

Лабораторна робота № 9-10

Створення програм для обробки текстових рядків

Мета: навчитись створювати текстові рядки, зчитувати інформацію з рядків, виконувати операції із рядками

Теоретичні відомості

Рядок у C++ - це послідовність байт (масив символів, величин типу *char*), що завершується в кінці спеціальною ознакою - байтом '\0'. Тобто в мові C++ використовується концепція «zero termination string». Це означає, що для позначення закінчення кожного рядка (масиву типу *char*) використовується нуль-символ ('\0').

Наприклад:

```
char err_str1[6] = "Error";  
char err_str2[6] = {'E', 'r', 'r', 'o', 'r', '\0'};
```

в даному прикладі оголошення рядків будуть ідентичні:

```
err_str1[0] = err_str2[0] = 'E';  
err_str1[1] = err_str2[1] = 'r';  
err_str1[2] = err_str2[2] = 'r';  
err_str1[3] = err_str2[3] = 'o';  
err_str1[4] = err_str2[4] = 'r';  
err_str1[5] = err_str2[5] = '\0';
```

Для друку цілого рядка за допомогою функції *printf* використовують ключ *%s*. Ніякий інший ключ не дозволить надрукувати рядок цілком (тобто до тих пір поки не зустрінеться символ закінчення рядка '\0'). Рядок так само може розглядатися як масив символів і друкуватися поелементно за допомогою ключа *%c*.

Наприклад:

```
char i = 0;  
char err_str1[6] = "Error";  
printf("%s\n", err_str1);  
while(err_str1[i]) printf("%c", err_str1[i++]);
```

Часто при оголошенні рядка з ініціалізацією можна не вказувати розмір рядка. При наявності ініціалізації розмір рядка буде підрахований автоматично - він дорівнює кількості символів у константі +1 (для зберігання нуль символу).

Наприклад:

```
char err_str1[] = "Error";
```

У цьому прикладі розмірність масиву *err_str1* буде дорівнювати $5 + 1 = 6$.

Примітка. Підключення функція для відображення символів кирилиці в режимі консольного виведення інформації:

1) Підключення стандартних бібліотек:

```
# include "locale"
# include "windows.h"
// заголовок програми int_main...
```

2) Підключення функцій кодування:

```
SetConsoleCP(1251);
SetConsoleOutputCP(1251);
```

Стандартна бібліотека мови C++ має багато функцій для обробки символів (табл. 1) та рядків (табл. 2).

Таблиця 1 – Функції для роботи із символами (<ctype.h>)

int isalnum (int c);	Перевірка, чи є символ літерою або цифрою
int isalpha (int c);	Перевірка, чи є символ літерою
int iscntrl (int c);	Перевірка, чи є символ керуючим
int isdigit (int c);	Перевірка, чи є символ десятковою цифрою
int isgraph (int c);	Перевірка, чи є символ видимим
int isprint (int c);	Перевірка, чи є символ видимим, включаючи пробіл
int islower (int c);	Перевірка, чи є символ літерою нижнього регістру
int ispunct (int c);	Перевірка, чи є символ знаком пунктуації
int isspace (int c);	Перевірка, чи є символ пробільним
int isupper (int c);	Перевірка, чи є символ літерою верхнього регістру
int isxdigit (int c);	Перевірка, чи є символ шістнадцятковою цифрою
int tolower (int c);	Перетворення символу в нижній регістр
int toupper (int c);	Перетворення символу у верхній регістр

Таблиця 2 – Функції бібліотеки введення/виведення (<stdio.h>)

getchar (void)	Вводить наступний символ зі стандартного пристрою введення і повертає його в форматі цілого.
char *gets (char *s)	Вводить символи зі стандартного пристрою введення в масив s до тих пір, поки не зустріне символ кінця рядка або індикатор кінця файлу.
int putchar (int c)	Виводить символ, який зберігається в с.
int puts (const char *s)	Виводить рядок s та перехід на наступний рядок
int sprintf (char *buf, const char *format [,arg1, ...])	Виконує форматоване виведення у рядок buf. Параметр format задає спосіб відображення значень змінних arg1, Дія функції sprintf аналогічна дії функції printf, але виведення виконується в рядок-буфер, а не на екран.
int sscanf (const char *s, const char *format [, address1, ...])	Виконує дії, еквівалентні scanf за винятком того, що введення здійснюється з масиву s, а не з клавіатури.

