

**Федеральное государственное автономное образовательное учреждение**

**высшего образования**

**«Национальный исследовательский университет ИТМО»**

Факультет программной инженерии и компьютерной техники

Лабораторная работа №4

Вариант 19

По дисциплине Информатика

(наименование учебной дисциплины согласно учебному плану)

Тема работы: Исследование протоколов, форматов обмена информацией и языков разметки документов. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Выполнил: студент гр. Р3110  Голиков Д.И.

(шифр группы) (подпись) (Ф.И.О.)

Проверил: доцент \_\_\_\_\_\_\_\_\_\_\_\_ Балакшин П.В.

(должность) (подпись) (Ф.И.О.)

Санкт-Петербург

2022

Оглавление

[Задание 2](#_Toc106753518)

[Основные этапы выполнения 2](#_Toc106753519)

[Дополнительное задание №1 3](#_Toc106753520)

[Дополнительное задание №2 3](#_Toc106753521)

[Дополнительное задание №3 4](#_Toc106753522)

[Вывод 4](#_Toc106753523)

[Список литературы: 5](#_Toc106753524)

# Задание

Исходя из структуры расписания конкретного дня, сформировать файл с расписанием в формате, указанном в задании в качестве исходного. Написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый.

**Вариант**

|  |  |  |  |
| --- | --- | --- | --- |
| № варианта | Исходный формат | Результирующий формат | День недели |
| 19 | JSON | XML | Четверг |

# Основные этапы выполнения

Основное обязательное задание

Изучив разницу между форматами JSON и XML, основной задачей написания программы является банальная замена синтаксиса. Например, начало строки в формате JSON обозначается как « “Tag”: » , а конец этой же строки простой запятой. В то время как в XML начало – «<Tag>» и конец «</tag>». Для этого в лабораторной работе используется метод find(), который в дальнейшим записывает всё в массив.

В данном задании практически всю основную работу делает оператор ветвления if-elif-else, он помогает изменять синтаксисы разных форматов, для достижения рабочего парсера.

# Дополнительное задание №1

a) Найти готовые библиотеки, осуществляющие аналогичный парсинг и конвертацию файлов.

b) Переписать исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.

c) Сравнить полученные результаты и объяснить их сходство/различие.

Для выполнения данного задания хорошо подойдет библиотека json2xml, так как она переводит JSON в XML без написания десятков строк кода. Все что остается сделать с этой библиотекой, так это правильно ее применить и записать полученный результат.

Отличий моего результата, от результата полученным с использованием библиотеки мало. Json2xml добавляет в тип тэга в него. Так, например, тэг title получил тип str. Удобно, когда еще в самом начале строчки приходит понимание, что это всего лишь строка, а не начало какого-нибудь списка или числа.

# Дополнительное задание №2

a) Переписать исходный код, добавив в него использование регулярных выражений.

b) Сравнить полученные результаты и объяснить их сходство/различие.

Импорт модуля re помогает для работы с регулярными выражениями своими функциями. Так, теперь можно отказаться от метода find() и использовать re.search(pattern,string), который позволяет найти в строке string первую строчку, подходящую под шаблон pattern.

Однако, использование регулярных выражений дело привычки. Пока мне даётся немного тяжеловато и приходится постоянно заглядывать в сторонние источники, чтобы вспомнить какой шаблон или квантификатор нужен в данном случае.

Результаты основного задания и дополнительного задания№2 одинаковые, что не может не радовать, потому что лично для меня использование регулярных выражений пока что ассоциируется с путаницей, где можно с лёгкостью пропустить или перепутать некоторые символы.

# Дополнительное задание №3

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.

Импортировав все предыдущие программы, сравниваем их время выполнения парсинга и конвертацию в цикле.

Основное задание = 0,00770

Дополнительное №1 = 0,10767

Дополнительное №2 = 0,00580

Я считаю, что такая большая разница между использованием библиотеки и регулярными выражениями из-за того, что библиотека содержит в себе ООП, в котором содержится большое количество объектов.

Большой разницей между использованием регулярных выражений и без них нет.

# Вывод

В итоге данной лабораторной работы, я изучил форму Бэкуса-Наура, особенности протоколов и форматов обмена информацией между системами: JSON, XML. Понял устройство страницы расписанием для своей группы и попрактиковался в программировании различными методами.

# Список литературы:

1. Лямин А.В., Череповская Е.Н. Объектно-ориентированное программирование. Компьютерный практикум. – СПб: Университет ИТМО, 2017. – 143 с. – Режим доступа: <https://books.ifmo.ru/file/pdf/2256.pdf>.
2. Википедия, Форма Бэкуса-Наура. Режим доступа: https://ru.wikipedia.org/wiki/Форма\_Бэкуса\_-\_Наура.
3. Статья “Пишем изящный парсер на Питоне” режим доступа: https://habr.com/ru/post/309242.