

Fundamental Computer Programming- C++ Lab(I)

LAB 6

Word Processing

Week 6, Fall 2020

International Bachelor Program in Informatics
College of Informatics
Yuan Ze University

Purposes

- Get familiar with the basic structure of a C++ program
- Get familiar with some control statements such as:
 - `if (...) { ... } else { ... }`
 - `while (..) { }`
 - `for(...) { ... },`
 - `switch(...) { ... }, etc.`
- Get familiar with the use of logical operators `&&`, `||`, `!`, etc.
- Get familiar with an array
- Develop problem solving skills

switch (...) {...}

```
switch (anIE)
{
    case aCIE1: cout << "1";
    case aCIE2: cout << "2";
    //execution starts at this case label
    case aCIE3: cout << "3";
    case aCIE4: cout << "4";
        break;
    //execution of subsequent statements is stopped
    case aCIE5: cout << "5";
    default:
        cout << "The rest";
        break;
}
```

anIE: an expression evaluated into an integer.

For example,

Int x=1;

Int y=3;

char z;

x+y is evaluated into an integer.

x+z is evaluated into an integer.

x+y+'x' is evaluated into an integer.

x, y, and z themselves each are evaluated into an integer. However, all the above expressions are not a constant integral expression.

aCIE1: an expression evaluated into a constant integer value. That is, it is a constant integral expression. The same is for **aCIE2**, **aCIE3**, and **aCIE4**. For example,

const int i=2;

const char c='a';

i+c is a constant integral expression.

Both i and c each are a constant integral expression.

break & continue

- **break** is used to exit from a loop or from a switch statement
- **continue** is used to skip the statements following it till the end of the loop and continue executing the next iteration.

```
for (int j = 0; j < 2; j++) {  
    for (int k = 0; k < 5; k++) {  
        //only this loop is affected by break  
        if (k == 2) break;  
        if (k == 3) continue;  
        cout << j << k << " ";  
    }  
    //cout << j+k << endl;  
}
```

If there is a match, the statements after the matched case and that in the following cases before a break is met will be executed.

```
int i = 3; // 3 4 is printed out.  
switch (i) {  
    case 1: cout << "1";  
    case 2: cout << "2";  
        break;  
    //execution starts at this case label  
    case 3: cout << "3";  
    case 4: cout << "4";  
        break;  
    //execution of subsequent statements is terminated  
    case 5: cout << "5";  
}
```

Word Processing Using switch Statement

■ Problem description

- You are given a number of words from keyboard. A word is formed by a sequence of symbols. Symbols can be alphabets or non-alphabets. Words are separated by space, tab, or newline. A word that starts with either **a, e, i, o, u, A, E, I, O, or U** is called a word started with a vowel.
- You are asked to do the following tasks:
 - ✓ Calculate the total number of words, total number of words started with an alphabet, total number of words that do not start with a vowel but start with an alphabet, total number of words starting with a digit, 0, 1, ..., 9, and the rest.
 - ✓ Calculate the distribution of words by word length, distribution of words by the first letter, and distribution of words whose first and last letters are the same.
- Requirements
 - ✓ Use **switch(...) {...}**. However, you can use only one loop.
 - ✓ When a punctuation mark **“,”, “.” or “:”** is the last character in a word, it is not counted when calculating word length. Also, in this situation when checking whether the last and first characters are the same, you should check the character before the last one (i.e., before **“:”, “.”, or “:”**).

Input

■ Input

- Your program should accept an unknown number of words. You should use “**ctrl z**” to terminate reading input from keyboard.

■ An example for input

```
cout << "Total number of words: " << initVowelNum + nonVowelNum +  
firstCharNum[26] + firstCharNum[27] << endl;
```

This is a java program to implement monoalphabetic cypher. In cryptography, a substitution cipher is a method of encoding by which units of plaintext are replaced with ciphertext, according to a regular system; the “units” may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the above, and so forth. The receiver deciphers the text by performing an inverse substitution. 134Abc, 7sdjk, wekjf0, 23i5, Ab AB ab aB cD Ef gh IJ KL Mn op Qr St UV wx YZ Y Za ZA za zA Ab bc cd dE ef Fg Gh hi IJ JK kl LM mN No Op pq Qr Rs St TU uV vw Wx Xy YZ Za For example, eve, ada, Noun, rear, Sears, pop, kick, Sos, etc. are words whose first letter and last letter are the same.

Input example for Code::Block

```
cout << "Total number of words: " << initVowelNum + nonVowelNum + firstCharNum[26] + firstCharNum[27]
<< endl;
This is a java program to implement monoalphabetic cypher. In cryptography, a substitution cipher is a
method of encoding by which units of plaintext are replaced with ciphertext, according to a regular s
ystem; the "units" may be single letters (the most common), pairs of letters, triplets of letters, mix
tures of the above, and so forth. The receiver deciphers the text by performing an inverse substitutio
n. 134Abc, 7sdjk, wekjf0, 23i5, ... . Ab AB ab aB cD Ef gh IJ KL Mn op Qr St UV wx YZ Y Za ZA za zA
Ab bc cd dE ef Fg Gh hi IJ JK kl LM mN No Op pq Qr Rs St TU uV vw Wx Xy YZ Za For example, eve, ada, N
oun, rear, Sears, pop, kick, Sos, etc. are words whose first letter and last letter are the same.
^Z
```

Pay attention to the use of ^Z to stop reading input, where ^Z corresponds to "ctrl z" on the keyboard. ctrl means control. ^Z should be followed by an Enter.

