

# **Fundamental Computer Programming- C++ Lab(I)**

LAB 5

Word Processing

Week 5, Fall 2021

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.

After “continue” is executed, k++ statement will be executed first and then  $k < 5$  is checked. That is, the statements are executed in order of

①, ②, ③

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

```
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";  
}
```

What if i = 4?

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.

# LAB 5: Word Processing

## ■ Problem description

- You are given a number of words from keyboard. A word is formed by a sequence of symbols. Words are separated by space, tab, or newline. A vowel is one of **a, e, i, o, u, A, E, I, O, and U**.
- You are asked to calculate
  - ✓ total number of words which has at least a letter, total number of words whose first letter is a vowel, and total number of words which contain some letter repeated at least twice. Note that the first letter of a word may not be its first symbol.
  - ✓ Calculate the distribution of words by word length where the length of a word counts only letters. A word whose length is greater than 15 is treated as if its length is equal to 15.
  - ✓ Calculate the distribution of words by the first letter. Here, the lower case letter and upper-case letter of the same alphabet are treated as the same letter. For example A and a are treated as the same letter.

# 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, **a**ccording to **a** regular system; the “units” may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the **a**bove, **a**nd so forth. The receiver deciphers the text by performing **a**n inverse substitution. 134**A**bc, 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 key.

# Output

● You should print out something like what is shown below

# Total number of words which has at least a letter: 189

# Total number of words whose first letter is a vowel: 49

# Total number of words which contain some letter repeated at least twice: 67

# Distribution of words by word length:

# L0 = 0

# L1 = 9

# L2 = 66

...

...

# L14 = 1

# L15 = 2

L1 = 9 means that the number of words with length equal to 1 is 9.

L15 = 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

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



# Input Example

```
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.

# Output for the Example

```
# Total number of words which has at least a letter: 148
# Total number of words whose first letter is a vowel: 47
# Total number of words which contain some letter repeated at least twice: 41
# Distribution of words by word length:
# L0 = 7
# L1 = 6
# L2 = 62
# L3 = 19
# L4 = 13
# L5 = 13
# L6 = 9
# L7 = 7
# L8 = 5
# L9 = 4
# L10 = 2
# L11 = 1
# L12 = 6
# L13 = 0
# L14 = 1
# L15 = 0
```

# Output for the Example cont.

```
# Distribution of words by the first letter:  
# a/A: 19  
# b/B: 4  
# c/C: 8  
# d/D: 2  
# e/E: 7  
# f/F: 6  
# g/G: 2  
# h/H: 1  
# i/I: 9  
# 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: 11  
# t/T: 13  
# u/U: 4  
# v/V: 1  
# w/W: 8  
# x/X: 1  
# y/Y: 3  
# z/Z: 5
```

# 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**
- **Print out the statistics and distributions**

# 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
- 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.