Национальный исследовательский университет компьютерных технологий, механики и оптики

Факультет ПИиКТ

Лабораторная работа №4

«Регулярные выражения»

Вариант №11

Работу выполнил: Кулаков Никита

Группа: P3130

Преподаватель: Балакшин П.В.

Город: Санкт-Петербург

2020 год

Задание №1:

**Задание на 70 баллов**

1) Реализуйте программный продукт на языке Python, используя регулярные выражения по варианту,

представленному в таблице.

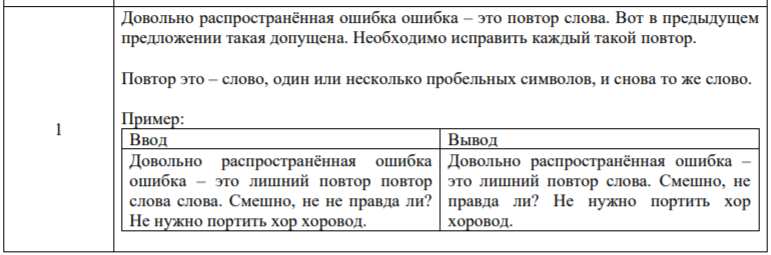
2) Для своей программы придумайте минимум 5 тестов. Каждый тест является отдельной сущностью,

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.



Решение:

import re

text = open('try\_one\_10', encoding='utf-8')

text\_w = text.read()

print(re.sub(r'\b(\w+)(\s+\1)\b', r'\1', text\_w, flags=re.I))

Тесты:

Довольно распространенная ошибка ошибка - это лишний повтор повтор слова слова. Смешно, не правда ли? Не нужно портить хор хороводом

Довольно распространенная ошибка

Ошибка - это лишний повтор слова! Слова! Смешно, не правда ли? Не нужно портить хор

хороводом

Давайте будем-будем более дружелюбными,

дружелюбными, именно поэтому нужно на на всех лезть обниматься

Как

Как

Ты это сделал?

Сделал?

Давай ты ты повторишь при мне

Окей окейс ну вот и и все все, теперь мы будем читать все заново. Все

все заново. Опять опять!

Результат:

Довольно распространенная ошибка - это лишний повтор слова. Смешно, не правда ли? Не нужно портить хор хороводом

Довольно распространенная ошибка - это лишний повтор слова! Слова! Смешно, не правда ли? Не нужно портить хор

хороводом

Давайте будем-будем более дружелюбными,

дружелюбными, именно поэтому нужно на всех лезть обниматься

Как

Ты это сделал?

Сделал?

Давай ты повторишь при мне

Окей окейс ну вот и все, теперь мы будем читать все заново. Все заново. Опять!

(все правильно)

**Доп. задание №1 (+12 баллов)**

1) Для выполнения задания скачайте текстовый документ по ссылке, соответствующей вашему

варианту

**https://drive.google.com/file/d/1lBo2ExJSIWcgoGx4Ton517dlIr2Q8qC/view**

* **Макбет**

2) Реализуйте программный продукт на языке Python, используя регулярные выражения, выделите

все предложения из вашего текста, соответствующие критерию по вашему варианту.

3) Придумайте способ, как проверить корректность работы вашего регулярного выражения и

проверьте его.

Вариант 3:

**3** | Все вопросительные предложения, в которых слов меньше 6

Решение:

import re

text = open('try\_one\_10', encoding='utf-8')

text\_w = '. ' + text.read()

matches = re.findall(r"(?<=[?.!])(?:\s?[,:;-]?\s+[\w'-]+){1,5}\?", text\_w)

for i in matches:

i = i.replace('\n', ' ')

while ' ' in i:

i = i.replace(' ', ' ')

i = i[1:]

i += ' (True)' if i.count(' ') - i.count('-') <= 4 else '(False)'

print(i)

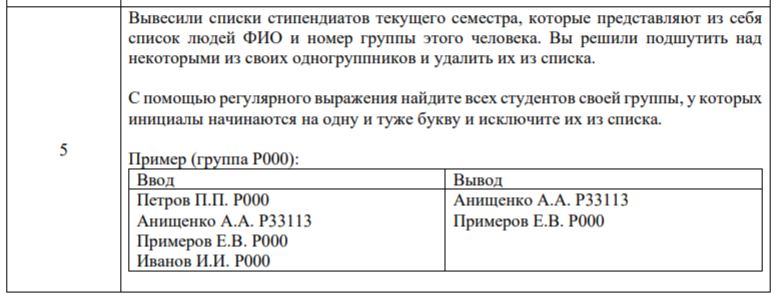
**Доп. задание №2 (+18 баллов)**

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.



Решение:

import re

text = open('try\_one\_10', encoding='utf-8')

regex = r"\b(\w)\w\* \1\.\1\. \w+\b"

group = r"\bP3130\b"

text\_w = text.readline()

while text\_w:

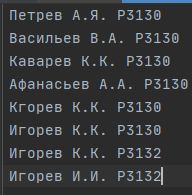
if re.search(group, text\_w):

if not re.match(regex, text\_w):

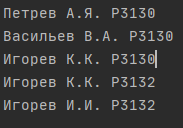
print(text\_w.replace("\n", ""))

else:

print(text\_w.replace("\n", "")) Ввод:

 text\_w = text.readline()

й Вывод:



В процессе выполнения данной лабораторной работы я узнал, что такое регулярные выражения и где их можно использовать. Научился пользоваться тернарным оператором.