МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ

Федеральное государственное автономное

образовательное учреждение высшего образования

«Санкт-Петербургский национальный исследовательский университет

информационных технологий, механики и оптики»

**ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНОЙ ТЕХНИКИ**

**ЛАБОРАТОРНАЯ РАБОТА № 4**

**“** Исследование протоколов, форматов обмена информацией и языков разметки документов**”**

по дисциплине

‘ИНФОРМАТИКА’

Вариант № 9

Выполнил:

Студент группы P3118  
Кравец Роман Денисович

Преподаватель:

Малышева Татьяна Алексеевна

Санкт-Петербург, 2021

Оглавление

[Задание: 3](#_Toc87490009)

[Обязательное задание (+65 баллов): 3](#_Toc87490010)

[Дополнительное задание задание №1 (+10 баллов): 3](#_Toc87490011)

[Дополнительное задание задание №2(+10 баллов): 3](#_Toc87490012)

[Дополнительное задание задание №3(+10 баллов): 3](#_Toc87490013)

[Дополнительное задание задание №4 (+5 баллов): 3](#_Toc87490014)

[Расписание: 4](#_Toc87490015)

[Основные этапы выполнения: 5](#_Toc87490016)

[Исходный файл: 5](#_Toc87490017)

[Основное задание: 5](#_Toc87490018)

[Дополнительное задание 1: 6](#_Toc87490019)

[Дополнительное задание 2: 7](#_Toc87490020)

[Дополнительное задание 4: 8](#_Toc87490021)

[Вывод 9](#_Toc87490022)

[Список литературы 9](#_Toc87490023)

# 

# Задание:

## **Обязательное задание (+65 баллов):**

1) написать программу на языке Python 3.x, которая бы осуществляла парсинг и конвертацию исходного файла в новый. 7.

2) Нельзя использовать готовые библиотеки, в том числе регулярные выражения в Python и библиотеки для загрузки XML-файлов

## **Дополнительное задание задание №1 (+10 баллов):**

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 задание №2(+10 баллов):**

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

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

## **Дополнительное задание задание №3(+10 баллов):**

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.

## **Дополнительное задание задание №4 (+5 баллов):**

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.



Figure 1. Вариант.

# Расписание:

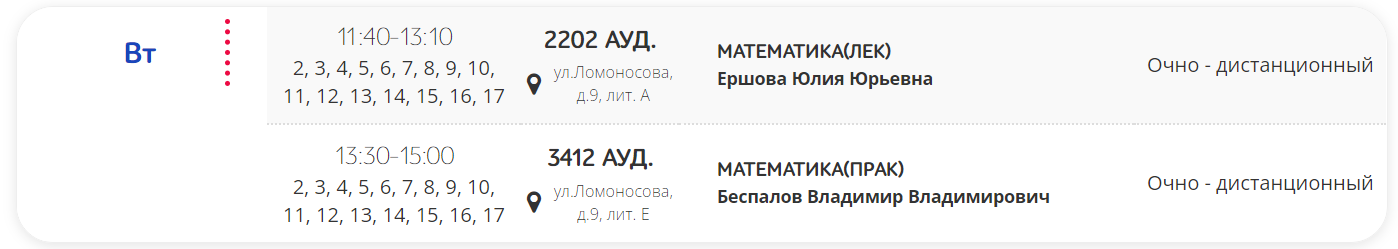


Figure 2. Расписание

# Основные этапы выполнения:

## **Исходный файл:**

{

"CLASS\_P3118": {

"date": "October 26, 2021",

"day": "TUESDAY",

"lessons": [

{

"first": {

"subject": "МАТЕМАТИКА",

"teacher": "Ершова Юлия Юрьевна",

"lesson-format": "Очно - дистанционный",

"room": "ул.Ломоносова, д.9, лит.А",

"time": "11:40-13:10",

"weeks": "2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17",

"aud": "2202"

}

},

{

"second": {

"subject": "МАТЕМАТИКА",

"teacher": "Беспалов Владимир Владимирович",

"lesson-format": "Очно - дистанционный",

"room": "ул.Ломоносова, д.9, лит.Е",

"time": "13:30-15:00",

"weeks": "2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17",

"aud": "3412"

}

}

]

}

}

## **Основное задание:**

import time

start\_time = time.perf\_counter()

print("Вариант " + str(9 %36))

#вар 9, исходный формат = JSON, результирующий формат = YAML, день недели = Вторник

arr =[]

file1 = open("json\_in.json","r",encoding="UTF-8")

while True:

line = file1.readline()

if not line:

break

stroka = str(line.strip()) #будут устранены пробельные символы

if '"' in stroka:

stroka = stroka.replace('",', '"') #УБИРАЕМ , В КОНЦЕ

stroka = stroka.replace('"',"",) #УБИРАЕМ "

n = 0

if stroka[n-1] == '{':

stroka = stroka.split()

count = len(stroka)

for x in range(count):

arr.append(stroka[x])

elif stroka[n-1] == "[":

stroka = stroka.split()

count = len(stroka)

for x in range(count):

arr.append(stroka[x])

else:

arr.append(stroka)

count = 0

for i in range(len(arr)):

if (arr[i] == '{') and (arr[i-1] != '['):

count +=1

if (count == 1) and (arr[i]!='{'):

print(arr[i])

if (count == 2) and (arr[i]!='{') and (arr[i]!='[') and (arr[i-2] != '['):

print(' ', arr[i])

if (count == 2) and (arr[i-1] =='{') and (arr[i-2] =='['):

print(' -',arr[i])

if (count == 3) and (arr[i] != '{') and (arr[i] != '}') and (arr[i] != '},'):

print(' ',arr[i])

if (count == 4) and (arr[i-1] == '{') and (arr[i-2] == '},'):

print(' -',arr[i])

if (count == 5) and (arr[i]!='{') and (arr[i]!='}') and (arr[i]!=']'):

print(' ',arr[i])

print('Прошло',10 \* (time.perf\_counter() - start\_time), 'сек.')

