Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

*Факультет Программной инженерии и компьютерной техники*

Дисциплина: Информатика

**Лабораторная работа №4**

Исследование протоколов, форматов обмена информацией и языков разметки документов

Вариант №11

Группа: P3132

Выполнил: Васильев Александр Дмитриевич

Проверил: Белозубов Александр Владимирович

г. Санкт-Петербург

2022 г.

Оглавление

[Задание 3](#_Toc124944988)

[Порядок выполнения работы 4](#_Toc124944989)

[Обязательное задание 6](#_Toc124944990)

[Исходный код 6](#_Toc124944991)

[Результат выполнения программы 6](#_Toc124944992)

[Заключение 8](#_Toc124944993)

[Литература 9](#_Toc124944994)

# **Задание**

|  |  |  |  |
| --- | --- | --- | --- |
| № варианта | Исходный формат | Результирующий формат | День недели |
| 11 | XML | YAML | Вторник |

# **Порядок выполнения работы**

1. Определить номер варианта как остаток деления на 36 порядкового номеров в списке группы в ISU. В случае, если в данный̆ 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 своей̆ 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 данный̆ 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 №1 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

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

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

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

1. Дополнительное задание №2 (позволяет набрать +10 процентов от максимального числа баллов БаРС за данную лабораторную).

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

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

10.Дополнительное задание No3 (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 стократное время выполнения парсинга + конвертации в цикле.

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

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

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

d) Проанализировать полученные результаты, объяснить особенности использования формата.

# **Обязательное задание**

Написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию XML файла в YAML

## **Расписание в формате XML**

<?xml version="1.0" encoding="UTF-8" ?>  
<root>  
 <day>Вторник</day>  
 <classes>  
 <firstclass>  
 <time>8:20-9:50</time>  
 <weekcount>5</weekcount>  
 <room>ZOOM</room>  
 <lesson>Информатика(Лек): ZOOM</lesson>  
 <teacher>Балакшин Павел Валерьевич</teacher>  
 <lesson-format>Дистанционный</lesson-format>  
 </firstclass>  
 <secondclass>  
 <time>8:20-9:50</time>  
 <weekcount>3, 7, 9, 11, 13, 15, 17</weekcount>  
 <room>ул.Ломоносова, д.9, лит. М</room>  
 <lesson>Информатика(Лек): Актовый зал</lesson>  
 <teacher>Балакшин Павел Валерьевич</teacher>  
 <lesson-format>Очно - дистанционный</lesson-format>  
 </secondclass>  
 <thirdclass>  
 <time>10:00-11:30</time>  
 <weekcount>3, 7, 9, 11, 13, 15, 17</weekcount>  
 <room>ул.Ломоносова, д.9, лит. М</room>  
 <lesson>Основы профессиональной деятельности(Лек): Актовый зал</lesson>  
 <teacher>Клименков Сергей Викторович </teacher>  
 <lesson-format>Очно - дистанционный</lesson-format>  
 </thirdclass>  
 <fourthclass>  
 <time>10:00-11:30</time>  
 <weekcount>5</weekcount>  
 <room>ZOOM</room>  
 <lesson>Основы профессиональной деятельности(Лек): ZOOM</lesson>  
 <teacher>Клименков Сергей Викторович </teacher>  
 <lesson-format>Дистанционны</lesson-format>  
 </fourthclass>  
 <fithclass>  
 <time>11:40-13:10</time>  
 <weekcount>5</weekcount>  
 <room>ZOOM</room>  
 <lesson>Программирование(Лек): ZOOM</lesson>  
 <teacher>Письмак Алексей Евгеньевич</teacher>  
 <lesson-format>Дистанционный</lesson-format>  
 </fifthclass>  
  
 <test>  
 <time>10:00-11:30</time>  
 <weekcount>5</weekcount>  
 <room>ZOOM</room>  
 <lesson>Основы профессиональной деятельности(Лек): ZOOM</lesson>  
 <teacher>Клименков Сергей Викторович </teacher>  
 <lesson-format>Дистанционны</lesson-format>  
 </test>  
 </classes>  
</root>

## **Исходный код**

inputFile = open("timetable.xml", "r", encoding = "utf-8")  
text = inputFile.read()  
inputFile.close()  
  
text = text.replace('<', '')  
text = text.replace('>', ': ')  
text = text.replace('/', ' /')  
text = text.replace('?xml version="1.0" encoding="UTF-8" ?: ', '')  
text = text.replace('root:', '')  
newtext = text.split(" ")  
final =(' '.join(filter(lambda word: not word.lower().startswith("/"), newtext)))  
test = final.split(" ")  
#print(test)  
for i in range (len(test)):  
 if test[i] == test[i-2]:  
 test[i] = '\n\t\t\t\t\t—'  
for i in range (len(test)):  
 if test[i] == test[i-4]:  
 test[i] = '\n\t\t\t\t\t—'  
  
final1 = (' '.join(filter(lambda word: not word.lower().startswith("/"),test )))  
print(final1)

## **Результат выполнения программы**

day: Вторник

classes:

firstclass:

time: 8:20-9:50

weekcount: 5

room: ZOOM

lesson: Информатика(Лек): ZOOM

teacher: Балакшин Павел Валерьевич

lesson-format: Дистанционный

secondclass:

time: 8:20-9:50

weekcount: 3, 7, 9, 11, 13, 15, 17

room: ул.Ломоносова, д.9, лит. М

lesson: Информатика(Лек): Актовый зал

teacher: Балакшин Павел Валерьевич

lesson-format: Очно - дистанционный

thirdclass:

time: 10:00-11:30

weekcount: 3, 7, 9, 11, 13, 15, 17

room: ул.Ломоносова, д.9, лит. М

lesson: Основы профессиональной деятельности(Лек): Актовый зал

teacher: Клименков Сергей Викторович

lesson-format: Очно - дистанционный

fourthclass:

time: 10:00-11:30

weekcount: 5

room: ZOOM

lesson: Основы профессиональной деятельности(Лек): ZOOM

teacher: Клименков Сергей Викторович

lesson-format: Дистанционны

fithclass:

time: 11:40-13:10

weekcount: 5

room: ZOOM

lesson: Программирование(Лек): ZOOM

teacher: Письмак Алексей Евгеньевич

lesson-format: Дистанционный

test:

time: 10:00-11:30

weekcount: 5

room: ZOOM

lesson: Основы профессиональной деятельности(Лек): ZOOM

teacher: Клименков Сергей Викторович

lesson-format: Дистанционны

# **Заключение**

В результате выполнения лабораторной работы я научился выполнять парсинг и конвертацию файла из XML формата в YAML и понял, что что можно как писать парсеры самостоятельно, так и с помощью библиотек и регулярных выражений, которые способны облегчить процесс написания кода

# **Литература**

1. Балакшин П.В., Соснин В. В., Машина Е. А. Информатика. – СПб: Университет ИТМО, 2020. – 122 с.
2. Балакшин П.В., Соснин В. В., Калинин И.В., Малышева Т.А., Раков С.В.,

Рущенко Н.Г., Дергачев А.М. Информатика: лабораторные работы и тесты.

СПб: Университет ИТМО, 2019. – 56 с.