Лабораторная работа №5

ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ИНФОРМАЦИОННЫХ СИСТЕМ

РЕГУЛЯРНЫЕ ВЫРАЖЕНИЯ

Вариант 13

**Цель работы**: разработать приложение по обработке строк с использованием регулярных выражений.

**Задания**

1. Общее для всех вариантов: написать регулярное выражение, определяющее является ли данная строка строкой "abcdefghijklmnopqrstuv5320736" или нет.

– пример правильных выражений: abcdefghijklmnopqrstuv5320736.

– пример неправильных выражений: abcdefghijklmnoasdfasdpqrstuv5320736.

1. Написать регулярное выражение, определяющее является ли данная строчка валидным URL адресом. В данной задаче правильным URL считаются адреса http и https, явное указание протокола также может отсутствовать. Учитываются только адреса, состоящие из символов, т.е. IP адреса в качестве URL не присутствуют при проверке. Допускаются поддомены, указание порта доступа через двоеточие, GET запросы с передачей параметров, доступ к подпапкам на домене, допускается наличие якоря через решетку. Однобуквенные домены считаются запрещенными. Запрещены спецсимволы, например «–» в начале и конце имени домена. Запрещен символ «\_» и пробел в имени домена. При составлении регулярного выражения ориентируйтесь на список правильных и неправильных выражений, заданных ниже.

– пример правильных выражений: http://www.example.com, http://example.com

– пример неправильных выражений: Just Text, http://a.com.

**Решение**

**1)** Для определения является ли данная строка валидным URL адресом, можно использовать следующее регулярное выражение:^ ­- символ, указывающий на начало строки

abcdefghijklmnopqrstuv5320736 – это конкретная последовательность символов, которая должна точно совпадать с введенной по условию строкой.

$ - символ, обозначающий конец строки.

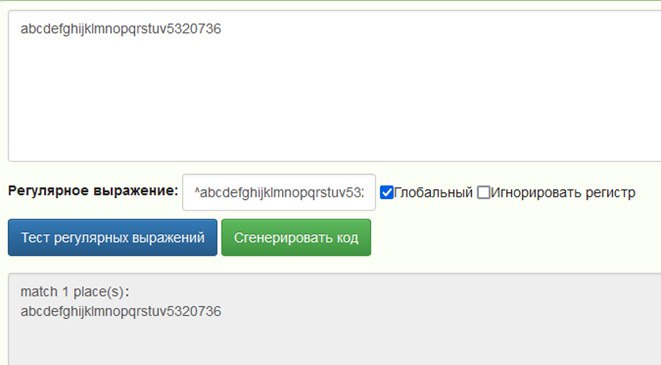


Рисунок 1 ­– Совпадение найдено

Теперь попробуем добавить еще один лишний символ в нашу строку:

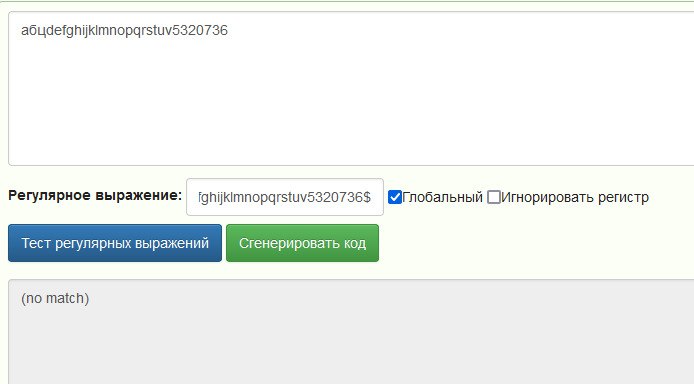


Рисунок 2 ­– Совпадений не найдено

Это регулярное выражение позволяет проверить является ли введенная строка строкой или нет.

**2)** Для определения является ли данная строка валидным URL адресом, можно использовать следующее регулярное выражение: ^(https?://)?(?:[a-zA-Z0-9](?:[a-zA-Z0-9-]{0,61}[a-zA-Z0-9])?\.)+[a-zA-Z]{2,}(?::\d{1,5})?(?:/\S\*)?(?:\?\S\*)?(?:#\S\*)?$

* ^: Начало строки.
* (https?://)?: Опциональная часть, которая представляет собой протокол http или https, с двойными слэшами.
* (?:[a-zA-Z0-9](?:[a-zA-Z0-9-]{0,61}[a-zA-Z0-9])?\.)+: Один или более поддоменов, каждый из которых начинается с буквы или цифры, может содержать от 1 до 62 букв, цифр или дефисов, и заканчивается буквой или цифрой, за которым следует точка.
* [a-zA-Z]{2,}: Домен верхнего уровня, состоящий как минимум из двух букв.
* (?::\d{1,5})?: Опциональная часть, представляющая порт, указанный после двоеточия, состоящий из одной или более цифр.
* (?:/\S\*)?: Опциональная часть, представляющая путь после доменного имени, начинающийся с косой черты и содержащий ноль или более символов, не являющихся пробелом.
* (?:\?\S\*)?: Опциональная часть, представляющая строку запроса, начинающуюся с вопросительного знака и содержащую ноль или более символов, не являющихся пробелом.
* (?:#\S\*)?: Опциональная часть, представляющая якорь, начинающийся с решетки и содержащий ноль или более символов, не являющихся пробелом.
* $: Конец строки.

