

Структура на програма в С

Част 9 - низове от символи, C-string

Иван Георгиев, Христо Иванов, Христо Стефанов

Технологично училище "Електронни системи",
Технически университет, София

15 май 2019 г.

- 1 Проблем
- 2 Низове от символи. C string

1 Проблем

2 Низове от символи. C string

Проблем

- При обработка на масиви чрез функции, (почти) винаги е нужно да се предава големината на масива в елементи заедно с указател към първия елемент от масива
- Това създава неудобство за програмиста, тъй като непрекъснато трябва да се предава допълнителен аргумент

```
/* prints up to 'size' elements in the array pointed by 'a' */
int print_array(int *a, int size) {
    // ...
}
/* copies the array pointed by 'a' into 'b' element by element */
int copy_array(int *a, int size_a, int *b, int size_b) {
    // ...
}
int main() {
    int a[4] = {10, 20, 30, 40}; int b[5] = {0};
    print_array(a, 4); copy_array(a, 4, b, 5); print_array(b, 5);
    // wishing for: print_array(a); copy_array(a, b); print_array(b);
}
```

- При обработване на текст в C, текста най-често се запазва в масив от тип 'char', като поредният елемент от масива съдържа кода от ASCII таблицата на поредния символ от текста

```
int capitalize_and_print(char *t, int size) {
    int i = 0;
    if (t[0] >= 'a' && t[0] <= 'z') {
        t[0] = t[i] - ('a' - 'A');
    }
    for (i = 0; i < size; i++) {
        printf("%c", t[i]);
    }
}

int main() {
    char text[7] = {'h', 'e', 'l', 'l', 'o', '!', '\n'};
    capitalize_and_print(text, 7);
}
```

Фрагмент 1: Пример за обработка на текст

1 Проблем

2 Низове от символи. C string

- Когато се работи със текстове, проблемът може да бъде решен по следния начин:
- ASCII таблицата съдържа специален символ наречен *нулев символ* (*null character*), който има код 0. Той няма графично представяне и съответно никога не може да бъде срещнат в текст
- В C е прието с нулев символ да се означава края на текста в един масив, което помага за обработването на текста в масива, тъй като масива може да бъде обработван докато не се стигне до нулевия символ
- Масив с елементи от тип 'char', съдържащ някъде в себе си кода на нулевия символ, се казва че съдържа C string - т.е. последователност (низ) от символи завършваща с 0
- В C символната константа, която съответства на кода на нулевия символ в ASCII таблицата, е '\0'

```
int capitalize_and_print(char *t) {
    int i = 0;
    if (t[0] >= 'a' && t[0] <= 'z') {
        t[0] = t[i] - ('a' - 'A');
    }
    for (i = 0; t[i] != '\0'; i++) {
        printf("%c", t[i]);
    }
}

int main() {
    char text[] = {'h', 'e', 'l', 'l', 'o', '!', '\n', '\0'};
    char text2[] = {'b', 'y', 'e', '!', '\n', '\0'};
    capitalize_and_print(text);
    capitalize_and_print(text2);
}
```

Фрагмент 2: Пример за обработка на текст записан като C string

Низове от символи. C string

- Езикът C има специална поддръжка за C string-ове, която цели да улесни използването им
- При инициализация на масив с елементи от тип 'char' е възможно да се използва низов литерал, като съответно поредният елемент от масива се инициализира с кода на поредния символ от низовия литерал
- Въпреки че не се изписва, низовият литерал винаги съдържа като последен символ нулевият символ - '\0'

```
int main() {  
    char a[] = {'h', 'e', 'l', 'l', 'o', '!', '\n', '\0'};  
    char b[] = "hello!\n"; /* 'b' is initialized in the same way as 'a',  
                           both 'a' and 'b' are C strings because  
                           they contain '\0' */  
}
```

Фрагмент 3: Стандартна инициализация на масив от символи и еквивалентно инициализиране с низов литерал

Низове от символи. C string

- Когато се използва извън инициализация на масив, низовият литерал се замества със стойност, която е адреса на първият елемент от масив, съдържащ символите от низовия литерал

```
int print(char *str) {
    int i = 0;
    for (i = 0; str[i] != '\0'; i++)
        printf("%c", str[i]);
    return 0;
}
int main() {
    char a[] = "hello!\n";
    char *ptr = "hello!\n";
    print(a); // prints 'hello\n'
    print(ptr); // prints 'hello\n'
    print("hello!\n"); // prints 'hello\n'
}
```

Фрагмент 4: Употреба на низов литерал в и извън инициализация на масив