

# UNIVERSITY<sup>OF</sup> BIRMINGHAM

## C++ for Finance Assessment 1 Spell Checker

1222781

Date: Nov 2016

School of Social Sciences University of Birmingham Birmingham, B15 2TT

### **Contents**

CONTENTS

| 1 | Briefing            | 2 |
|---|---------------------|---|
| 2 | Functionality       | 2 |
| 3 | suggestion creation | 4 |
| 4 | Source Codes        | 5 |
|   |                     |   |

#### 1 Briefing

In this assessment I have created a prototype for a spell checker. The program is written in ISO C++ 11 using DevC++ 5.11 compiler. The project contains four files :

- · main.cpp
- · small\_dictionary.h
- small\_dictionary.cpp
- utility.h

The program contains a dictionary of 200 common English words. The program use this dictionary to check if a word is correct or to provide suggestions for the words that do not exist in the dictionary. You can find the dictionary in the source code number 3.

#### 2 Functionality

The program starts by asking the user to input a word. If the word was found in the dictionary, it would inform the user and then asks you to to either continue or quite the program. If you chose to continue, the process starts again, and inf you chose to exit the program, the program ends. This is demonstrated in the Figure 1

#### 

Figure 1: Screen-shot of the console. The console closes after entering the q

However, if you enter a word that does not exist in the dictionary, the program will try to give you suggestions from the dictionary. It will then asks the user to chose one of the suggestions or proceed. To chose a suggestion one should enter the suggestion number or to proceed should enter an integer outside the suggestion range. If the user enters a suggestion number, the program will inform the user of the decision, and ask the user to either quit or continue with the next word. The program is designed to ignore inputs that are out of range. This processes can be seen in Figure 2

```
Please enter your word : follow
The word is found in the dictionary
Enter:
    -q :to quit the program
    -c :to continue entering words
your choice : c
Please enter your word : ge
** The entered word could not be found in the dictionary **
The suggested word(s) are/is:
                       1-be
                       2-he
                       3-we
                       4-go
                       5-get
If you meant any of the suggestions enter the suggestion number
otherwise press any integer to proceed
Enter your choice: h
Enter your choice: f
Enter your choice: 1
The correct word is now : <be>
Enter:
    -q :to quit the program
    -c :to continue entering words
your choice : q
```

Figure 2: Screen-shot of the console. The console closes after entering the q

However if the user decides not to go for any of the suggested words and enter a number outside of the range, the program will ask the user to either add the new word to the dictionary or ignore the word. If the user likes to add the word to the dictionary, the user will be notified, but in either case the user will be asked to chose to either continue with the next word or to quit the program. This process is illustrated in Figure 3

In the same way as the previous part, the user input is checked for validity. If the user inputs a string that is not recognised by the program, the user will be asked to input another response. This also applies to the input at quit or continue level.

```
Please enter your word : gollow
** The entered word could not be found in the dictionary **
The suggested word(s) are/is:
                       1-follow
If you meant any of the suggestions enter the suggestion number
otherwise press any integer to proceed
Enter your choice: 2
Enter:
     -add :to add the word to the dictionary
     -ignore :to ignore the mistake
your choice : ignore
Enter:
     -q :to quit the program
     -c :to continue entering words
your choice :
Please enter your word : ge
** The entered word could not be found in the dictionary **
The suggested word(s) are/is:
                        1-be
                        2-he
                        3-we
                       4-go
                        5-get
If you meant any of the suggestions enter the suggestion number
otherwise press any integer to proceed
Enter your choice: 7
Enter:
     -add :to add the word to the dictionary
     -ignore :to ignore the mistake
your choice : add
The Word has been added to the dictionary
     -q :to quit the program
     -c :to continue entering words
your choice : c
Please enter your word : ge
The word is found in the dictionary
Enter:
     -q :to quit the program
     -c :to continue entering words
your choice : q
```

Figure 3: Screen-shot of the console. The console closes after entering the q

#### 3 suggestion creation

Creating a list of suggestions from the dictionary is the most sophisticated part of this program. In order to find suggestions, the program calculates the Damerau-Levenshtein distance between every word in the dictionary and the user's input. It will then supply the words in the dictionary that have the distance of at most 51% of the size of the input word as

suggestions.

