Санкт-Петербургский Национальный Исследовательский Университет ИТМО

Факультет Программной Инженерии и Компьютерной Техники

**Лабораторная работа №3**

по дисциплине «**Информатика**»

Выполнил студент: Фам Данг Чунг Нгиа

Группа: P3121

Преподаватель: Болдырева Елена Александровна

Санкт- Петербург

2023

**Оглавление**

[**Задания** 3](#_Toc148347970)

[**Задание 1** 3](#_Toc148347971)

[**Задание 2** 3](#_Toc148347972)

[**Задание 3** 3](#_Toc148347973)

[**Ответы** 5](#_Toc148347974)

[**Задание 1** 5](#_Toc148347975)

[**Задание 2** 5](#_Toc148347976)

[**Задание 3** 6](#_Toc148347977)

[**Дополнительное задание 1** 8](#_Toc148347978)

[**Дополнительное задание 2** 9](#_Toc148347979)

[**Вывод** 10](#_Toc148347980)

[**Список литературы** 10](#_Toc148347981)

**Номер ИСУ: 374806**

# **Задания**

## **Задание 1**

1) Реализуйте программный продукт на языке Python, используя регулярные выражения по варианту, представленному в таблице.

2) Для своей программы придумайте минимум 5 тестов. Каждый тест является отдельной сущностью, передаваемой регулярному выражению для обработки. Для каждого теста необходимо самостоятельно (без использования регулярных выражений) найти правильный ответ. После чего сравнить ответ, выданный программой, и полученный самостоятельно.

3) Программа должна считать количество смайликов определённого вида (вид смайлика описан в таблице вариантов) в предложенном тексте. Все смайлики имеют такую структуру: [глаза][нос][рот]. Вариантом является различные наборы глаз, носов и ртов.

Вариант: 125

|  |  |  |
| --- | --- | --- |
| Глаза | Нос | Рот |
| ; | -{ | / |

## **Задание 2**

1) Реализуйте программный продукт на языке Python, используя регулярные выражения по варианту, представленному в таблице.

2) Для своей программы придумайте минимум 5 тестов.

3) Протестируйте свою программу на этих тестах.

Вариант: 4

Дан текст. Требуется найти в тексте все фамилии, отсортировав их по алфавиту.

Фамилией для простоты будем считать слово с заглавной буквой, после которого идут инициалы.

Пример:

|  |  |
| --- | --- |
| Ввод | Вывод |
| Студент Вася вспомнил, что на своей лекции Болдырева Е.А. упоминала про старшекурсников, которые будут ей помогать: Анищенко А.А. и Машина Е.А. | Анищенко  Болдырева  Машина |

## **Задание 3**

1. Определить номер варианта как остаток деления номера в ИСУ на 36. В случае, если в данный день недели нет занятий, то увеличить номер варианта на восемь.

|  |  |  |  |
| --- | --- | --- | --- |
| № варианта | Исходный формат | Результирующий формат | День недели |
| 374 806 % 36 = **10** | YAML | JSON | Вторник |

2. Изучить форму Бэкуса-Наура.

3. Изучить особенности протоколов и форматов обмена информацией между системами: JSON, YAML, XML.

4. Понять устройство страницы с расписанием для своей группы: https://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.

6. Обязательное задание: 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 исходный код, применив найденные библиотеки. Регулярные выражения также нельзя использовать.

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

9. **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.

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

# **Ответы**

## **Задание 1**

import re

input=open('file.txt','r')

pattern=';-{/'

for x in input:

if(pattern in x):

print(x, end='')

print(x.count(pattern))

|  |  |
| --- | --- |
| Тест в “File.txt” | Вывод |
| ;-{/{ | 1 |
| abcd;-{/ | 1 |
| ;-{/;-{/ddee; | 2 |
| ;-{/23{4ab/c;-{/;-{/ | 3 |
| a;-{/b/;-{/c;-{/d/;-{/; | 4 |