## **Дополнительное задание 1:**

import json

import time

from ruamel.yaml import YAML

start\_time = time.perf\_counter()

in\_file = 'json\_in.json'

out\_file = 'yaml\_out.yaml'

yaml = YAML(typ='safe')

yaml.default\_flow\_style = False

with open(in\_file, 'r', encoding='utf-8') as i:

data = json.load(i)

with open(out\_file,'w', encoding='utf-8') as o:

yaml.dump(data, o)

print('Прошло', 10\*(time.perf\_counter() - start\_time),'сек.')

## **Дополнительное задание 2:**

import re

import time

start\_time = time.perf\_counter()

print("Вариант " + str(9 %36))

#вар 9, исходный формат = JSON, результирующий формат = YAML, день недели = Вторник

arr =[]

file1 = open("json\_in.json","r",encoding="UTF-8")

while True:

line = file1.readline()

if not line:

break

stroka = str(line.strip()) #будут устранены пробельные символы

if '"' in stroka:

stroka = re.sub(r'",','"',stroka) #замена replace

stroka = re.sub(r'"','',stroka)

n = 0

if re.findall('\{', stroka):

stroka = stroka.split()

count = len(stroka)

for x in range(count):

arr.append(stroka[x])

elif re.findall('\[', stroka):

stroka = stroka.split()

count = len(stroka)

for x in range(count):

arr.append(stroka[x])

else:

arr.append(stroka)

count = 0

for i in range(len(arr)):

if (arr[i] == '{') and (arr[i-1] != '['):

count +=1

if (count == 1) and (re.findall('\w+',arr[i])):

print(arr[i])

if (count == 2) and (re.findall('^[^f]\w+',arr[i])):

print(' ', arr[i])

if (count == 2) and (re.findall('^f',arr[i])):

print(' -',arr[i])

if (count == 3) and (re.findall('\w+',arr[i])):

print(' ',arr[i])

if (count == 4) and (re.findall('^s',arr[i])):

print(' -',arr[i])

if (count == 5) and (re.findall('\w+',arr[i])):

print(' ',arr[i])

print('Прошло',10 \*(time.perf\_counter() - start\_time),'сек.')

## **Дополнительное задание 4:**

#json -> csv формат

import re

all =[]

spisok =[]

file1 = open("input.json","r",encoding="UTF-8")

for x in range(31):

f = file1.readline()

if 6 <= x <= 25:

if x == 6:

kl1 =re.findall('\w+',f)

if 7 <=x <=13:

if x == 12 or x == 10:

key = re.search('\w+',f)

all.append(key.group())

val1 = re.search(': "[\w,."]+',f)

val1 = val1.group().replace('",', '"')

s = val1.replace(": ", '')

spisok.append(s)

else:

key = re.search('\w+',f)

#print(key.group())

#print(kl.group() +'/' + key.group() +',')

all.append(key.group())

val1 = re.search(': "[\w\s,-.:]+',f)

s = val1.group().replace(': "', '')

spisok.append(s)

if x== 17:

kl2 = re.findall('\w+',f)

if 18<=x<=24:

if x == 23 or x == 21:

val2 = re.search(': "[\w,."]+',f)

val2 = val2.group().replace('",', '"')

s = val2.replace(": ", '')

spisok.append(s)

else:

val2 = re.search(': "[\w\s,-.:]+',f)

s = val2.group().replace(': "', '')

spisok.append(s)

for x in range(7):

x = kl1[0] + '/' + all[x] + ','

print(x, end = '')

for x in range(7):

y = kl2[0]+ '/' + all[x] + ','

if x==6:

print(kl2[0]+ '/' + all[x])

else:

print(y, end = '')

for x in range(14):

if 0<=x<=5:

print(spisok[x]+ ',',end ='')

if x==6:

print(spisok[x]+ ','+','\*6)

if x== 7:

print(','\*7, end='')

if 7<=x<=12:

print(spisok[x]+ ',',end ='')

if x ==13:

print(spisok[x])

# 

# Вывод

В ходе выполнения лабораторной работы я изучил особенности протоколов и форматов обмена информацией между системами: JSON и YAML. Также познакомился с готовой библиотекой, осуществляющей аналогичный парсинг и конвертацию.

**Список литературы**

Балакшин П.В. Соснин В.В. Информатика. Методическое пособие.

<https://docplayer.com/136642876-Balakshin-p-v-sosnin-v-v-informatika-metodicheskoe-posobie-sankt-peterburg-2015-g.html>

Алексеев Е.Г., Богатырев С.Д. Информатика. Мультимедийный электронный учебник.

<https://qo.do.am/index/multimedijnyj_ehlektronnyj_uchebnik_alekseev_e_g_bogatyrev_s_d/0-33>