The maximum distance has a linear positive relationship withe size of the input word. On the other hand the bigger the word is the more mistakes could happen and therefore more distance should be allowed. If the factor is much bigger than 51% it will give suggestions are are too far away from the input word. However if the factor is anything less than 51% the program will not suggest well for very small words. For example for two letter words, the maximum distance would be computed to be 0 and no word will be suggested. But with 51% the program suggest similar words to two digit words.

#### 4 Source Codes

```
2 // main.cpp
3 // ID1222781
5 #include <iostream>
6 #include <string>
7 #include <algorithm>
8 #include "small_dictionary.h"
9 #include "utility.h"
10
12 // main
15 int main ()
16 {
17
    std::vector<std::string> Dict = dct::Dictionary200 ();
    // creating the dictionary vector.
18
19
    bool run = true; // while run is true the following while loop, loops.
20
    while (run)
21
22
23
       ut::Printline(); //print a line to distinguish new word's journey
       std::string input = ut::GetString("Please enter your word ");
24
       //Ask the user for the input
25
26
       if (dct::existance( Dict , input )) std::cout <<</pre>
27
         "The word is found in the dictionary" << std::endl;
29
30
       else
31
       {
         std::cout << "** The entered word could not be found in the dictionary ** \n";
32
33
         std::vector<std::string> suggestions = dct::suggestions ( Dict , input );
34
         //creating suggestions from the dictionary.
35
         long SugSize = suggestions.size();
         //find the size to see if there is any suggestions at all
37
         bool correction = false; // true when the user has chosen a correction
39
40
         if (SugSize > 0) // if there was any suggestion
42
         {
            std::cout \ll "The suggested word(s) are/is: \n";
43
            for (long i = 0; i != SugSize; ++i)
45
            std::cout << "
i+1 << "-" << suggestions[i] << std::endl;</pre>
            // print the suggestions
48
            std::cout << "If you meant any of the suggestions enter the suggestion number"
50
            << std::endl << " otherwise press any integer to proceed " << std::endl;
51
            long choice = ut::GetLong ("Enter your choice");
53
```

```
55
              if (choice <= SugSize)</pre>
              {
                 correction = true;
57
                 std::cout << "The correct word is now : <" << suggestions[choice-1] << ">\n";
59
60
              else correction = false;
62
           }
63
           while (correction != true) //if the user did not use the sugestions loop starts.
                                // This loops until the user enters a valid input.
65
66
              std::string add = "add" , ignore = "ignore";
std::cout << "Enter: \n" ;</pre>
67
68
              std::cout << "
                                  -add : to add the word to the dictionary \n";
69
              std::cout << "
                                 -ignore : to ignore the mistake";
70
71
              std::string choice = ut::GetString("your choice");
72
              if (choice == add)
73
74
75
                 Dict.push_back(input);
                 std::cout << "The Word has been added to the dictionary \n" ;
76
77
                 break:
78
              if (choice == ignore) break;
79
81
82
        }
83
        while (run) //this loops until a valid respond is captured
84
85
           std::string quit = "q" , fcontinue = "c"; std::cout << "Enter: \n" ;
86
87
           std::cout << "
                             -q :to quit the program \n" ;
-c :to continue entering words" ;
           std::cout << "
89
           std::string choice = ut::GetString("your choice ");
91
           if (choice == quit) run = false;
92
93
           if (choice == fcontinue) break;
        }
94
95
     }
97
98
      return 0;
99 }
100
102 // end
```

Source Code 1: main.cpp.

