**Федеральное государственное автономное**

**образовательное учреждение высшего образования**

**«Санкт-Петербургский**

**национальный исследовательский университет**

**информационных технологий, механики и оптики»**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

## 

**Лабораторная работа №4**

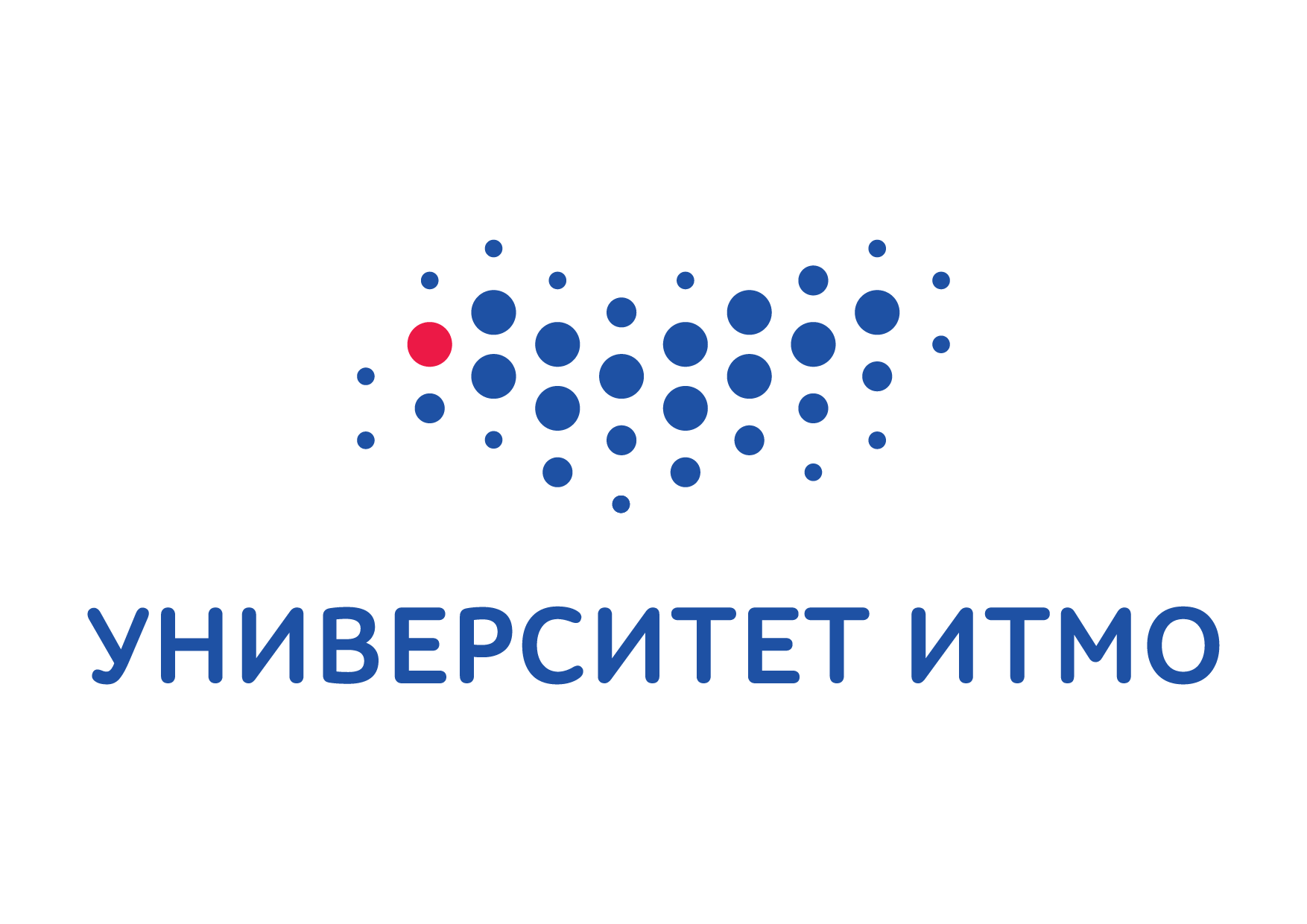
**по Информатике**

Вариант: 9

Группа: Р3110

Выполнил: Данилов Павел Юрьевич

Преподаватель: Балакшин Павел Валерьевич

****

**САНКТ-ПЕТЕРБУРГ**

**2020**

# Лабораторная работа №4 «Регулярные выражения»

Для определения варианта используйте свой порядковый номер по списку группы, которые можно найти в ИСУ. Задание на 70 баллов

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. Пример тестов приведён в таблице.

**Задание на 70 баллов**

Анатолий выложил пост с расписанием доп. занятий по информатике, но везде перепутал время. Поэтому нужно заменить все вхождения времени на строку (TBD). Время – это строка вида HH:MM:SS или HH:MM, в которой HH – число от 00 до 23, а MM и SS – число от 00 до 59.