Это регулярное выражение удовлетворяет описанным в задании требованиям для валидного URL адреса.

Для проверки правильности регулярного выражения запишем адрес http://example.com в онлайн-сервисе для проверки регулярных выражений.

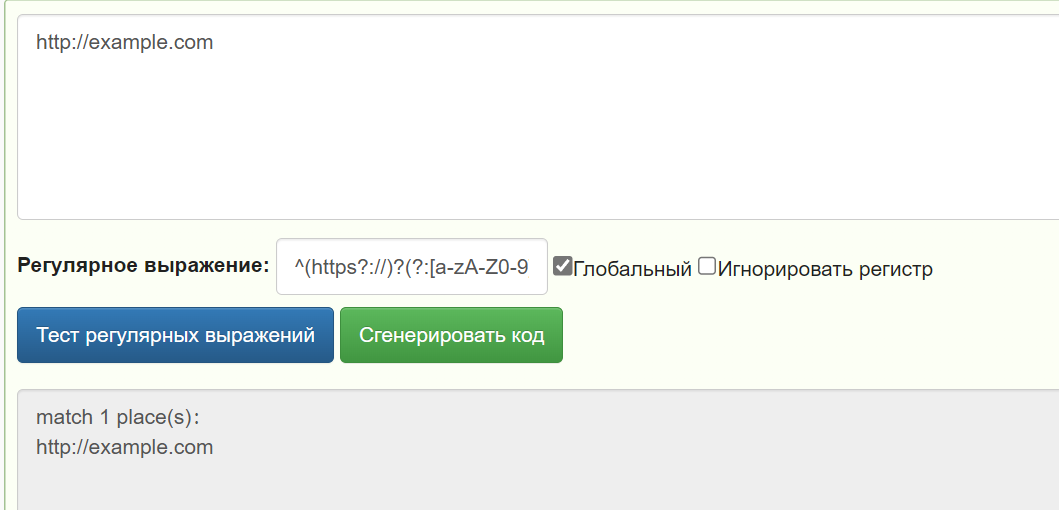


Рисунок 3 ­– Совпадение найдено

А теперь попробуем добавить еще один лишний символ в наш адрес:

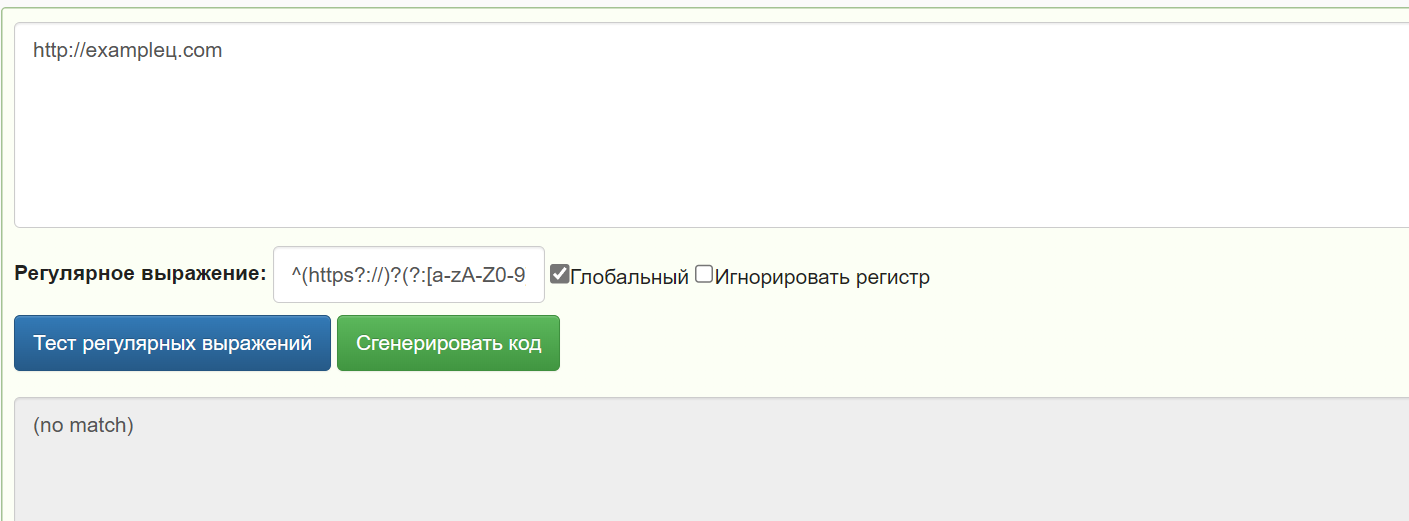


Рисунок 4 ­– Совпадений не найдено

Таким образом, это регулярное выражение действительно позволяет проверить является ли строка IP адресом или нет.

**Вывод**: регулярные выражения используются в различных языках программирования, таких как Perl, Python, JavaScript, Java, C#, и других. Они позволяют искать определенные шаблоны или последовательности символов в строках текста и выполнять различные операции с этими шаблонами, такие как поиск, замена, извлечение и проверка.