## **Задание 2**

import re  
with open('test1\_ex2.txt', 'r') as input\_file:   
 text=input\_file.read()  
result=re.findall('[A-Z][a-z]+\s[A-Z]\.[A-Z]\.', text) result=sorted(result)  
for x in result:  
 y=re.findall('[A-Z][a-z]+', x)  
 print(y)

|  |  |
| --- | --- |
| Тест | Вывод |
| Student Vasya remembered that at his lecture Boldyreva E.A. mentioned senior students who would help her: Anishchenko A.A. and Machine E.A. | ['Anishchenko']  ['Boldyreva']  ['Machine'] |
| Almakael A.A., Boychenko V.M., and Alisheva K.R. are outstanding individuals who have made significant contributions to their respective fields. Almakael A.A. has tirelessly strived to research and contribute to the field of computer science. Boychenko V.M. is a technical expert known for creating innovative solutions. Alisheva K.R. is renowned for her in-depth research in the field of medicine and healthcare. We take great pride in their contributions to society and science. | ['Alisheva']  ['Alisheva']  ['Almakael']  ['Almakael']  ['Boychenko']  ['Boychenko'] |
| Buzin D.V., Dvoeglazova N.N., Egorova V.A., and Zehir D.A. have all been standout individuals in the academic realm, each leaving an indelible mark on the educational landscape. Buzin D.V. has been a pioneer in the field of technology and has shown a relentless commitment to innovation. Dvoeglazova N.N. is celebrated for her exceptional contributions to the world of scientific research, always pushing the boundaries of knowledge. Egorova V.A. is a distinguished figure in the world of academia, known for her astute leadership and vision. Zehir D.A. has been a luminary in the realm of education, particularly in the field of medicine, where their dedication to research and patient care has been truly remarkable. Together, they have shaped the educational sphere and continue to inspire future generations. | ['Buzin']  ['Buzin']  ['Dvoeglazova']  ['Dvoeglazova']  ['Egorova']  ['Egorova']  ['Zehir']  ['Zehir'] |
| Yildiz S.S. and Karov A.D. are best friends | ['Karov']  ['Yildiz'] |
| Тест | Вывод |
| Ha H.H. and Vy T.T. are Vietnameses and they are partners | ['Ha']  ['Vy'] |

## **Задание 3**

---  
Время : Вторник/8  
Занятие\_1 :  
 - Время : 08:20~09:50  
 - Ауд./Корпус : aуд.1229/Кронверский  
 - Предмет/Преподаватель : Математический\_Анализ/Правдин.К.В  
 - Формат\_занятий : Лекция  
Занятие\_2 :  
 - Время : 10:00~11:30  
 - Ауд./Корпус : ауд.2416/Кронверский  
 - Предмет/Преподаватель : Математический\_Анализ/Правдин.К.В  
 - Формат\_занятий : Практика  
...

Текст YAML

Исходный код

yaml\_data =""  
with open("schedule.yaml", "r", encoding = "utf-8") as f:  
 yaml\_data=f.read()  
  
def yaml\_to\_json(yaml\_data):  
 lines = yaml\_data.split('\n') #chia văn bản thành bộ các dòng  
 json\_data="" #tạo giá trị trả về của hàm  
 i=0  
 for line in lines: #xét lần lượt từng dòng  
 words=line.split(' ') #chia dòng thành bộ các từ  
 new\_words=[] #tạo 1 danh sách các từ mới  
 for word in words: #xét lần lượt từng từ (trong dòng)  
 if(word=="---" or word=="..."):  
 break  
 if(word=='-' or word==''):  
 continue  
 if(word==':'):  
 new\_words.append(word)  
 continue  
 else: new\_words.append("\""+word+"\"")  
 line=''.join(new\_words)  
  
 if(i==0): json\_data="{"+json\_data+"\n"  
 if(i==1): json\_data=json\_data+line+",\n"  
 if(i==2): json\_data=json\_data+line+"[\n"  
 if(i==3 or i==4 or i==5):  
 json\_data= json\_data+"{"+ line+"},\n"  
 if(i==6): json\_data= json\_data+"{"+ line+"}],\n"  
 if(i==7): json\_data=json\_data+line+"[\n"  
 if(i==8 or i==9 or i==10):  
 json\_data= json\_data+"{"+ line+"},\n"  
 if(i==11): json\_data= json\_data+"{"+ line+"}]}\n"  
 i=i+1  
 return json\_data  
  