**Пример**:

Ввод: Уважаемые студенты! В эту субботу в 15:00 планируется доп. занятие на 2 часа. То есть в 17:00:01 оно уже точно кончится.

Вывод: Уважаемые студенты! В эту субботу в (TBD) планируется доп. занятие на 2 часа. То есть в (TBD) оно уже точно

**Доп. задание 1 (+12 баллов)**

**Текст:**

<https://drive.google.com/file/d/1j2pEAwainLQ7dbOf2CFAxDTjQ9O7tNl/view>: Гамлет

1)Реализуйте программный продукт на языке Python, используя регулярные выражения, выделите все предложения из вашего текста, соответствующие критерию по вашему варианту.

2)Придумайте способ, как проверить корректность работы вашего регулярного выражения и проверьте его

**Регулярное выражение:**

Все восклицательные предложения, содержащие хотя бы одну запятую

**Доп. задание 2 (+18 баллов)**

С помощью регулярного выражения найти числа, в которых в средине числа стоит 0 и затем перенести его в конец числа.

**Пример:**

Ввод: 101 + 2430123 = 2430224

Вывод: 110 + 2431230 = 2432240

# Задание 1:

**Код**

import re  
  
ans = open(r"C:\Users\admin\PycharmProjects\Parser\out", "w", encoding="utf-8")  
file = open(r"C:\Users\admin\PycharmProjects\Parser\txt", "r", encoding="utf-8")  
string = file.read()  
string = re.sub("(^|\s)((?:[0**-**1][0**-**9]|2[0**-**3]):[0**-**5][0**-**9]:[0**-**5][0**-**9]|(?:[0**-**1][0**-**9]|2[0**-**3]):[0**-**5][0**-**9])([.?!;,\s]|$)", r"\1(TBD)\3", string)  
ans.write(string)

### Тесты:

1)

Ввод: 25:00:01

Вывод: 25:00:01

2)

Ввод: AAA 22:50:01, abcdef

Вывод: AAA (TBD), abcdef

3)

Ввод: not22:00:01

Вывод: not22:00:01

4)

Ввод: 01:01. not

Вывод: (TBD). not

5)

Ввод: 01:01not

Вывод: 01.01not

# Задание 2:

### Код

import re  
  
ans = open(r"C:\Users\admin\PycharmProjects\Parser\out", "w", encoding="utf-8")  
file = open(r"D:\Google Downloads\Hamlet.txt", "r", encoding="utf-8")  
string = file.read()  
strings = re.findall("[^!.?,]+,[^!.?]+[!]", string)  
for s in strings:  
 ans.write(s)  
 ans.write("\n")

**Вывод:**

Stand, ho!  
 Holla, Bernardo!  
 See, it stalks away!  
 Speak, speak!  
 Lo, where it comes again!  
   
 If thou art privy to thy country's fate,  
 Which happily foreknowing may avoid,  
 O, speak!  
 Stay, and speak!  
- Stop it, Marcellus!  
 Time be thine,  
 And thy best graces spend it at thy will!  
 [aside] A little more than kin, and less than kind!  
 O that this too too solid flesh would melt,  
 Thaw, and resolve itself into a dew!  
  
 How weary, stale, flat, and unprofitable  
 Seem to me all the uses of this world!  
 ah, fie!  
 Why, she would hang on him  
 As if increase of appetite had grown  
 By what it fed on; and yet, within a month-  
 Let me not think on't!

(продолжение см. в файле)

# Задание 3:

### Код

import re  
  
ans = open(r"C:\Users\admin\PycharmProjects\Parser\out", "w", encoding="utf-8")  
file = open(r"C:\Users\admin\PycharmProjects\Parser\txt2", "r", encoding="utf-8")  
string = file.read()  
strings = re.split("[^0**-**9]", string)  
container = []  
substitute = []  
for i in range(len(strings)):  
 st = strings[i]  
 if len(st) % 2 != 0:  
 length = int((len(st) - 1) / 2)  
 template = "(.{" + str(length) + "})0(.{" + str(length) + "})"  
 temp = re.findall(template, st)  
 if len(temp) > 0:  
 if len(temp[0]) == 2:  
 container.append(temp[0][0] + temp[0][1] + "0")  
 substitute.append(temp[0][0] + "0" + temp[0][1])  
for i in range(len(container)):  
 string = re.sub(substitute[i], container[i], string)  
ans.write(string)

### Тесты:

1)

Ввод: 1010101 + 5000001

Вывод: 1011010 + 5000010

2)

Ввод: abc1230123dfe

Вывод: abc1231230dfe

3)

Ввод: abc1230123d101fe

Вывод: abc1231230d110fe

4)

Ввод: 1230123 1230124

Вывод: 1231230 1231240

5)

Ввод: 1.011

Вывод: 1.011

# Вывод:

В результате выполнения данной лабораторной работы я научился работать с регулярными выражениями и применять их, используя я.п. Питон. Теперь, глядя на шаблоны поиска, я способен примерно оценить, что они делают.