

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ «Санкт-Петербургский национальный исследовательский
университет ИТМО»

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНОЙ
ТЕХНИКИ

ЛАБОРАТОРНАЯ РАБОТА №4

по дисциплине

«ИНФОРМАТИКА»

Вариант № 9

Выполнил:

Студент группы Р3116

Билошицкий Михаил Владимирович

Преподаватель:

Машина Екатерина Алексеевна

Санкт-Петербург, 2022

Содержание

Задание.....	3
Основные этапы вычисления	5
Вывод.....	6
Список использованных источников.....	7

Задание

1. Определить номер варианта как остаток деления на 36 порядкового номера в списке группы в ISU. В случае, если в данный день недели нет занятий, то увеличить номер варианта на восемь.
2. Изучить форму Бэкуса-Наура.
3. Изучить особенности языков разметки/форматов JSON, YAML, XML.
4. Понять устройство страницы с расписанием для своей группы:
<http://itmo.ru/ru/schedule/0/P3110/schedule.htm>
5. Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного. При этом необходимо, чтобы в выбранном дне было не менее двух занятий (можно использовать своё персональное). В случае, если в данный день недели нет таких занятий, то увеличить номер варианта ещё на восемь.
6. 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.
7. Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов.
8. Дополнительное задание No1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
 1. а) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.
 2. б) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.
 3. с) Сравнить полученные результаты и объяснить их сходство/различие.
9. Дополнительное задание No2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
 - а) Переписать исходный код, добавив в него использование регулярных выражений.
 - б) Сравнить полученные результаты и объяснить их сходство/различие.
10. Дополнительное задание No3 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).
 1. а) Используя свою исходную программу из обязательного задания, программу из дополнительного задания No1 и программу из дополнительного задания No2, сравнить стократное время выполнения парсинга + конвертации в цикле.

2. б) Проанализировать полученные результаты и объяснить их сходство/различие.

11.Дополнительное задание No4 (позволяет набрать +5 процентов от максимального числа баллов БаРС за данную лабораторную).

с) Переписать исходную программу, чтобы она осуществляла парсинг и конвертацию исходного файла в любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.

d) Проанализировать полученные результаты, объяснить особенности использования формата.

12.Проверить, что все пункты задания выполнены и выполнены верно.

13.Написать отчёт о проделанной работе.

14.Подготовиться к устным вопросам на защите.

Основные этапы вычисления

Выполненная программа состоит из следующих модулей.

`main.py` – модуль, в котором используются все остальные модули для конвертации, а также замерка времени 100 кратного исполнения.

`convertors/yaml_json` – модуль для конвертации файла `yaml` в `json` без использования сторонних библиотек и регулярных выражений.

`convertors/yaml_json_lib` – модуль для конвертации файла `yaml` в `json` с использованием сторонней библиотеки.

`convertors/yaml_json_reg` – модуль для конвертации файла без использования сторонних библиотек, но с использованием регулярных выражений.

`convertors/yaml_csv` – модуль для конвертации файла `yaml` в `csv`.

В каждом модуле каждая функция задокументирована, чтобы было проще понять ее смысл, структуру, аргументы и возвращаемое значение.

Файлы для конвертации можно найти в корне каталога `program` с именем `<формат>_data`

Вывод

Во время выполнения лабораторной работы я изучил типы хранения данных в разных форматах (yaml, json, xml, html и др.), а также научился реализовывать парсер из yaml в json тремя способами и тестировать их. Полученные знания пригодятся мне в будущем и в процессе дальнейшего обучения.

Список использованных источников

1. <https://www.json2yaml.com> – онлайн парсер yaml<->json.
2. <https://pandas.pydata.org> – библиотека конвертации json в csv.
3. <https://pyyaml.org/wiki/PyYAMLDocumentation> - библиотека для работы с yaml.
4. <https://docs.python.org/3/library/json.html> - библиотека для работы с json.
5. Лекция номер 4 Павла Валерьевича Балакшина.