

Санкт-Петербургский университет Информационных Технологий Механики и Оптики (ИТМО)

Факультет программной инженерии и компьютерной техники (ФПИиКТ)

Кафедра вычислительной техники (ВТ)

Лабораторная работа №4

Исследование языков разметки для документов

Вариант №29

**Преподаватель:**

Балакшин Павел Валерьевич

**Выполнил:**

Цыпандин Николай Петрович

Студент, группы Р3110

Санкт-Петербург

2020

**Задание №1**

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

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.

Вариант №4

import re  
s = open('D:\other\ITMO\Informatics\Labs\lab4\input11.txt', 'r', encoding = 'utf8')  
text = s.read().split('\n')  
pattern = r'([01]\d:([0-5]\d:?){1,2})|(2[0-3]:([0-5]\d:?){1,2})'  
for i in range(len(text)):  
 print(re.sub(pattern,r'(TBD)',text[i]))  
s.close()

|  |  |
| --- | --- |
| Ввод | Вывод |
| Я обычно выхожу из общаги в 07:30 при этом покупаю 40:50:60 бананов, огурцов и яблок соответственно | Я обычно выхожу из общаги в (TBD) при этом покупаю 40:50:60 бананов, огурцов и яблок соответственно |
| Я сделал опечатку и написал время встречи 24:24 вместо 00:24 | Я сделал опечатку и написал время встречи 24:24 вместо (TBD) |
| У меня первая пара по информатике начинается в 08:20 а вторая в 10:00:00 | У меня первая пара по информатика начинается в (TBD) а вторая в (TBD) |
| Петя в 8:20 времени сказал: « 9:3 = 3 это факт ». | Петя в 8:20 времени сказал: « 9:3 = 3 это факт ». |
| Уважаемые студенты! В эту субботу в 15:00 планируется доп. занятие на 2 часа. То есть в 17:00:01 оно уже точно кончится. | Уважаемые студенты! В эту субботу в (TBD) планируется доп. занятие на 2 часа. То есть в (TBD) оно уже точно кончится. |
| Планируется захватить здание в 08:37 по местному времени, далее уходим в 9:40 по часам на стенах. | Планируется захватить здание в (TBD) по местному времени, далее уходим в 9:40 по часам на стенах. |

**Задание №2**

Вариант №2

import re  
s = open('D:\other\ITMO\Informatics\Labs\lab4\input21.txt', 'r', encoding = 'utf8')  
text = s.read().split('\n')  
pattern = r'[A-Z][^?!.\n;]\*!'  
patternzap = r'\,'  
for i in range(len(text)):  
 if(re.search(pattern,text[i])):  
 dops=list()  
 dops=re.findall(pattern,text[i])  
 for j in range(len(dops)):  
 if(re.search(patternzap,dops[j])):  
 helps = re.search(pattern, dops[j])  
 splited = text[i].split(helps.group(), maxsplit=1)  
 text[i] = splited[0] + helps.group().upper() + splited[1]  
 print(text[i])  
s.close()

Вывод:

...

Cap. O heavens! O WIFE, LOOK HOW OUR DAUGHTER BLEEDS!

Wife. O me! this sight of death is as a bell

That warns my old age to a sepulchre.

Enter Montague [and others].

Prince. Come, Montague; for thou art early up

To see thy son and heir more early down.

Mon. ALAS, MY LIEGE, MY WIFE IS DEAD TO-NIGHT!

...

Проверял по своим тестам где рассмотрел множество разных вариантов расположеия и количества запятых, восклицательных знаков.

**Задание №3**

import re  
s = open('D:\other\ITMO\Informatics\Labs\lab4\input31.txt', 'r', encoding = 'utf8')  
text = s.read().split('\n')  
group='P000'  
pattern21=r'[А-Я][а-я]+\s[А-Я]\.[А-Я]\. '+group  
pattern22=r'[А-Я]\.'  
for i in range(len(text)):  
 if(re.search(pattern21,text[i])):  
 dops=list()  
 dops=re.findall(pattern22,text[i])  
 check=re.findall(pattern21,text[i])  
 if(dops[0][0]==dops[1][0]==check[0][0]):  
 continue  
 else:  
 print(text[i])  
   
 else:  
 print(text[i])  
s.close()

|  |  |
| --- | --- |
| Ввод | Вывод |
| Group=’P000’  Криштиану К.К. M3110  Месси М.М. M3110  Кержаков К.К. G3110  Кокорин К.Г. M3110  Насри Н.Н. M3110 | Кержаков К.К. G3110  Кокорин К.Г. M3110 |
| Group=’P3112’  Панчук П.П. P3112  Цыпандин Ц.Ц. P3112  Цивцин Ц.С. P3112  Клясюков М.О. P3141  Петров С.Ж. P0000 | Цивцин Ц.С. P3112  Клясюков М.О. P3141  Петров С.Ж. P0000 |
| Group=’P3110’  Путин П.П. P3110  Медведев М.М. P3110  Кличко К.К. P3110  Коворкинг К.К. P3111  Балакшин Б.Б. P3110 | Коворкинг К.К. P3111 |
| Group=’bad’  Степанов М.М. P312312  Степанчук С.С. bad  Степник С.С. bad  Стальной С.С. sad  Старый С.С. sak | Степанов М.М. P312312  Стальной С.С. sad  Старый С.С. sak |
| Group=’P123123’  Успешный У.У. P123123  Бессмертный Б.Б. P123123  Красивый К.К. P123321  Добрый Д.Д. P123321  Ленивый Я.Я. P123123 | Красивый К.К. P123321  Добрый Д.Д. P123321  Ленивый Я.Я. P123123 |

**Вывод:**

В данной лабораторной работе была освоен навык работы с регулярными выражениями на языке Python 3