Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

*Факультет Программной инженерии и компьютерной техники*

**Лабораторная работа №3**

Регулярные выражения

Вариант №336765

Группа: P3131

Выполнил: Хайкин О. И.

Проверил:

к.т.н. преподаватель Белозубов А.В.

Санкт-Петербург

2021г

Оглавление

[Задание 3](#_Toc85978632)

[Задание на 60 баллов (Смайлики) 3](#_Toc85978633)

[Доп. задание №1 (+18 баллов) 3](#_Toc85978634)

[Доп. задание №2 (+22 баллов) 3](#_Toc85978635)

[Выполнение работы 5](#_Toc85978636)

[Задание №1 5](#_Toc85978637)

[Доп. задание №1 5](#_Toc85978638)

[Доп. задание №2 5](#_Toc85978639)

[Вывод программы 6](#_Toc85978640)

[Заключение 9](#_Toc85978641)

# Задание

## Задание на 60 баллов (Смайлики)

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.

Для варианта 336765 смайлик: :<O

## Доп. задание №1 (+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. 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.

Для варианта 336765 задание:

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.

## Доп. задание №2 (+22 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

Для варианта 336765 задание:

С помощью регулярного выражения найти в тексте слова, в которых встречается строго одна гласная буква (встречаться она может несколько раз). Пример таких слов: окно, трава, молоко, etc.

После чего данные слова требуется отсортировать по увеличению длины слова.

# Выполнение работы

## Задание №1

import re

print("Задание №1:")

test1=[

":<O0^,:<0^^^:<<O:O<:<O:<O:<<O",

":<OO<O::<0:<):):( : <O :<O ",

" :<О,: <O",

"::<О :<0 :>O",

":<O O>: <>O:<O"

]

for i in range (len(test1)):

print("test №", i+1)

print("Ищем: :<O")

print("Вход:", test1[i])

print(" Ответ:", len(re.findall(":<O", test1[i])))

print()

print(“Все ответы сошлись с найденными вручную”)

## Доп. задание №1

import re

print("Доп. задание №1:")

test2=[

"Студент Вася вспомнил, что на своей лекции Балакшин П.В. упоминал про старшекурсников, которые будут ему помогать: Анищенко А.А. и Машина Е.А.",

"По моему объективному мнению Кореневская Е.В. лучший школьный учитель физики, а Печелийская О.А. - лучший учитель русского языка.",

"Если вы напишите Хайкин О. И., то ничего не будет найдено",

"А вот если Хайкин О.И., то всё в порядке",

"На английском фамилии не воспринимаются: Haikin O.I."

]

for i in range (len(test2)):

print()

print("test №", i+1)

print("Ищем фамилии")

print("Вход:", test2[i])

c=re.findall(r'(\b[А-Я][а-я]+)\s[А-Я]{1}\.[А-Я]{1}\.', test2[i])

c=sorted(c)

print(" Ответ:", end="")

for i in range (len(c)):

print(" ", c[i], end=" ")

if c==[]:

print(" Фамилий не найдено")

print()

print(“Все ответы сошлись с найденными вручную”)

## Доп. задание №2

import re

print("Доп. задание №2")

test3=[

"Классное слово - обороноспособность, которое должно идти после слов: трава и молоко",

"Вот объясните мне: кому нужно искать в тексте слова, содержащие одну гласную букву? В чём сокральный смысл?",

"Вот сидишь такой и думаешь о том, как же хочется найти в тексте все слова, содержащие только одну гласную букву, это ведь так необходимо",

"Лично я подобным каждый день занимаюсь",

"Просто для заметки: всё сказанное выше было сказано шутки ради"

]

for i in range (len(test3)):

e=test3[i]

g=[]

print()

print("test №", i+1)

print("Ищем слова с одной гласной")

print("Вход:", test3[i])

print(" Ответ:", end="")

e=re.sub(r"[^А-Яа-яёЁ ]+", "", e)

e=e.split(" ")

for j in range (len(e)):

f=re.findall(r'[ЮИЯЭОАЫЕУЁюияэоаыеуё]', e[j])

if all(k==f[0] for k in f) and len(f)>=2:

g.append(e[j])

g=sorted(g, key=len)

for j in range (len(g)):

print(" ", g[j], end=" ")

if g==[]:

print(" Таких слов нет")

print(“Все ответы сошлись с найденными вручную”)

