

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

`re.compile()`

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

So that backslashes do not have to escape. Using raw string backslash is treated as a literal character

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

Returns the matched objects

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

Match objects have a `Group` method that will return the actual matched text from the searched string

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

Group zero cover the entire match

Group 1 covers the first set in parenthesis and Group 2 covers the second parenthesis

6. 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?

Using escape character `\.` `\(` `\)`

7. 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 find all method has regex with no groups ; a list of strings is returned

If the find all method has regex with groups ; a list of tuples of string is returned

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

The character `|` signifies 'either, or' between groups

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

`?` character can either mean 'match zero or one' of the preceding group

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

`*` Matches at least one or more preceding characters

`+` Matched zero or more preceding characters

11. What is the difference between {4} and {4,5} in regular expressions?

{4} 4 repetition of preceding character

{4,5} at least 4 repetition and at most 5 repetition of the preceding character

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

\d Matches any digit. Equivalent to the class [0-9]

\w Matches any alphanumeric character. Equivalent to class [a-zA-Z0-9_]

\s Matches any whitespace character, Equivalent to class [\t\n\r\f\v]

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

\D Matches any non-digit character, Equivalent to the class [^0-9]

\W Matches any non-alphanumeric character. Equivalent to the class [^a-zA-Z0-9_]

\S Matches any non-whitespace character. Equivalent to the class [^\t\n\r\f\v]

14. What is the difference between .*? and .*?

? is a non greedy qualifier. The engine advances a character at a time, retrying match every step.

.*? matches any character (.) any number of times (*), as few times as possible to make the regex match (?)

If .* matches any character zero or more times til the end of string. Consume as much pattern as possible

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

[a-z0-9]

16. What is the procedure for making a normal expression in regex case insensitive?

Flag : re.IGNORECASE

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

. character matches any character except newline

Flag : re.DOTALL Make '.' special character match any character at all, including a newline

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

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

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

'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)

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.'

name of the feature responsible for generating Regex objects?

2. Why do raw strings often appear in Regex objects?
3. What is the return value of the search() method?
4. From a Match item, how do you get the actual strings that match the pattern?
5. 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?
6. 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?
7. The findall() method returns a string list or a list of string tuples. What causes it to return one of the two options?
8. In standard expressions, what does the | character mean?
9. In regular expressions, what does the character stand for?
10. In regular expressions, what is the difference between the + and * characters?
11. What is the difference between {4} and {4,5} in regular expression?
12. What do you mean by the \d, \w, and \s shorthand character classes signify in regular expressions?
13. What do means by \D, \W, and \S shorthand character classes signify in regular expressions?
14. What is the difference between .*? and .*?
15. What is the syntax for matching both numbers and lowercase letters with a character class?
16. What is the procedure for making a normal expression in regex case insensitive?
17. What does the . character normally match? What does it match if re.DOTALL is passed as 2nd argument in re.compile()?
18. If numReg = re.compile(r'\d+'), what will numRegex.sub('X', '11 drummers, 10 pipers, five rings, 4 hen') return?

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

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

'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)

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.'