**МИНОБРНАУКИ РОССИИ**

**Санкт-Петербургский государственный**

**электротехнический университет**

**«ЛЭТИ» им. В.И. Ульянова (Ленина)**

**Кафедра МО ЭВМ**

отчет

**по лабораторной работе №1**

**по дисциплине «Программирование»**

**Тема: Регулярные выражения**

|  |  |  |
| --- | --- | --- |
| Студент гр. 3342 |  | Малахов А.И. |
| Преподаватель |  | Глазунов С.А. |

Санкт-Петербург

2024

## Цель работы

Изучение и работа с регулярными выражениями и группами, а также использование их в языке программирования Си.

## Задание

Вариант 1.

На вход программе подается текст, представляющий собой набор предложений с новой строки. Текст заканчивается предложением "Fin." В тексте могут встречаться ссылки на различные файлы в сети интернет. Требуется, используя регулярные выражения, найти все эти ссылки в тексте и вывести на экран пары <название сайта> - <имя файла>. Гарантируется, что если предложение содержит какой-то пример ссылки, то после ссылки будет символ переноса строки.

Ссылки могут иметь следующий вид:

* Могут начинаться с названия протокола, состоящего из букв и **://** после
* Перед доменным именем сайта может быть www
* Далее доменное имя сайта и один или несколько доменов более верхнего уровня
* Далее возможно путь к файлу на сервере
* И, наконец, имя файла с расширением.

## Выполнение работы

В начале работы программы компилируется регулярное выражение. После этого, используя цикл, перебираются все предложения до тех пор, пока не встретится «Fin.».

В каждой итерации цикла идет считывание строки, которая далее проверяется на соответствие регулярному выражению. Если строка подходит, то запускается цикл, который проходится по всем требуемым группам, и выводит результат на экран в формате <название сайта> - <имя файла>.

В конце происходит очистка памяти, выделенной под регулярное выражение.

Разработанный программный код см. в приложении А.

## Тестирование

Результаты тестирования представлены в табл. 1.

Таблица 1 – Результаты тестирования

|  |  |  |
| --- | --- | --- |
| № п/п | Входные данные | Выходные данные |
|  | This is simple url:  http://www.google.com/track.mp3  May be more than one upper level  domain http://www.google.com.edu/hello.avi  Many of them.  Rly. Look at this!  http://www.qwe.edu.etu.yahooo.org.net.ru/qwe.q  Some other protocols  ftp://skype.com/qqwe/qweqw/qwe.avi  Fin. | google.com - track.mp3  google.com.edu - hello.avi  qwe.edu.etu.yahooo.org.net.ru - qwe.q  skype.com - qwe.avi |

## Выводы

Были изучены и применены на практике регулярные выражения в языке программирования Си. Была написана программа, которая находит в тексте все ссылки на различные файлы и выводит их.

# Приложение А Исходный код программы

Название файла: main.c

#include <stdio.h>

#include <string.h>

#include <regex.h>

#define PATTERN "([a-zA-Z]+://)?(www\\.)?([a-zA-Z0-9-]+(\\.[a-zA-Z0-9]+)+)/((\\w+/)\*)([a-zA-Z0-9-]+(\\.[a-zA-Z0-9-]+))\*\n$"

#define MAX\_GROUP 9

#define END\_OF\_TEXT "Fin."

void print\_match(char \*sentence, regmatch\_t groupArray[], int groupIndex)

{

for (int i = groupArray[groupIndex].rm\_so; i < groupArray[groupIndex].rm\_eo; i++)

{

printf("%c", sentence[i]);

}

}

int main(){

regex\_t regex;

char sentence [100];

regmatch\_t groupArray[MAX\_GROUP];

regcomp(&regex, PATTERN, REG\_EXTENDED);

while (fgets(sentence, 100, stdin) != NULL && strstr(sentence, END\_OF\_TEXT) == NULL) {

if(regexec(&regex, sentence, MAX\_GROUP, groupArray, 0) == 0){

print\_match(sentence, groupArray, 3);

printf(" - ");

print\_match(sentence, groupArray, 7);

printf("\n");

}

}

regfree(&regex);

return 0;

}