# Chuyển đổi YAML thành JSON  
json\_data = yaml\_to\_json(yaml\_data)  
  
# In JSON kết quả  
with open("schedule.json", "w", encoding = "utf-8") as f:  
 f.write(json\_data)

## 

## **Дополнительное задание 1**

import json  
  
from ruamel.yaml import YAML  
  
in\_file = "schedule.yaml"  
out\_file="schedule.json"  
  
yaml=YAML(typ="safe") #чтобы избежать потенциальных угроз безопасности  
  
with open(in\_file, "r", encoding="utf-8") as i:  
 data = yaml.load(i) #Используется для анализа данных YAML из другого источника данных и преобразования их в структуру данных Python.  
  
with open(out\_file, "w", encoding="utf-8") as o:  
 json.dump(data, o, indent=4, ensure\_ascii=False)

## **Дополнительное задание 2**

import re  
yaml\_data=""  
with open ("schedule.yaml", "r", encoding="utf-8") as file:  
 yaml\_data=file.read()  
  
def yaml\_to\_json(yaml\_data):  
 json\_data=""  
 lines=yaml\_data.split('\n')  
 i=0  
 for line in lines:  
 words=line.split(' ')  
 newWords=[]  
 for word in words:  
 if(word=='-'or word=="---"or word=="..."): continue  
 if (re.match('[0-9]+:[0-9]+(~)?', word)):  
 newWords.append('\"'+word+'\"')  
 continue  
 if (re.match('[\w./]+', word)):  
 newWords.append('"'+word+'"')  
 else: newWords.append(word)  
 line=''.join(newWords)  
  
 if(i==0): json\_data="{"+json\_data+"\n"  
 if(i==1): json\_data=json\_data+line+",\n"  
 if(i==2): json\_data=json\_data+line+"[\n"  
 if(i==3 or i==4 or i==5):  
 json\_data= json\_data+"{"+ line+"},\n"  
 if(i==6): json\_data= json\_data+"{"+ line+"}],\n"  
 if(i==7): json\_data=json\_data+line+"[\n"  
 if(i==8 or i==9 or i==10):  
 json\_data= json\_data+"{"+ line+"},\n"  
 if(i==11): json\_data= json\_data+"{"+ line+"}]}\n"  
 i=i+1  
  
 return json\_data  
  
json\_data=yaml\_to\_json(yaml\_data)  
  
with open ("schedule.json", 'w', encoding="utf-8") as file:  
 file.write(json\_data)

# 

# Текст JSON

{  
"Время":"Вторник/8",  
"Занятие\_1":[  
{"Время":"08:20~09:50"},  
{"Ауд./Корпус":"aуд.1229/Кронверский"},  
{"Предмет/Преподаватель":"Математический\_Анализ/Правдин.К.В"},  
{"Формат\_занятий":"Лекция"}],  
"Занятие\_2":[  
{"Время":"10:00~11:30"},  
{"Ауд./Корпус":"ауд.2416/Кронверский"},  
{"Предмет/Преподаватель":"Математический\_Анализ/Правдин.К.В"},  
{"Формат\_занятий":"Практика"}]}

# **Вывод**

У меня есть базовое понимание XML, JSON, YAML и способов преобразования между ними.

# **Список литературы**

1. <https://apps.openedu.ru/learning/course/course-v1:ITMOUniversity+COMTEC+self_2023/block-v1:ITMOUniversity+COMTEC+self_2023+type@sequential+block@f19f9de48dec4952af3e0c14dced3bc1/block-v1:ITMOUniversity+COMTEC+self_2023+type@vertical+block@4679a611039d48e8b7dce2f245e7ad86>