Федеральное государственное образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Дисциплина «Информатика»

**Отчёт по лабораторной работе №3**

**Вариант №115**

Регулярные выражения

**Выполнил:**Смирнов Вадим Константинович

Группа: P3119

**Проверил:**  
Авксентьева Елена Юрьевна, доцент

г. Санкт-Петербург

2024

**Оглавление**

[**Оглавление** 2](#_Toc114735310)

[**Задание** 3](#_Toc114735311)

[**Решение** 4](#_Toc114735312)

[**Дополнительное задание**](#_Toc114735313) 5

[**Вывод**](#_Toc114735314) 6

[**Список литературы** 7](#_Toc114735310)

**Задание**

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. Все 5 тестов необходимо показать при защите.

3) Программа должна считать число 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.

Смайлик варианта 115 – ( ;-/ )

**Решение**

import re

pattern = ";</"

for i in range(5):

    string = input()

    print(len(re.findall(pattern, string, flags=0)))

**Дополнительное задание**

Дополнительное задание 1

import re

pattern = r'(\w+)(\s+\1\b)+'

'''

\w+ - для "захвата" символов

\s+ - ищет пробелы (больше одного)

\1 - первая группа букв, то что нашло (\w+)

\b - обозначение границы слова

плюс в конце обозначает что последовательность может повторяться сколько угодно раз

'''

for i in range(5):

    string=input()

    print(re.sub(pattern, '\\1', string))

# \\1 ссылается на первое совпадаемое слово

# Довольно распространённая ошибка ошибка – это лишний повтор повтор слова слова. Смешно, не не правда ли? Не нужно портить хор хоровод.

Дополнительное задание 2

import re

string = "Петров П.П. P0000\nАнищенко А.А. P33113\nПримеров Е.В. P0000\nИванов И.И. P0000\n"

group = "P0000"

pattern = r"[А-ЯЁа-яё-]+ ([А-ЯЁ]\.)\1 " + group.strip() + r"(\n)"

print(re.sub(pattern, "", string).strip())

# последовательность букв найти

# найти заглавную букву с точкой

# круглые скобки как группа захвата символов

# найти повторения (.1 опирается на предыдущее)

**Вывод**

В ходе выполнения работы я закрепил и улучшил свои знания по теме «Регулярные выражения», убедился в их большой необходимости.

**Список литературы**

## Майкл Фицжеральд "Введение в регулярные выражения" (2012)

## Феликс Лопез, Виктор Ромеро "Освоение регулярных выражений на Python" (2014)