Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»

**Факультет программной инженерии и компьютерной техники**

**Информатика**

Лабораторная работа №3

«Регулярные выражения»

Выполнил: Герасимов Артём Кириллович

Группа: P3108

Вариант: 339052

Преподаватель: Малышева Татьяна Алексеевна

Санкт-Петербург

2021

Оглавление

[Задание 2](#_Toc88653255)

[Ход работы 4](#_Toc88653256)

[Основное задание 4](#_Toc88653257)

[Доп. задание 1 5](#_Toc88653258)

[Доп. задание 2 6](#_Toc88653259)

[Вывод 7](#_Toc88653260)

[Список литературы 7](#_Toc88653261)

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.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Номер в ИСУ % 5 | Глаза | Номер в ИСУ % 4 | Нос | Номер в ИСУ % 7 | Рот |
| 2 | X | 0 | - | 0 | ( |

**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.

|  |  |
| --- | --- |
| Номер в ИСУ % 6 | Задание |
| 4 | 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 | Вывод | | 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. | |

**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.

|  |  |
| --- | --- |
| Номер в ИСУ % 4 | Задание |
| 0 | Написать регулярное выражение, которое проверяет корректность email и в качестве ответа выдаёт почтовый сервер (почтовый сервер – часть email идущая после «@»).    Для простоты будем считать, что почтовый адрес может содержать в себе буквы, цифры, «.» и «\_», а почтовый сервер только буквы и «.». При этом почтовый сервер, обязательно должен содержать верхний уровень домена («.ru», «.com», etc.)    Пример:   |  |  | | --- | --- | | Ввод | Вывод | | students.spam@yandex.ru | yandex.ru | | example@example | Fail! | | example@example.com | example.com | |

Ход работы

Основное задание

**Исходный код:**

import re  
  
smile = 'X-\('  
smileTests = ['Мой номер ИСУ 339052, поэтому мой смайлик - X-(.',  
 'Мой друг отправил мне X-( X-( X-(.',  
 'У тебя ошибка. Правильно X-(, а не X-C.',  
 'X-( и X-) - показывают абсолютно разные эмоции.',  
 'Мне написали X-C.']  
i = 0  
for test1 in smileTests:  
 i += 1  
 smileMatch = re.findall(smile, test1)  
 smileCount = len(smileMatch)  
 print(f'Текст: {test1} \nКол-во смайликов: {smileCount} \n')

**Результат:**

Текст: Мой номер ИСУ 339052, поэтому мой смайлик - X-(.

Кол-во смайликов: 1

Текст: Мой друг отправил мне X-( X-( X-(.

Кол-во смайликов: 3

Текст: У тебя ошибка. Правильно X-(, а не X-C.

Кол-во смайликов: 1

Текст: X-( и X-) - показывают абсолютно разные эмоции.

Кол-во смайликов: 1

Текст: Мне написали X-C.

Кол-во смайликов: 0

Доп. задание 1

**Исходный код:**

import re  
  
textTests = ['Уважаемые студенты! В эту субботу в 15:00:23 планируется доп. занятие на 2 часа. '  
 'То есть в 27:68:01 оно уже точно кончится.',  
 'Я буду дома примерно в 17:10.',  
 'Рейс перенесли с 16:30 на 18:00.',  
 'Мой часы показывают, что сейчас 14:30:10.',  
 'Встреча состоится 03.10 ровно в 13:00.',  
 'Я уйду через часа 2, у меня поезд в 19:20.',  
 'Время 23:45:54 указано правильно',  
 'Время 14:47:70 - указано неправильно']  
  
timeCheck = r'(?<![:\d])([01]\d|2[0-3])(:[0-5]\d){1,2}(?![:\d])'  
i = 0  
for test2 in textTests:  
 i += 1  
 textFinal = re.sub(timeCheck, '(TBD)', test2)  
 print(f'Исходный текст {i}: {test2} \nИзмененный текст {i}: {textFinal} \n')

**Результат:**

Исходный текст 1: Уважаемые студенты! В эту субботу в 15:00:23 планируется доп. занятие на 2 часа. То есть в 27:68:01 оно уже точно кончится.

Измененный текст 1: Уважаемые студенты! В эту субботу в (TBD) планируется доп. занятие на 2 часа. То есть в 27:68:01 оно уже точно кончится.

Исходный текст 2: Я буду дома примерно в 17:10.

Измененный текст 2: Я буду дома примерно в (TBD).

Исходный текст 3: Рейс перенесли с 16:30 на 18:00.

Измененный текст 3: Рейс перенесли с (TBD) на (TBD).

Исходный текст 4: Мой часы показывают, что сейчас 14:30:10.

Измененный текст 4: Мой часы показывают, что сейчас (TBD).

Исходный текст 5: Встреча состоится 03.10 ровно в 13:00.

Измененный текст 5: Встреча состоится 03.10 ровно в (TBD).

Исходный текст 6: Я уйду через часа 2, у меня поезд в 19:20.

Измененный текст 6: Я уйду через часа 2, у меня поезд в (TBD).

Исходный текст 7: Время 23:45:54 указано правильно

Измененный текст 7: Время (TBD) указано правильно

Исходный текст 8: Время 14:47:70 - указано неправильно

Измененный текст 8: Время 14:47:70 - указано неправильно

Доп. задание 2

**Исходный код:**

import re  
  
emailTests = ['artyom0104@gmail.com', 'pupalupa', 's339052@niuitmo.ru',  
 'dima\_1912@yandex', 'se.ifmo.ru', 'VaNya2010@outlook.com']  
  
emailCheck = r'[a-z0-9\_.]+@[a-z]+\.[a-z]+'  
i = 0  
for test3 in emailTests:  
 i += 1  
 emailMatch = re.fullmatch(emailCheck, test3, re.IGNORECASE)  
 if emailMatch:  
 emailServer = re.split('@', test3)[1]  
 print(f'Тестовый email {i}: {emailServer}')  
 else:  
 print(f'Тестовый email {i}: Fail!')

**Результат:**

Тестовый email 1: gmail.com

Тестовый email 2: Fail!

Тестовый email 3: niuitmo.ru

Тестовый email 4: Fail!

Тестовый email 5: Fail!

Тестовый email 6: outlook.com

Вывод

Во время выполнения данной лабораторной работы я изучил как работают регулярные выражения, а также написал несколько небольших программ на языке Python, в которых они используются.



**Рисунок 1 - американский математик и логик Стивен Коул Клини**

Список литературы

1. [ShashkovS](https://habr.com/ru/users/ShashkovS/) Регулярные выражения в Python от простого к сложному. Подробности, примеры, картинки, упражнения. 2018г. Режим доступа:

<https://habr.com/ru/post/349860/>

1. Балакшин П.В., Соснин В.В., Калинин И.В., Малышева Т.А., Раков С.В., Рущенко Н.Г., Дергачев А.М. Информатика: лабораторные работы и тесты. – СПб: Университет ИТМО, 2019. – 56 с. - Режим доступа:

<https://books.ifmo.ru/file/pdf/2464.pdf>