**МИНОБРНАУКИ РОССИИ**

**Санкт-Петербургский государственный**

**электротехнический университет**

**«ЛЭТИ» им. В.И. Ульянова (Ленина)**

**Кафедра информационных систем**

отчет

**по практической работе №4**

**по дисциплине «Программирование»**

Тема: Работа с текстом

|  |  |  |
| --- | --- | --- |
| Студент гр. 9893 |  | Росихин А.Ю. |
| Преподаватель |  | Глущенко А.Г. |

Санкт-Петербург

2019

**Цель работы.**

Изучить способы обработки текстовых данных, чтение из файла и ввод с клавиатуры.

**Основные теоретические положения.**

Работа с файлами. Для работы с файлами необходимо подключить заголовочный файл <fstream>. В <fstream> определены несколько классов и подключены заголовочные файлы <ifstream> — файловый ввод. [1]

Функция getline() имеет конструкцию getline(stream, string, separator),где stream — это поток данных, string — переменная, в которую запишется строка и separator – строковый разделитель, показывающий на конец строки [2].

С помощью функции file.eof() проверяет окончание файла.

Функция s.find(sub) ищет первое вхождение подстроки *sub* в строке *s*. Это необходимо для поиска дублирующихся символов с последующим их удалением.

Функция tolower() переводит символ в нижний регистр.

Функция wholeText.erase(0,position) удаляет из строки *wholeText* подстроку с первой позиции до позиции номер *position*.

Функция word.length() возвращает количество символов в указанной строке.

**Постановка задачи.**

С клавиатуры вводится последовательность, содержащая от 1 до 50 слов, в каждом из которых от 1 до 10 строчных латинских букв и цифр. Между соседними словами произвольное число пробелов. За последним словом точка.

Необходимо отредактировать входной текст:

- удалить лишние пробелы;

- удалить лишние знаки препинания (под «лишними» подразумевается несколько подряд идущих знаков.К примеру, «,,», «;;;;» и тому подобное);

- исправить регистр букв, если это требуется (пример некорректного регистра букв: пРиМЕр);

Вывести на экран только те слова последовательности, в которых встречаются одинаковые буквы.

**Выполнение работы.**

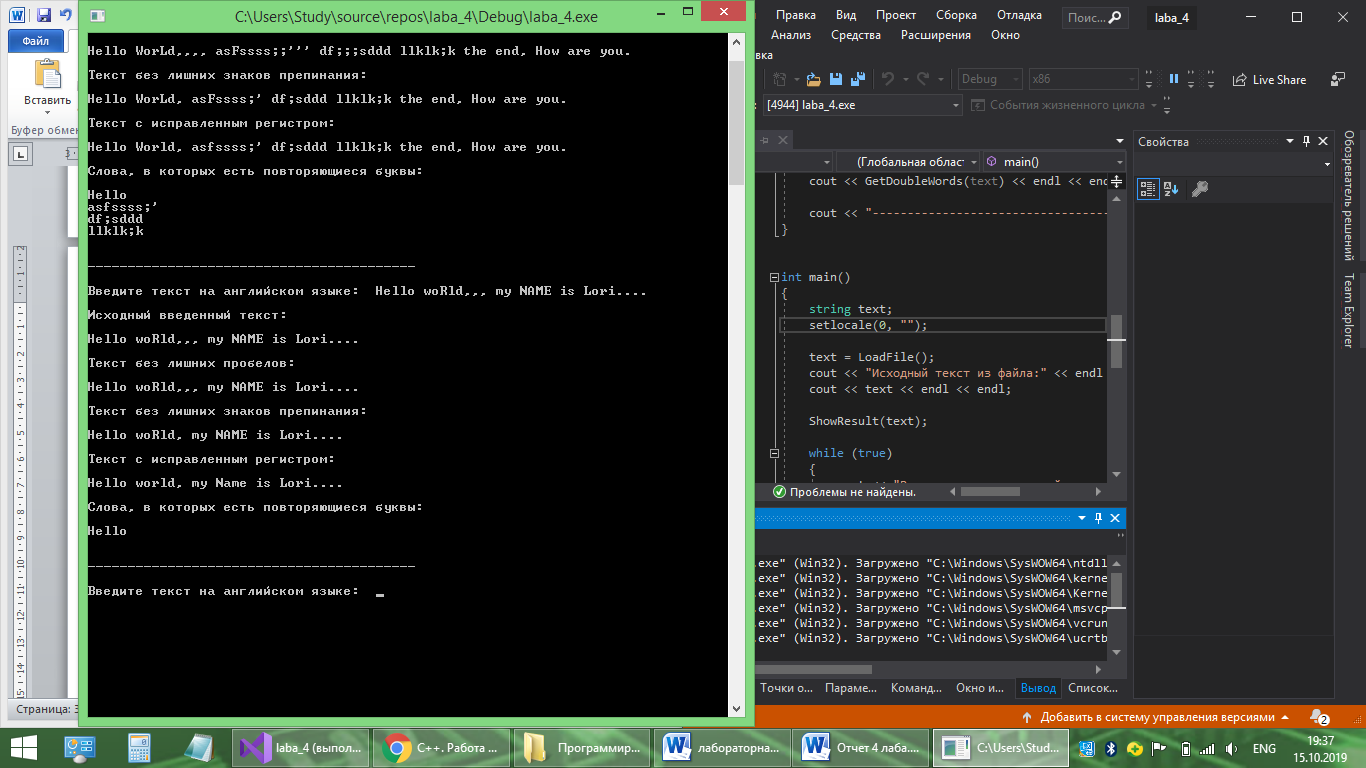
Загрузить текстовый файл в переменную типа string.

