Федеральное государственное автономное

образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

**Лабораторная работа № 4**

**«Исследование протоколов, форматов обмена информацией и языков разметки документов»**

По дисциплине «Информатика»

Вариант 4

Выполнила:

Студентка группы P3117

Русакова Е.Д.

Преподаватель:

Машина Е.А.

Санкт-Петербург

2022

Оглавление

[Задание 2](#_Toc123075446)

[Ход работы: 3](#_Toc123075447)

[Исходный файл timetable.yml: 3](#_Toc123075448)

[Обязательное задание: 3](#_Toc123075449)

[Код программы main.py: 3](#_Toc123075450)

[Результирующий файл timetable.xml: 4](#_Toc123075451)

[Дополнительное задание №1: 5](#_Toc123075452)

[Дополнительное задание №2: 5](#_Toc123075453)

[Дополнительное задание №3: 5](#_Toc123075454)

[Дополнительное задание №4: 5](#_Toc123075455)

# Задание

* 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. Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов.
* 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 исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.

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

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

a) Переписать исходный код, добавив в него использование регулярных выражений. 2

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

* Дополнительное 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 стократное время выполнения парсинга + конвертации в цикле.

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

* Дополнительное задание №4 (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 любой другой формат (кроме JSON, YAML, XML, HTML): PROTOBUF, TSV, CSV, WML и т.п.

d) Проанализировать полученные результаты, объяснить особенности использования формата.

Вариант 4:

День недели – Понедельник

Исходный язык разметки – YAML

Результирующий язык разметки – XML

# Ход работы:

## Исходный файл timetable.yml:

timetable:

-group: P3117

-day: понедельник

lesson1:

-time: 15:20-16:50

-room: 3210 ауд.

-place: ул. Ломоносова, д.9, лит. Е

-lesson: Иностранный язык

-teacher: Аганина Екатерина Александровна

lesson2:

-time: 18:40-20:00

-room: 1419 ауд.

-place: Кронверкский пр., д. 49, лит. А

-lesson: Дискретная математика (продвинутый уровень)

-teacher: Карпов Дмитрий Валерьевич

lesson3:

-time: 20:20-21:50

-room: 2337 ауд.

-place: Кронверкский пр., д. 49, лит. А

-lesson: Дискретная математика (продвинутый уровень)

-teacher: Карпов Дмитрий Валерьевич

# Обязательное задание:

## Код программы main.py:

**import** **time**

start = time.time()

f = open('timetable.yml', 'r', encoding="utf-8")

list = f.readlines()

f.close()

f1 = open('timetable.xml', 'w', encoding="utf-8")

**for** line **in** list:

line = line.replace('-','',**1**)

**for** i **in** range(len(line)):

**if** line[i] != ' ':

s = i

**break**

**if** line.find(":"):

line = line[**0**:s] + '<' + line[s:len(line)]

line = line.replace(' ','**\t**',s)

**if** line.find(":"):

line = line.replace(': ',':',**1**)

line = line.replace(':','>',**1**)

**if** line.find('<') != -**1**:

s1 = line.find('<')

s2 = line.find('>')

tag = line[s1+**1**:s2]

**if** line[len(line)-**2**] != '>':

line = line[**0**:len(line)-**1**] + '</' + tag + '>' + '**\n**'

**if** line.find("lesson2") != -**1**:

f1.write('**\t**</lesson1>**\n**')

**if** line.find("lesson3") != -**1**:

f1.write('**\t**</lesson2>**\n**')

f1.write(line)

f1.write('**\t**</lesson3>**\n**')

f1.write('</timetable>**\n**')

f1.close()

end = time.time()

**print**((end - start)\***100**)

## Результирующий файл timetable.xml:

<timetable>

<group>P3117</group>

<day>понедельник</day>

<lesson1>

<time>15:20-16:50</time>

<room>3210 ауд.</room>

<place>ул. Ломоносова, д.9, лит. Е</place>

<lesson>Иностранный язык</lesson>

<teacher>Аганина Екатерина Александровна</teacher>

</lesson1>

<lesson2>

<time>18:40-20:00</time>

<room>1419 ауд.</room>

<place>Кронверкский пр., д. 49, лит. А</place>

<lesson>Дискретная математика (продвинутый уровень)</lesson>

<teacher>Карпов Дмитрий Валерьевич</teacher>

</lesson2>

<lesson3>

<time>20:20-21:50</time>

<room>2337 ауд.</room>

<place>Кронверкский пр., д. 49, лит. А</place>

<lesson>Дискретная математика (продвинутый уровень)</lesson>

<teacher>Карпов Дмитрий Валерьевич</teacher>

</lesson3>

</timetable>

# Дополнительное задание №1:

# Дополнительное задание №2:

## Код программы dop2.py:

**import** **re**

**import** **time**

start = time.time()

f = open('timetable.yml', 'r', encoding="utf-8")

list = f.readlines()

f.close()

f1 = open('timetable2.xml', 'w', encoding="utf-8")

**for** line **in** list:

line = re.sub('-','',line,**1**)

**for** i **in** range(len(line)):

**if** line[i] != ' ':

s = i

**break**

**if** re.search(':',line):

line = line[:s] + '<' + line[s:]

line = re.sub(' ','**\t**',line,s)

**if** re.search(':',line):

line = re.sub(':\s\*','>',line,**1**)

tag1 = re.search('<\w+>',line)

tag = tag1[**0**]

**if** line[len(line)-**1**] != '>':

line = line[**0**:len(line)-**1**] + '</' + tag[**1**:] + '**\n**'

**else**:

line = line + '**\n**'

**if** line.find("lesson2") != -**1**:

f1.write('**\t**</lesson1>**\n**')

**if** line.find("lesson3") != -**1**:

f1.write('**\t**</lesson2>**\n**')

f1.write(line)

f1.write('**\t**</lesson3>**\n**')

f1.write('</timetable>**\n**')

f1.close()

end = time.time()

**print**((end - start)\***100**)

## Результирующий файл timetable2.xml:

<timetable>

<group>P3117</group>

<day>понедельник</day>

<lesson1>

<time>15:20-16:50</time>

<room>3210 ауд.</room>

<place>ул. Ломоносова, д.9, лит. Е</place>

<lesson>Иностранный язык</lesson>

<teacher>Аганина Екатерина Александровна</teacher>

</lesson1>

<lesson2>

<time>18:40-20:00</time>

<room>1419 ауд.</room>

<place>Кронверкский пр., д. 49, лит. А</place>

<lesson>Дискретная математика (продвинутый уровень)</lesson>

<teacher>Карпов Дмитрий Валерьевич</teacher>

</lesson2>

<lesson3>

<time>20:20-21:50</time>

<room>2337 ауд.</room>

<place>Кронверкский пр., д. 49, лит. А</place>

<lesson>Дискретная математика (продвинутый уровень)</lesson>

<teacher>Карпов Дмитрий Валерьевич</teacher>

</lesson3>

</timetable>

# Дополнительное задание №3:

Основное задание:

0.06954669952392578

Дополнительное №1:

Дополнительное №2:

0.33354759216308594