#### 31.2 Регулярные выражения: модуль re и его методы

Практика

Задача 1. Скороговорка

Дан текст вот такой английской скороговорки:

How much wood would a woodchuck chuck if a woodchuck could chuck wood?

С помощью модуля re реализуйте программу, которая выполняет следующие действия по порядку:

1. Определить, начинается ли текст с шаблона wo.
2. Найти первое упоминание шаблона wo в тексте.
3. Определить содержимое найденной по шаблону подстроки из пункта 2.
4. Найти позицию, с которого начинается первое упоминание шаблона wo.
5. Найти позицию, на которой заканчивается первое упоминание шаблона wo.
6. Получить список из каждого упоминания шаблона wo в тексте.
7. Заменить в тексте все совпадения по шаблону wo на слово «ЗАМЕНА».

По каждому действию вывести соответствующий результат. Не используйте методы строк. Не забывайте использовать приписку r в шаблонах.

Ожидаемый результат работы программы:

Поиск шаблона в начале строки: None

Поиск первого найденного совпадения по шаблону: <re.Match object; span=(9, 11), match='wo'>

Содержимое найденной подстроки: wo

Начальная позиция: 9

Конечная позиция: 11

Список всех упоминаний шаблона: ['wo', 'wo', 'wo', 'wo', 'wo']

Текст после замены: How much ЗАМЕНАod ЗАМЕНАuld a ЗАМЕНАodchuck chuck if a ЗАМЕНАodchuck could chuck ЗАМЕНАod?

Задача 2. Экранирование спецсимволов

В уроке при написании шаблона мы брали только обычные строки, без всяких специальных знаков. Часть из этих спецсимволов является неотъемлемой частью регулярных выражений (о чём мы поговорим уже в следующем уроке), они даже выделяются своими цветами при написании паттернов. Например:



Поэтому если мы хотим найти в тексте спецсимвол, а не использовать его как часть паттерна, то нужно его дополнительно экранировать — добавить обратный слеш перед этим знаком. Например, если нам нужно будет найти шаблон wd+?. в виде полноценного текста, то это будет выглядеть так:



А теперь сама задача. Дан немного изменённый текст скороговорки:

How much \wwood+?, would a \wwood+?chuck \dwwood+, chuck if a \wwood+?,chuck could chuck \wwood?,

Напишите программу, которая выводит список из всех упоминаний подстроки \wwood+?,

Ожидаемый результат:

Список всех упоминаний шаблона: ['\\wwood+?,', '\\wwood+?,']