Рядкові функції працюють з масивами символів (рядками), що закінчуються символом кінця рядка. У мові C++ для роботи з рядковими функціями використовується заголовок `<string.h>`. У ньому визначено тип `size_t` – тип результату, який утворюється після застосування оператора `sizeof()` і являє собою різновид цілого без знака. У таблиці 3 наведено найбільш популярні і підтримувані більшістю компіляторів функції.

Таблиця 3 – Функції роботи з рядками (`<string.h>`)

Функція	Заголовок	Опис
<code>strcpy(s1,s2)</code>	<code><cstring></code>	Копіювання рядка <code>s2</code> в рядок <code>s1</code>
<code>strcat(s1,s2)</code>	<code><cstring></code>	Рядок <code>s2</code> приєднуємо до рядку <code>s1</code>
<code>strcmp(s1,s2)</code>	<code><cstring></code>	Порівняння рядків <code>s1</code> та <code>s2</code> : якщо рядки рівні, повертаємо значення 0
<code>strchr(s,ch)</code>	<code><cstring></code>	Вказівник на першу позицію символу <code>ch</code> у рядку <code>s</code>
<code>strstr(s1,s2)</code>	<code><cstring></code>	Вказівник на першу позицію підрядка <code>s2</code> у рядку <code>s1</code>
<code>atoi(s)</code>	<code><cstdlib></code>	Перетворення рядка з цифр <code>s</code> в ціле число типу <code>int</code>
<code>atol(s)</code>	<code><cstdlib></code>	Перетворення рядка з цифр <code>s</code> в ціле число типу <code>long</code>
<code>atof(s)</code>	<code><cstdlib></code>	Перетворення рядка з цифр <code>s</code> в дійсне число типу <code>double</code>
<code>tolower(ch)</code>	<code><cctype></code>	Перетворення рядка літер <code>ch</code> до рядкового формату
<code>toupper(ch)</code>	<code><cctype></code>	Перетворення рядка літер <code>ch</code> до прописного формату

Приклад розробки програми для роботи з рядками і символами

Задача. З клавіатури вводиться текстовий рядок. Скласти програму, яка:
а) інвертує рядок, подаючи його у зворотному вигляді; б) підраховує кількість чисел у тексті; в) видаляє всі слова, що починаються з цифри.

Формалізація задачі

- Вхідний рядок S представимо у вигляді масиву символів.
- Підраховуємо n – кількість символів у рядку.
- Далі програму розділимо на три послідовних частини.
 - Отримання інвертованого рядка:
 - підготуємо рядок $S1$ такого самого розміру n ;
 - здійснимо в циклі ($i=0, \dots, n-1$) присвоєння:
 - $S1[i] = S[n-i-1]$. (тут враховуємо те, що останній символ рядка S дорівнює `'\0'`);
 - додаємо в масив $S1$ символ закінчення (щоб обірвати рядок).
 - Виділення окремих слів.
 - запам'ятовуємо S в $S1$ (щоб не зіпсувати початковий рядок);
 - переглядаємо рядок S по словах (виділяємо частину рядка до пробілу). Це доцільно зробити в циклі.
 - Видалення слів, що починаються з цифри:
 - відновлюємо S з $S1$;

- готуємо рядок S2, в якому будемо зберігати ті слова, що не починаються з цифри;
- переглядаємо рядок S по словах (виділяємо частину рядка до пробілу). Якщо слово починається з цифри, не додаємо його до S2, в іншому випадку – додаємо;
- додаємо в масив S2 символ закінчення (щоб обірвати рядок).

Лістинг програми роботи з рядками і символами

```
#include <stdio.h>
#include <iostream>
#include <string.h>
#include <ctype.h>
#include <locale.h>
// #include <windows.h>

using namespace std;
int main()
{
    setlocale (LC_ALL, "Rus");
    // SetConsoleCP(1251);
    // SetConsoleOutputCP(1251);
    cout << "\n Input String: ";
    char s[80];

    cin.getline (s, 80);
    int n=strlen(s); // довжина рядка
    char *s1 = new char[n];
    for (int i=0; i<n; i++)// перевертаємо рядок
        s1[i]=s[strlen(s)-i-1];
    s1[n]='\0';
    cout << "\n Reverse String: " << s1;
    // виділяємо слова у реченні
    strcpy(s1,s); // запам'ятовуємо початкове речення
    cout << "\n\n Words in string: \n";
    char *word;
    word = strtok(s, " ");
    while (word != NULL)
    {
        printf("\n\t%s",word);
        word=strtok(NULL, " ");
    }
    // ще раз виділяємо слова, але не виводимо ті, що починаються з цифри
    strcpy(s,s1); // відновлюємо початкове речення
    cout << "\n\n String without first digits: \n";
    char *s2 = new char[n];
    strcpy(s2,"");
    word = strtok(s, " ");
    while (word != NULL)
```

```

{
    if (strpbrk(word, "0123456789")==NULL)
        strcat(s2, word);
    strcat(s2, " ");
    word=strtok(NULL, " ");
}
strcat(s2, "\\0");
printf("\\n\\t%s", s2);
delete[] s1;
delete[] s2;

return 0;
}

