1. What is the name of the feature responsible for generating Regex objects?

The re.compile() function returns Regex objects

1. Why do raw strings often appear in Regex objects?

Raw strings are used so that backslashes do not have to be escaped.

1. What is the return value of the search() method?

The search() method returns Match objects.

1. From a Match item, how do you get the actual strings that match the pattern?

findall() module is used to search for “all” occurrences that match a given pattern

1. In the regex which created from the r'(\d\d\d)-(\d\d\d-\d\d\d\d)', what does group zero cover? Group 2? Group 1?

Group 0 is the entire match, group 1 covers the first set of parentheses, and group 2 covers the second set of parentheses.

1. In standard expression syntax, parentheses and intervals have distinct meanings. How can you tell a regex that you want it to fit real parentheses and periods?

Periods and parentheses can be escaped with a backslash: \., \(, and \).

1. The findall() method returns a string list or a list of string tuples. What causes it to return one of the two options?

If the regex has no groups, a list of strings is returned. If the regex has groups, a list of tuples of strings is returned.

1. In standard expressions, what does the | character mean?

The | character signifies matching "either, or" between two groups.

1. In regular expressions, what does the character stand for?

Regular expressions use two types of characters:

a) Meta characters: As the name suggests, these characters have a special meaning, similar to \* in wild card.

b) Literals (like a,b,1,2…)

1. 10.In regular expressions, what is the difference between the + and \* characters?

+ 1 or more occurrences of the pattern to its left

\* 0 or more occurrences of the pattern to its left

1. What is the difference between {4} and {4,5} in regular expression?

{4} is exact match

{4,5} is ,minimum 4 and max 5 matches

1. What do you mean by the \d, \w, and \s shorthand character classes signify in regular expressions?

\d Any digit from 0 to 9

\w any word character [a-zA-Z0-9]

\s space character

1. What do means by \D, \W, and \S shorthand character classes signify in regular expressions?

\D Any character except digit

\W Any word character

\S any character except space character

1. What is the difference between .\*? and .\*?

\*Any number including zero

? Atmost one (either or 1)

1. What is the syntax for matching both numbers and lowercase letters with a character class?

Numbers[0-9]

Lowercase[a-z]

1. What is the procedure for making a normal expression in regax case insensitive?

We can pass re.IGNORECASE to the flags parameter of search, match, or sub

1. What does the . character normally match? What does it match if re.DOTALL is passed as 2nd argument in re.compile()?

.(dot) 🡪 Any character including special characters

(Dot.) In the default mode, this matches any character except a newline. If the DOTALL flag has been specified, this matches any character including a newline.

1. If numReg = re.compile(r'\d+'), what will numRegex.sub('X', '11 drummers, 10 pipers, five rings, 4 hen') return?

import re

numReg = re.compile(r'\d+')

numReg.sub('X', '11 drummers, 10 pipers, five rings, 4 hen')

o/p: 'X drummers, X pipers, five rings, X hen'

1. What does passing re.VERBOSE as the 2nd argument to re.compile() allow to do?

e.compile(Regular Expression, re.VERBOSE). re.compile() returns a RegexObject which is then matched with the given string.

1. 20. How would you write a regex that match a number with comma for every three digits? It must match the given following:

Ans

import re

numReg = re.compile(r'^\d{1,3}(,\d{3})\*$')

numReg.findall('42')

'42'

'1,234'

'6,368,745'

but not the following:

'12,34,567' (which has only two digits between the commas)

'1234' (which lacks commas)

21. How would you write a regex that matches the full name of someone whose last name is Watanabe? You can assume that the first name that comes before it will always be one word that begins with a capital letter. The regex must match the following:

'Haruto Watanabe'

'Alice Watanabe'

'RoboCop Watanabe'

but not the following:

'haruto Watanabe' (where the first name is not capitalized)

'Mr. Watanabe' (where the preceding word has a nonletter character)

'Watanabe' (which has no first name)

'Haruto watanabe' (where Watanabe is not capitalized)

Ans

import re

def printMatch(name):

pat = re.compile(r'\b[A-Z][a-zA-Z]+\sWatanabe')

if pat.search(name):

print ("matches",name)

else:

print ("does not match",name)

printMatch('Haruto Watanabe')

printMatch('Alice Watanabe')

printMatch('RoboCop Watanabe')

printMatch('haruto Watanabe' )

printMatch('Mr. Watanabe' )

printMatch('Watanabe' )

printMatch('Haruto watanabe' )

output

matches Haruto Watanabe

matches Alice Watanabe

matches RoboCop Watanabe

does not match haruto Watanabe

does not match Mr. Watanabe

does not match Watanabe

does not match Haruto watanabe

22. How would you write a regex that matches a sentence where the first word is either Alice, Bob, or Carol; the second word is either eats, pets, or throws; the third word is apples, cats, or baseballs; and the sentence ends with a period? This regex should be case-insensitive. It must match the following:

'Alice eats apples.'

'Bob pets cats.'

'Carol throws baseballs.'

'Alice throws Apples.'

'BOB EATS CATS.'

but not the following:

'RoboCop eats apples.'

'ALICE THROWS FOOTBALLS.'

'Carol eats 7 cats.'

Ans

import re

def printMatch(name):

pat = re.compile(r'(Alice|Bob|Carol)\s+(eats|pets|throws)\s+(apples|cats|baseballs)\.', re.IGNORECASE)

if pat.search(name):

print ("matches",name)

else:

print ("does not match",name)

printMatch('Alice eats apples.')

printMatch('Bob pets cats.')

printMatch('Carol throws baseballs.')

printMatch('Alice throws Apples.')

printMatch('BOB EATS CATS.')

printMatch('RoboCop eats apples.')

printMatch('ALICE THROWS FOOTBALLS.')

printMatch('Carol eats 7 cats.')

output

matches Alice eats apples.

matches Bob pets cats.

matches Carol throws baseballs.

matches Alice throws Apples.

matches BOB EATS CATS.

does not match RoboCop eats apples.

does not match ALICE THROWS FOOTBALLS.

does not match Carol eats 7 cats.