Output⁽¹⁾

● You should print out (using the input example)

Total number of words: 157

Total number of words starting with a vowel: 44

Total number of words not starting with a vowel, but with an alphabet: 98

Total number of words started with an alphabet: 142

Total number of words started with a digit: 3

Distribution of words by word length:

L=0 0

L=1 9

L=2 66

...

...

L=14 1

L>= 15 2

L=1 means that the number of words with length equal to 1 is 9.

L>=15 2 means that the number of words with length greater than or equal to 15 is 2.

Distribution of words by first letter:

A/a: 18

B/b: 4

C/c: 8

...

...

Y/y: 3

Z/z: 5

0~9: 3

Other: 12

A/a: 18 means that the number of words whose first letter is either **A** or **a** is 14.

0~9: 3 means that the number of words with the first symbol being a digit of 0, 1, ..., and 9 is 3.

Other: 12 means that the number of words whose first symbol is non-alphabet and non-digit is 12.

Output ⁽²⁾

Distribution of words whose first and last letters are the same:

A/a: 1

B/b: 0

C/c: 0

A/1: 1 means that the number of words with the first letter **A** or **a** and last letter **A** or **a** is 1.

X/x: 0

Y/y: 0

Z/z: 0

B/b: 1 means that the number of words with the first letter **B** or **b** and last letter **B** or **b** is 1.

For example, eve, ada, Noun, rear, Sears, pop, kick, Sos, etc. are words whose first letter and last letter are the same.

Output for the Example

```
Total number of words: 157
Total number of words starting with a vowel: 44
Total number of words not starting with a vowel, but with an alphabet: 98
Total number of words started with an alphabet: 142
Total number of words started with a digit: 3
Distribution of words by word length:
L=0    1
L=1    9
L=2   66
L=3   17
L=4   13
L=5   12
L=6   10
L=7   10
L=8    5
L=9    4
L=10   2
L=11   1
L=12   4
L=13   0
L=14   1
L>=15  2
```

Output for the Example cont.

Distribution of words by the first letter:

A/a: 18
B/b: 4
C/c: 8
D/d: 2
E/e: 7
F/f: 6
G/g: 2
H/h: 1
I/i: 8
J/j: 2
K/k: 3
L/l: 7
M/m: 7
N/n: 4
O/o: 8
P/p: 6
Q/q: 2
R/r: 5
S/s: 10
T/t: 11
U/u: 3
V/v: 1
W/w: 8
X/x: 1
Y/y: 3
Z/z: 5
0~9: 3
Other: 12

Output for the Example cont.

Distribution of words whose first and last letters are the same:

A/a: 1
B/b: 0
C/c: 0
D/d: 0
E/e: 2
F/f: 0
G/g: 0
H/h: 0
I/i: 0
J/j: 0
K/k: 1
L/l: 0
M/m: 0
N/n: 1
O/o: 0
P/p: 1
Q/q: 0
R/r: 3
S/s: 2
T/t: 1
U/u: 0
V/v: 0
W/w: 0
X/x: 0
Y/y: 0
Z/z: 0

Hint: Basic structure of code

• Declaration

- An string variable to hold the input read from keyboard one at a time.
- Some int arrays to hold the distributions of words.

• A **while loop** to read the words from keyboard

- Inside the loop, use `switch(...)` {...} to count the words by their types and create the distributions by the specified attributes (特徴).

• Print out the statistics and distributions

- Can not use any loops.

C++ string and characters

- Each string is a character array. For example, if aStr is a string variable that stores a string “He loves you.”, aStr[0] will contain ‘H’, aStr[1] contains ‘e’, aStr[2] contains ‘ ’, aStr[3] contains ‘l’, aStr[4] contains ‘o’, We can use a function aStr.length() to get the length of the string stored in aStr. It will return 13.
- Each character is stored in computers using ASCII. It is an integer. So we can have the following code.

```
int aDigit, anInt;  
aDigit = '9' - '0'; // aDigit will have a value of 9.  
anInt = 'z' - 'a';  // anInt will have a value of 25.
```

You may need use `static_cast<char>(90)` to convert an integer to a character.

Follow All Requirements

- Input formats
- Output formats
- All constraints on input data, especially not accepting invalid inputs
- Coding styles
 - Avoiding using variables which do not have expressive power. That is, a variable name should carry the meaning of the matter in which the variable intends to represent.

If you don't follow the requirements, up to 30% of the points for your lab will be deduced.

Rules for Program Submission

- Put all the relevant files in the same folder.
- Name your folder SID_LabX, where ID is your student ID number and X is the number assigned to the lab. If a lab has N parts, $N > 1$, then create N sub-folders with their names SID_LabX_N in the the folder SID_LabX.
 - For example, for Lab 2 with only one part and with student ID number 1041544, the name of the folder must be S1041544_Lab2. N is omitted if there is only one part.
 - Another example, similar to the above but Lab 2 has two parts. Then, you have to create a folder S1041544_Lab2 and two sub-folders S1041544_Lab2_1 and S1041544_Lab2_2
- Compress the folder into a file named SID_LabX.zip, for example, S1041533_Lab2.zip. Then, submit the compressed file
- If you violate this rule, your lab will not be graded. If graded other penalty will be applied.