Source Code 2: small\_dictionary.h.

```
_2 // small_dictionary.cpp
3 // ID1222781
6 #include "small_dictionary.h"
7 #include "utility.h"
9 #include <string>
10 #include <vector>
11 #include <algorithm>
13 long DLdistance (std::string A , std::string B);
16 // dict::Dictionary200 ()
           returns a vector of first initial 200 words
20 std::vector<std::string> dct::Dictionary200 ()
21 {
22
          std::vector<std::string> Dict
                "the", "be", "of", "and", "a", "to", "in", "he", "have", "it", "that", "for", "they", "I", "with", "as", "not", "on", "she", "at", "by", "this", "we", "you", "do", "but", "from", "or", "which", "one", "would", "all", "will", "there", "say", "who", "make", "when", "can", "more", "if", "no", "man", "out", "other", "so", "what", "time", "up", "go", "about", "than", "into", "could", "state", "only", "new", "year", "some", "take", "come", "these", "know", "see", "use", "get", "like", "then", "first", "any", "work", "now", "may", "such", "give", "over", "think", "most", "even", "find", "day", "also", "after", "way", "many", "must", "look", "before", "great", "back", "through", "long", "where", "much", "should", "well", "people", "down", "own", "just", "because", "good", "each", "those", "feel", "seem", "how", "high", "too", "place", "little", "world", "very", "still", "nation", "hand", "old", "life", "tell", "write", "become", "here", "show", "house", "both", "between", "need", "mean", "call", "develop", "under", "last", "right", "move", "thing", "general", "school", "never", "same", "another", "begin", "while",
23
25
26
27
28
29
30
31
32
33
34
35
36
37
38
                 mean, carr, develop, under, fast, right, move, thing, "general", "school", "never", "same", "another", "begin", "while", "number", "part", "turn", "real", "leave", "might", "want", "point", "form", "off", "child", "few", "small", "since", "against", "ask", "late", "home", "interest", "large", "person", "end", "open", "public",
39
41
42
                 "follow", "during", "present", "without", "again", "hold", "govern",
"around", "possible", "head", "consider", "word", "programme",
"problem", "however", "lead", "system", "set", "order", "eye", "plan",
"run", "keep", "face", "fact", "group", "play", "stand", "increase",
43
44
45
                  "early", "course", "change", "help", "line"
47
48
          //The vector comes from the assessment given file.
50
51
          return Dict;
52 }
53
```

```
56 // dict::existance ()
57 // true if the input exists in the dictionary and false if it doesn't
60 bool dct::existance (std::vector<std::string> & Dict , std::string & input)
61 {
    std::vector<std::string>::iterator it;
    it = find (Dict.begin(), Dict.end(), input); // find the location of the
63
64
                                  //input in the dictionary
     if (it != Dict.end()) return true; // if the pointer points at the end of
                            // the vector, input was not found
66
67
     else return false;
68 }
69
71 // disct::suggestions
72 //
     give suggestive words from the dictionary
75 std::vector<std::string> dct::suggestions (std::vector<std::string> & Dict , std::string & input)
76 {
    std::vector<std::string> suggestions;
77
78
    long DicSize = Dict.size();
79
     for (long i = 0; i != DicSize; ++i)
80
       long \ Distance = DLdistance (input , Dict[i] );
82
83
       if (Distance < input.size() *0.51) suggestions.push_back(Dict[i]);
       // The suggestion holds the words with distance smaller than 51% of the input size.
84
       // The logic is: the bigger the word, the more mistakes can be made, the bigger the
85
       // distance is to the real word. 51% was chosen to cover two letter words.
86
       // anything less than 51% does not cover two letter words
87
88
     return suggestions;
89
90 }
93 // DLdistance()
94 // Find the distance between two strings
     local function to the script
97 long DLdistance (std::string A , std::string B)
98 {
99
     long lal = A.size();
                            //number of characters in A
     long lbl = B.size();
                            //number of characters in B
100
101
102
     A.insert(0, "");
103
     B.insert(0,"");
104
      //inserting null characters. It helps to point at the right character when
105
      //pointing at the character number.
106
107
     std::vector< std::vector<long> > d_j_i; // the distance matrix
108
109
110
     std::vector<long> temp; //temporary vector defined local to the function
111
112
      for (long i = 0; i != lal+1; ++i)
114
115
       temp.push_back(i);
116
    d_j_i.push_back(temp); // filling the first row the matrix
117
                      //releasing the memory for temp
118
     temp.clear();
119
120
121
      for (long j = 1; j != lbl+1; ++j) // filling the matrix row by row
122
123
```

```
124
        temp.push_back(j); // The first colume of the matrix is the 1,2,3,...
125
126
         for (long i = 1; i != lal+1; ++i)
127
128
129
            long option1 = temp[i-1]+1;
            long option2 = d_j_i[j-1][i]+1;
130
            long option3 = d_j_i[j-1][i-1];
131
132
            if (A[i] != B[j]) ++option3;
133
            long the_minimum; // The minimum will be allocated to the minimum option
134
135
            if (i > 1 \text{ and } j > 1 \text{ and } A[i] == B[j-1] \text{ and } A[i-1] == B[j])
136
137
            {
               long option 4 = d_j[i-2][i-2]+1; // Under this condition, and onother option exists
138
               if (option1 < option2 and option1 < option3 and option1 < option4) the_minimum = \leftrightarrow
139
                   option1;
               if (option2 < option1 and option2 < option3 and option2 < option4) the_minimum = \leftrightarrow
140
                   option2;
               if (option3 < option1 and option3 < option2 and option3 < option4) the_minimum = \leftrightarrow
                   option3;
142
               else the_minimum = option4;
            }
143
144
            if (i == 1 \text{ or } j == 1 \text{ or } A[i] != B[j-1] \text{ or } A[i-1] != B[j])
145
            {
               if (option1 < option2 and option1 < option3) the_minimum = option1;</pre>
147
               if (option2 < option1 and option2 < option3) the_minimum = option2;</pre>
148
               else the_minimum = option3;
149
150
151
            temp.push_back(the_minimum); // The minimum is found and the next element of row is added↔
152
                 to the temp
153
154
155
        d_j_i.push_back(temp); // The newly created row saved in the temp, is now pushed back to the \leftrightarrow
             matrix.
156
157
        temp.clear(); //Temporary file is cleared and memory is released.
158
159
       return d_j_i[lbl][lal]; // The distance is defined as the final element of the matrix.
160
161 }
  162
163 // end
```

