

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО
ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНОЙ
ТЕХНИКИ

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 3
курса «Информатика»
по теме: «Регулярные выражения»
Вариант № 335047

Выполнил студент:
Тюрин Иван Николаевич
группа: Р3110

Преподаватель:
Балакшин П. В.,
Рудникова Т. В.

Санкт-Петербург, 2021 г.

Содержание

Лабораторная работа № 3 Регулярные выражения	2
1. Задание варианта № 335047	2
2. Выполнение задания 1 и 2	4
3. Выполнение обязательного задания	4
4. Выполнение дополнительного задания № 1	5
5. Выполнение дополнительного задания № 2	5
6. Выполнение дополнительного задания № 3	5
7. Выполнение дополнительного задания № 4	5
8. Вывод	6

Лабораторная работа № 3

Регулярные выражения

1. Задание варианта № 335047

, , ,

1. Понять устройство страницы с расписанием для своей группы:
<https://itmo.ru/ru/schedule/0/P3110/schedule.htm>
2. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного.
3. Обязательное задание (позволяет набрать до 65 процентов от максимального числа баллов БаРС за данную лабораторную): написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.
4. Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов.
5. Дополнительное задание задание №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
 - а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
 - б) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.
 - с) Сравнить полученные результаты и объяснить их сходство/различие.
6. Дополнительное задание задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

- a) Переписать исходный код, добавив в него использование регулярных выражений.
 - b) Сравнить полученные результаты и объяснить их сходство/различие.
7. Дополнительное задание задание №3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
- a) Используя свою исходную программу из обязательного задания, программу из дополнительного задания №1 и программу из дополнительного задания №2, сравнить десятикратное время выполнения парсинга + конвертации в цикле.
 - b) Проанализировать полученные результаты и объяснить их сходство/различие.
8. Дополнительное задание задание №4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).
- a) Переписать исходную, чтобы она осуществляла парсинг и конвертацию исходного файла в любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.
 - b) Проанализировать полученные результаты, объяснить особенности использованного формата.

‘ ‘ ‘

2. Выполнение задания 1 и 2

Были просмотрена структура веб-страницы с расписанием моей группы, она представляет собой текстовый документ, в формате HTML.

Было составлено собственное расписание пар в среду в формате YAML. Текстовый документ размещен в личном удаленном репозитории. Содержимое этого текстового документа можно найти по ссылке: <https://github.com/elturin/itmo-informatics/blob/main/lab-4/assets/timetable.yaml>.

3. Выполнение обязательного задания

Была написана программа на языке программирования Python, которая парсит исходный файл с расписанием (составленный самостоятельно) в формате YAML, строит соответствующую структуру данных внутри ЯП, конвертирует эту структуру в формат XML, при этом имеется возможность вывести полученный текстовый документ в стандартный поток вывода (консоль) или записать в файл.

Исходный код программы размещен в личном удаленном репозитории. Код главного скрипта `main.py` можно найти по ссылке:

<https://github.com/elturin/itmo-informatics/blob/main/lab-4/main.py>.

Для его работы были написаны мини-библиотеки: `YAML.py` для работы с YAML-файлом, `XML.py` для работы с XML-файлом, `exceptions.py` для обработки ошибок пользователя.

Исходный код библиотек можно найти по ссылкам:

- `YAML.py` — <https://github.com/elturin/itmo-informatics/blob/main/lab-4/src/YAML.py>
- `XML.py` — <https://github.com/elturin/itmo-informatics/blob/main/lab-4/src/XML.py>
- `exceptions.py` — <https://github.com/elturin/itmo-informatics/blob/main/lab-4/src/exceptions.py>

4. Выполнение дополнительного задания № 1

Была написана программа на языке программирования Python, которая выполняет конвертацию YAML-файла в XML-файл с использованием готовых библиотек для ЯП Python.

Исходный код программы размещен в личном удаленном репозитории. Код можно найти по ссылке: <https://github.com/elturin/itmo-informatics/blob/main/lab-4/main1.py>.

5. Выполнение дополнительного задания № 2

Была написана программа на языке программирования Python, которая выполняет конвертацию YAML-файла в XML-файл без использования готовых библиотек для этого, но с использованием регулярных выражений.

Исходный код программы размещен в личном удаленном репозитории. Код можно найти по ссылке: <https://github.com/elturin/itmo-informatics/blob/main/lab-4/main2.py>.

6. Выполнение дополнительного задания № 3

Была написана программа на языке программирования Python, которая сравнивает скорость выполнения программы написанной в качестве обязательного задания (см. с. 4) и программы написанной в качестве дополнительного задания № 2 (см. с. 5)

Исходный код программы размещен в личном удаленном репозитории. Код можно найти по ссылке: <https://github.com/elturin/itmo-informatics/blob/main/lab-4/main3.py>.

7. Выполнение дополнительного задания № 4

Была написана программа на языке программирования Python, которая выполняет конвертацию YAML-файла в BSON-файл (Бинарный JSON) с использованием готовых библиотек для ЯП Python.

Исходный код программы размещен в личном удаленном репозитории. Код можно найти по ссылке: <https://github.com/e1turin/itmo-informatics/blob/main/lab-4/main4.py>.

8. Вывод

Изучил форматы хранения данных и языки разметки документов. Научился работать с библиотеками ЯП Python для YAML, XML. Укрепил знания по работе с Python.