## Доп. задание (дробные числа)

import re

print("Доп. задание: найти дробные числа")

h="0.5 1.5 13363.05 13,363.05 13 363.05, 1334.05, 13,34.05 13 34.05 1531541395618956.3"

print("Вход: ", h)

print("Ответ:")

t=re.findall(r'[^,0-9]\b(\d{0,3}(?:[ ,]?\d{3})\*\.\d+)', h)

for i in range (len(t)):

print(t[i])

print("Не проходят числа, в которых запятая/пробел стоят в неправильных местах (т.е. числа по типу 13,34.05 или 13 34.05)")

## Вывод программы

Задание №1:

test № 1

Ищем: :<O

Вход: :<O0^,:<0^^^:<<O:O<:<O:<O:<<O

Ответ: 3

test № 2

Ищем: :<O

Вход: :<OO<O::<0:<):):( : <O :<O

Ответ: 2

test № 3

Ищем: :<O

Вход: :<О,: <O

Ответ: 0

test № 4

Ищем: :<O

Вход: ::<О :<0 :>O

Ответ: 0

test № 5

Ищем: :<O

Вход: :<O O>: <>O:<O

Ответ: 2

Все ответы сошлись с найденными вручную

Доп. задание №1:

test № 1

Ищем фамилии

Вход: Студент Вася вспомнил, что на своей лекции Балакшин П.В. упоминал про старшекурсников, которые будут ему помогать: Анищенко А.А. и Машина Е.А.

Ответ: Анищенко Балакшин Машина

test № 2

Ищем фамилии

Вход: По моему объективному мнению Кореневская Е.В. лучший школьный учитель физики, а Печелийская О.А. - лучший учитель русского языка.

Ответ: Кореневская Печелийская

test № 3

Ищем фамилии

Вход: Если вы напишите Хайкин О. И., то ничего не будет найдено

Ответ: Фамилий не найдено

test № 4

Ищем фамилии

Вход: А вот если Хайкин О.И., то всё в порядке

Ответ: Хайкин

test № 5

Ищем фамилии

Вход: На английском фамилии не воспринимаются: Haikin O.I.

Ответ: Фамилий не найдено

Все ответы сошлись с найденными вручную

Доп. задание №2

test № 1

Ищем слова с одной гласной

Вход: Классное слово - обороноспособность, которое должно идти после слов: трава и молоко

Ответ: идти слово трава должно молоко обороноспособность

test № 2

Ищем слова с одной гласной

Вход: Вот объясните мне: кому нужно искать в тексте слова, содержащие одну гласную букву? В чём сокральный смысл?

Ответ: букву тексте

test № 3

Ищем слова с одной гласной

Вход: Вот сидишь такой и думаешь о том, как же хочется найти в тексте все слова, содержащие только одну гласную букву, это ведь так необходимо

Ответ: букву сидишь тексте только

test № 4

Ищем слова с одной гласной

Вход: Лично я подобным каждый день занимаюсь

Ответ: Таких слов нет

test № 5

Ищем слова с одной гласной

Вход: Просто для заметки: всё сказанное выше было сказано шутки ради

Ответ: Просто

Все ответы сошлись с найденными вручную

Доп. задание: найти дробные числа

Вход: 0.5 1.5 13363.05 13,363.05 13 363.05, 1334.05, 13,34.05 13 34.05 1531541395618956.3

Ответ:

1.5

13363.05

13,363.05

13 363.05

1334.05

34.05

1531541395618956.3

Не проходят числа, в которых запятая/пробел стоят в неправильных местах (т.е. числа по типу 13,34.05 или 13 34.05)

# Заключение

В ходе выполнения данной работы я научился использовать регулярные выражения в целом и с использованием языка программирования Python в частности.