Source Code 3: small\_dictionary.cpp.

```
ntility h
3 // ID1222781
6 #include <iostream>
7 #include < string >
10 // namespace UtilityFunctions
12
13 namespace UtilityFunctions
14 {
15
  std::string GetString(const std::string &);
16
17
  long GetLong(const std::string &);
18
```

Source Code 4: utility.h.

```
2 // utility.cpp
3 // ID1222781
6 #include "utility.h"
8 #include <string>
9 #include <iostream>
10 #include <stdexcept>
11 #include <iomanip>
12 #include <sstream>
13 #include <stdexcept>
16 // declaration of internal function
19 long StringToLong(const std::string & s, bool & success);
22 // GetString()
25 std::string UtilityFunctions::GetString(const std::string & mess)
26 {
  std::cout << std::endl << mess + ": ";
27
28
29
  std::string text;
  std::cin >> text;
30
31
  return text:
32 }
33
35 // GetLong()
38 long UtilityFunctions::GetLong(const std::string & prompt)
39 {
  bool valid_input = false;
41
  long value;
42
  while (valid_input == false)
43
44
45
    std::string input = GetString(prompt);
    value = StringToLong(input, valid_input);
  }
47
48
  return value;
49
50 }
54 // Printline()
```

```
57 void UtilityFunctions::Printline()
58 {
59
               std::cout <<
60
               << std::endl;
61
62 }
63
66 // Validation function for get long used to make sure user inputs long
67 // Function is taken from the course materials.
notation n
71 {
                success = false;
72
73
              std::istringstream i(s);
              double x;
74
             if (!(i \gg x)) return long(x);
76
             long y = long(x);
77
              if (y == x) success = true;
              return y;
79
80 }
83 // end
```

Source Code 5: utility.cpp.