



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ _____ «Информатика и системы управления»

КАФЕДРА _____ «Теоретическая информатика и компьютерные технологии»

Лабораторная работа №12
по курсу «Языки и методы программирования»
«Обработка текстовых файлов»

Студент группы ИУ9-22Б Лавров Р. Д.

Преподаватель Посевин Д. П.

Москва 2025

1 Задание

Выриант 11: Найти все файлы с расширением «txt» в указанном каталоге, в тексте каждого файла найти вхождения целых десятичных чисел и сформировать в текущем каталоге файл sums.txt, в котором будут перечислены суммы найденных чисел в каждом файле. Каждая строка файла sums.txt должна содержать имя файла и сумму чисел в нём. Имена файлов должны быть отсортированы лексикографически. Работоспособность программы нужно проверить на наборе предварительно созданных текстовых файлов.

Выриант 12: Найти все файлы с расширением «html» в указанном каталоге, для каждого файла определить множество гиперссылок и сохранить объединение полученных множеств в файле links.txt в текущем каталоге. Каждая гиперссылка в сформированном файле должна располагаться в отдельной строке. Гиперссылки должны быть отсортированы лексикографически.

Гиперссылка в HTML-файле задаётся тегом «a»:

Работоспособность программы нужно проверить на наборе HTML-файлов, загруженных из интернета.

2 Результаты

Исходный код 1– 3.

Листинг 1 — var1.cpp (11 вариант)

```
1 #include <sys/types.h>
2 #include <sys/stat.h>
3 #include <unistd.h>
4 #include <iostream>
5 #include <dirent.h>
6 #include <fstream>
7
8 using namespace std;
```

Листинг 2 — var1.cpp (11 вариант) (продолжение)

```
1 int sumInFile(string name) {
2     ifstream f;
3     int sum = 0;
4     f.open(name);
5     string line;
6     while (getline(f, line)){
7         sum += stoi(line);
8     }
9     return sum;
10 }
11
12 void writeInFile(vector<pair<string, int>> data) {
13     ofstream f;
14     f.open("testVar1/sum.txt");
15     for (int i = 0; i < data.size(); i++) {
16         f << data[i].first << " " << data[i].second << endl;
17     }
18 }
19
20 int main(int argc, char** argv) {
21     vector<pair<string, int>> ans;
22
23     DIR* mainFolder = opendir(argv[1]);
24     if (mainFolder == NULL) {
25         cout << "wrong folder" << endl;
26         return 0;
27     }
28
29     dirent* curDir;
30     while ((curDir = readdir(mainFolder)) != NULL) {
31         string nameFile = curDir->d_name;
32         if (nameFile.size() > 4 && nameFile.substr(nameFile.size() - 3, nameFile.size() - 1)
33             == "txt") {
34             pair<string, int> p = {nameFile, sumInFile(string(argv[1]) + "/" + nameFile)};
35             ans.push_back(p);
36         }
37     }
38
39     sort(ans.begin(), ans.end());
40     writeInFile(ans);
41
42     return 0;
43 }
```

Листинг 3: var2.cpp (12 вариант)

```

1 #include <sys/types.h>
2 #include <sys/stat.h>
3 #include <unistd.h>
4 #include <iostream>
5 #include <dirent.h>
6 #include <fstream>
7
8 using namespace std;
9
10 string getLinkFromStr(string str) {
11     bool linkStart = false;
12     string ans = "";
13     for (int i = 0; i < str.size(); i++) {
14         if (str[i] == '"' || str[i] == '\\') {
15             if (linkStart) {
16                 break;
17             } else {
18                 linkStart = true;
19                 continue;
20             }
21         }
22         if (linkStart) {
23             ans += str[i];
24         }
25     }
26     return ans;
27 }
28
29 void getLinks(vector<string> *links, string name) {
30     ifstream f;
31     f.open(name);
32     string line;
33     while (getline(f, line)) {
34         if (line.find("href") != string::npos) {
35             links->push_back(getLinkFromStr(line));
36         }
37     }
38 }
39
40
41 void writeInFile(vector<string> data) {
42     ofstream f;
43     f.open("testVar2/links.txt");
44     for (int i = 0; i < data.size(); i++) {
45         f << data[i] << endl;

```

```

46     }
47 }
48
49 int main(int argc, char** argv) {
50     DIR* mainFolder = opendir(argv[1]);
51     if (mainFolder == NULL) {
52         cout << "wrong folder" << endl;
53         return 0;
54     }
55
56     dirent* curDir;
57     vector<string> links;
58     while ((curDir = readdir(mainFolder)) != NULL) {
59         string nameFile = curDir->d_name;
60         if (nameFile.size() > 5 && nameFile.substr(nameFile.size() - 4, nameFile.size() - 1)
61             == "html") {
62             getLinks(&links, string(argv[1]) + "/" + nameFile);
63         }
64     }
65
66     sort(links.begin(), links.end());
67     writeInFile(links);
68
69     return 0;
70 }

```

3 Вывод

Я научился работать с файлами C++