **one less** than one character or the misspelled character should be replaced by another character. For example, the word "aroundd" is misspelled due to one more character 'd' the correct word "around".
* Any symbol, space, or other special characters in the paragraph should also be saved in the "output.txt".
* You may create several functions to reduce the complexity of your program and improve readability.
* In some situations, if you want to clear the array, the function memset(array, ‘\0’, sizeof(array)) is a good choice. Or, you may simply iterate all the elements in the array one by one and set its value to 0.
* If your program has been considered comprehensively, you can solve the problem by scanning the whole paragraph only once.

**Example 1:**

|  |
| --- |
| Please input paragraph in the console:  tha town revolved aroundd the river. in summmmer, when the blazing sun beat down, it dozed under the weight of sultry days. on main street a sow and her litter of pigs might root along the woodon sidewalks, sharing the deeply rooted raoadway with foraging hens and a hound languidly scratching his fleas.  **(Below are console outputs and selections)**  The spelling of tha is incorrect! Please select one of the following suggestions:  1. aha  2. ha  3. tea  4. th  5. thai  6. than  7. that  8. thaw  9. the  10. tho  11. give up  Your selection is: 9  The spelling of aroundd is incorrect! Please select one of the following suggestions:  1. around  2. give up  Your selection is: 1  The spelling of summmmer is incorrect! Please select one of the following suggestions:  No suggestion.  1. give up  Your selection is: 1  The spelling of woodon is incorrect! Please select one of the following suggestions:  1. wooden  2. give up  Your selection is: 3  Your input is invalid! Please input again!  Your selection is: 1  The spelling of raoadway is incorrect! Please select one of the following suggestions:  1. roadway  2. give up  Your selection is: 5  Your input is invalid! Please input again!  Your selection is: 1  The corrected paragraph has been written into output.txt. |

Output content of **Output.txt**:

|  |
| --- |
| the town revolved around the river. in summmmer, when the blazing sun beat down, it dozed under the weight of sultry days. on main street a sow and her litter of pigs might root along the wooden sidewalks, sharing the deeply rooted roadway with foraging hens and a hound languidly scratching his fleas. |

**Example 2:**

|  |
| --- |
| Please input paragraph in the console:  all questions should be submitted in anvas. note that all solutions will be assessed by (1) correctness, (2) programmeing styles (including comments), and (3) non-redundancy in expressing solutions, if applicable. this assignment contaings 5 questions nameded as Q1 to Q5. Q5 has 10% bonus. Your solutions are forbidden to includz any library other than <iostream>, <iomanip>, <cstring>, <fstream>, and/or <cmath>. each solution using any function/facility in other libraries will receive 0 mark. if possible, not using <cstring> in your solution.  **(Below are console outputs and selections)**  The spelling of anvas is incorrect! Please select one of the following suggestions:  1. angas  2. annas  3. canvas  4. give up  Your selection is: 3  The spelling of programmeing is incorrect! Please select one of the following suggestions:  1. programming  2. give up  Your selection is: 1  The spelling of contaings is incorrect! Please select one of the following suggestions:  1. contains  2. give up  Your selection is: 1  The spelling of nameded is incorrect! Please select one of the following suggestions:  No suggestion.  1. give up  Your selection is: 1  The spelling of includz is incorrect! Please select one of the following suggestions:  1. include  2. give up  Your selection is: 4  Your input is invalid! Please input again!  Your selection is: 1  The spelling of iostream is incorrect! Please select one of the following suggestions:  No suggestion.  1. give up  Your selection is: 1  The spelling of iomanip is incorrect! Please select one of the following suggestions:  No suggestion.  1. give up  Your selection is: 1  The spelling of cstring is incorrect! Please select one of the following suggestions:  1. string  2. give up  Your selection is: 2  The spelling of fstream is incorrect! Please select one of the following suggestions:  1. stream  2. give up  Your selection is: 2  The spelling of cmath is incorrect! Please select one of the following suggestions:  1. math  2. give up  Your selection is: 2  The spelling of cstring is incorrect! Please select one of the following suggestions:  1. string  2. give up  Your selection is: 2  The corrected paragraph has been written into output.txt. |

Output content of **Output.txt**:

|  |
| --- |
| all questions should be submitted in canvas. note that all solutions will be assessed by (1) correctness, (2) programming styles (including comments), and (3) non-redundancy in expressing solutions, if applicable. this assignment contains 5 questions nameded as Q1 to Q5. Q5 has 10% bonus. Your solutions are forbidden to include any library other than <iostream>, <iomanip>, <cstring>, <fstream>, and/or <cmath>. each solution using any function/facility in other libraries will receive 0 mark. if possible, not using <cstring> in your solution. |