```

Хід роботи

Розв'язати задачі в консольному режимі в середовищі програми Code::Blocks (Microsoft Visual Studio).

Задача 1. Ввести з клавіатури символний рядок і замінити в ньому всі букви «а» на «б» і всі букви «б» на «а» (великі на великі, малі на малі).

Приклад:

Введіть рядок:

ааббААББссСС

Результат:

ббааББААссСС

Задача 2. Ввести з клавіатури символний рядок і визначити, скільки в ньому слів. Словом вважається послідовність непробільних символів, відокремлена з двох сторін пробілами (або стоїть з краю рядка). Слова не можуть бути розділені декількома пробілами, на початку і в кінці рядка теж не можуть бути пробіли.

Задача 3. Ввести з клавіатури символний рядок і знайти найдовше слово і його довжину. Словом вважається послідовність непробільних символів, відокремлена з двох сторін пробілами (або стоїть з краю рядка). Слова можуть бути розділені декількома пробілами, на початку і в кінці рядка теж можуть бути пробіли.

Задача 4. Ввести з клавіатури в один рядок прізвище, ім'я та по батькові, розділивши їх пробілом. Вивести прізвище та ініціали.

Приклад:

Введіть прізвище, ім'я та по батькові:

Іванов Петро Семенович

П.С. Іванов

Задача 5. Ввести адресу файла і «розібрати» його на частини, розділені знаком '/'. Кожну частину вивести в окремому рядку.

Приклад:

Введіть адресу файла:

C:/Фото/2013/Похід/vasya.jpg

C:

Фото

2013

Похід

vasya.jpg

Задача 6. Напишіть програму, яка замінює у всьому рядку одну послідовність символів на іншу.

Приклад:

Введіть рядок:

(X > 0) and (Y < X) and (Z > Y) and (Z <> 5)

Що міняємо : and

Чим замінити : &

Результат

(X > 0) & (Y < X) & (Z > Y) & (Z <> 5)

Задача 7. Напишіть програму, яка обчислює вираз, що складається з трьох чисел і двох знаків (допускаються тільки знаки «+» або «-»). Вираз вводиться як символьний рядок, всі числа цілі.

Приклад:

Введіть вираз :

12-3+45

Відповідь: 54

Задача 8. Напишіть програму, яка знаходить середнє арифметичне всіх чисел, записаних у файлі в стовпчик, і виводить результат в інший файл.

Задача 9. Напишіть програму, яка знаходить мінімальне і максимальне серед парних позитивних чисел, записаних у файлі, і виводить результат в інший файл. Врахуйте, що таких чисел може взагалі не бути.

Задача 10. Напишіть програму, яка обчислює вираз, що складається з трьох чисел і двох знаків (допускаються знаки «+», «-», «*» і «/»). Вираз вводиться як символьний рядок, всі числа цілі. Операція «/» виконується як цілочисельне ділення.

Приклад:

Введіть вираз :

12*3+45

Відповідь: 81

Контрольні питання

1. Який заголовочний файл використовується у мові C++ для роботи з символами?
2. Наведіть приклади використання функцій обробки символів.
3. Які функції існують для введення і виведення символів?
4. Наведіть основні функції введення\виведення рядків і їх призначення. Який заголовочний файл містить опис цих функцій?
5. Чим відрізняються різні функції введення\виведення рядків? Наведіть приклади їх використання.
6. Який заголовочний файл використовується у мові C для обробки рядків?
7. Назвіть основні функції роботи з рядками.
8. Як можна оголосити рядок? Наведіть різні способи оголошення та ініціалізації рядків.
9. Для чого у рядках використовується завершуючий символ?
10. Як отримати довжину рядка?
11. Як компілятор мови C контролює вихід за межі масиву символів?
12. Як можна запрограмувати виділення слів у введенному реченні?
13. Яким чином можна об'єднати рядки символів та виділити підрядки за певними ознаками? Наведіть приклади.