Найти и удалить двойные пробелы и двойные знаки препинания.

Найти верхний регистр «внутри» слова и заменить на нижний регистр.

Разделить текст на отдельные слова, проверить повторяются ли там буквы и вывести полученный результат.

Получить на «входе» исходный текст и повторить все операции, указанные ранее. Исходный код представлен в приложении А.Скриншот работы экрана см. рис. 1.

 Рисунок 1 — Скриншот работы программы

**Выводы.**

Для работы с большим объемом текста очень удобно использовать класс string и все функции этого класса.

Благодаря библиотеке <fstream> я научился взаимодействовать с файлами. При работе с файлами для предотвращения ошибок обязательно надо проверять открылся ли файл. Это позволяет нам проверить корректность загружаемого файла.

Также я научился работать с библиотекой <string> с основными ее функциями find(), erase(), length(), что позволяет удобно искать первое вхождение подстроки в строке, удалять необходимые символы и узнать длину подстроки.

Для удаления дублирующихся знаков препинания очень удобно искать не посимвольно, а сразу вести поиск двойных символов.

Приложение А

**ИСХОДНЫЙ КОД ПРОГРАММЫ**

#include<iostream>

#include<string>

#include<fstream>

usingnamespacestd;

stringLoadFile()

{

stringoneString="", allStrings="" ;

ifstream file("text.txt");

if (file)

do

{

getline(file, oneString);

allStrings+=oneString+" ";

} while (!file.eof());

elsecout<<"Ошибка открытия файла."<<endl;

returnallStrings;

}

stringDeleteDoubleChar(stringdeleteSymbols,stringwholeText)

{

intpos = 0;

do

{

pos = wholeText.find(deleteSymbols);

if (pos>0) wholeText.erase(pos, 1);

} while (pos != -1);

returnwholeText;

}

stringLowerCase(strings)

{

//transform(s.begin(), s.end(), s.begin(), tolower); // преобразуетвесьтекст #include <algorithm>

int i = 1;

while (s[i])

{

if (s[i-1]!=' ')

s[i] = tolower(s[i]);

i++;

}

returns;

}

stringGetDoubleWords(stringwholeText)

{

string word = "";

stringresultString = "";

int position=0;

do

{

position = wholeText.find(" ");

if (position == -1) position = wholeText.find(".");

if (position == -1) position = wholeText.find("\n");

if (position != -1)

{

word=wholeText.substr(0, position);

wholeText.erase(0, position + 1);

int repeat = 0;

for (int i = 0; i <word.length(); i++)

for (int ii = 0; ii <word.length(); ii++)

if ((i != ii) && (word[i] == word[ii])) repeat++;

if (repeat > 0) resultString+= word +"\n";

}

} while (position != -1);

returnresultString;

}

voidShowResult(stringtext)

{

cout<<"Текст без лишних пробелов:"<<endl<<endl;

text=DeleteDoubleChar(" ", text);

cout<<text<<endl<<endl;

cout<<"Текст без лишних знаков препинания:"<<endl<<endl;

text=DeleteDoubleChar(",,", text);

text=DeleteDoubleChar("!!", text);

text=DeleteDoubleChar("??", text);

text=DeleteDoubleChar(";;", text);

text=DeleteDoubleChar("::", text);

text=DeleteDoubleChar("''", text);

text=DeleteDoubleChar("))", text);

text=DeleteDoubleChar("((", text);

cout<<text<<endl<<endl;

cout<<"Текст с исправленным регистром:"<<endl<<endl;

text=LowerCase(text);

cout<<text<<endl<<endl;

cout<<"Слова, в которых есть повторяющиеся буквы:"<<endl<<endl;

cout<<GetDoubleWords(text) <<endl<<endl;

cout<<"-----------------------------------------"<<endl<<endl;

}

int main()

{

string text;

setlocale(0, "");

text=LoadFile();

cout<<"Исходный текст из файла:"<<endl<<endl;

cout<< text <<endl<<endl;

ShowResult(text);

while (true)

{

cout<<"Введите текст на английском языке: ";

stringtextFromKeyboard;

getline(cin,textFromKeyboard);

if (textFromKeyboard!="")

{

cout<<endl<<"Исходныйвведенныйтекст:"<<endl<<endl;

cout<<textFromKeyboard<<endl<<endl;

ShowResult(textFromKeyboard);

}

}

system("pause");

}

**Список использованных источников.**

1. http://cppstudio.com/post/446/
2. https://proginfo.ru/getline/