CS205 C/ C++ Programming - Lab Assignment4

Name: 何宜芮(Yirui He)
SID: 11811031

Part 1 – Analysis

First, the file require us to read and load a file, called "Blocks.txt". Then load information in it in to a array which is used to store structures of blocks.

Then read the file that we need to analysis and analyze it, if one character appear one times, then we add 1 to count. Finally, we find which blocks' count is max and print the name of the block.

To analyze the file, I need to know how to use the function "utf8_to_codepoint" first. Then use the function.

Part 2 – Code

```cpp
#include <iostream>
#include <string>
#include <fstream>
#include <sstream>
#include <algorithm>
#include <vector>
#include <cmath>
#include <cstring>
#include "utf8.h"

using namespace std;

string trim(string);

char *inputString(FILE *fp, size_t size);

struct block {
    string name;
    unsigned long start;
    unsigned long end;
};

int main() {
    block blockAr[300];
    string str;
    ifstream fin("Blocks.txt");
    if (!fin.good()) {
        cout << "Fail to open the file.";
        return 0;
```

```cpp
    }
    int count = 0;
    while (getline(fin, str)) {
        str = trim(str);
        if (str.find('#') == 0)continue;
        if (str.length() == 0)continue;
        istringstream sin(str);
        string name, start, end;
        getline(sin, start, '.');
        getline(sin, end, '.');
        getline(sin, end, ';');
        getline(sin, name);
        sin.clear();


        stringstream s1, s2;
        s1 << hex << trim(start);
        s1 >> blockAr[count].start;
        s2 << hex << trim(end);
        s2 >> blockAr[count].end;
        blockAr[count].name = trim(name);
        count++;
    }
    count--;
    fin.close();


    unsigned char *p;
    int bytes_in_char;
    int i;
    unsigned long codepoint;
    unsigned long long countAr[300];
    for (int k = 0; k < 300; ++k) {//Initialize
        countAr[k] = 0;
    }
    char *input;

    input = inputString(stdin, 100);


    p = (unsigned char *) (&input[0]);
    while (*p) {
        p++;
    }
    p = (unsigned char *) (&input[0]);
```

```cpp
    while (*p) {
        codepoint = utf8_to_codepoint(p, &bytes_in_char);
        if (codepoint) {
            i = 0;
            while (codepoint > blockAr[i].end)i++;
            if (codepoint <= blockAr[i].end && blockAr[i].start <= codepoint) {
                countAr[i]++;
            }
            // p += bytes_in_char;    // Same as the line that follows
            _utf8_incr(p);
        } else {
            p++;   // Try the next character
        }
    }



    unsigned long long max = 0;
    int maxIn;
    for (int j = 0; j < 300; j++) {
//          if(countAr[j]!=0)cout << j<<": "<<countAr[j] << " ";
        if (countAr[j] > max) {
            maxIn = j;
            max = countAr[j];
        }
    }
    cout << blockAr[maxIn].name << endl;
    free(input);
}

string trim(string st) {
    while (isspace(st[0])) {
        st = st.substr(1, st.length() - 1);
    }
    while (isspace(st[st.length() - 1])) {
        st = st.substr(0, st.length() - 1);
    }
    return st;
}


//This function is found in Stackoverflow, the webside is:
//https://stackoverflow.com/questions/16870485/how-can-i-read-an-input-string-of-unknown-length
char *inputString(FILE *fp, size_t size) {
//The size is extended by the input with the value of the provisional
    char *str;
```

```c
    int ch;
    size_t len = 0;
    str = (char *) realloc(NULL, sizeof(char) * size);//size is start size
    if (!str)return str;
    while (EOF != (ch = fgetc(fp))) {//&& ch != '\n'
        str[len++] = ch;
        if (len == size) {
            str = (char *) realloc(str, sizeof(char) * (size += 16));
            if (!str)return str;
        }
    }
    str[len++] = '\0';
    return (char *) realloc(str, sizeof(char) * len);
}
```

Part 3 - Result & Verification

Test case #1:

```
hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample.txt
Armenian

hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample2.txt
Georgian

hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample3.txt
Lao

hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample4.txt
Malayalam

hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample5.txt
Devanagari

hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample6.txt
Georgian
```

The answer is correct.

Test case #2(remove the "Blocks.txt")

```
hyr@DESKTOP-B34IR65 /cygdrive/c/users/hyr/clionprojects/assignment_4/cmake-build
-debug
$ ./assignment_4<sample.txt
Fail to open the file.
```

The answer is correct. And the program terminated.

Part 4 - Difficulties & Solutions

1. First, I do not know whether I need to download the "Blocks.txt" by myself or make the program download itself. So I asked my classmates.

2. At the beginning, I do not understand what is the meaning of:

"Read a file from the standard input - that means that your program will be called like

this: ./your_program < name_of_file_to_analyze". I asked my classmates.

3. I do not know how to compile C++ and C together, so I searched first. However, I did not get the right answer. After asking SA, I know how to do it right.