

Теория Формальных Языков

Лабораторная работа №4

Вариант 14

Мамаев А. А.

ИУ9-52Б

Цель лабораторной работы.

Научиться использовать классы и функции библиотеки регулярных выражений языка C++ для анализа текстовых файлов.

Порядок выполнения лабораторной работы.

1. Подготовить текст (более десяти страниц), в котором должны быть фрагменты, удовлетворяющие шаблону (приложение 3).
2. Написать программу на языке C++, которая формирует XML-файл заданного в приложении 3 формата.

Приложение 3

Шаблон:

Выделить из текста предложения со словами, удовлетворяющими следующим паттернам: *Москв**, *Нижн*-Новгород**.

Сформировать XML-файл заданного формата.

Формат XML-файла:

<t>предложение с паттерном<\t>

Результат выполнения лабораторной работы.

Листинг 1

```
#include <iostream>
#include <fstream>
#include <regex>
#include <map>

using namespace std;

string read_file(const string & path) {
    ifstream source_stream(path);
    ostringstream string_stream;
    string_stream << source_stream.rdbuf();
    source_stream.close();
    return string_stream.str();
}

vector < string > get_cities_sentences(const string & source) {
    vector < string > cities_sentences;
    regex cities_regex(R "([^\n\t\.\?!]*([Нижн[а-я]+
Новгород|Москв)[^\n\t\.\?!]*[\.\?!])", regex::optimize);
    sre_iterator regex_it(source.begin(), source.end(), cities_regex),
end;
    for (; regex_it != end; regex_it++) {
        smatch math_res = * regex_it;
        cities_sentences.push_back(math_res.str());
        cout << math_res.str() << endl;
    }
    return cities_sentences;
}
```

```

}

void write_sentences(const string & path,
    const vector < string > & sentences) {
    ofstream output_stream(path);
    output_stream << "<cities_sentences>" << endl << endl;
    for (const string & s: sentences) {
        output_stream << "\t<t>" << s << "</t>" << endl << endl;
    }
    output_stream << "</cities_sentences>" << endl;
    output_stream.close();
}

int main() {
    string input_path = "/home/alexey/TFL/soloview.txt";
    string output_path = "/home/alexey/TFL/output.txt";

    string input = read_file(input_path);
    write_sentences(output_path, get_cities_sentences(input));

    return 0;
}

```