**Rishabh Kant Jha’s assignment – 7**

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

re.compile() function

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

r”your string here” : here this helps in Regex as It would not require to enter an extra backslash.

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

True or False, depending on the condition being met or not.

re.search(pattern, target)

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

re.search(pattern, arget).

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?

Group 0 :would return the fully matched string

Group 1 : will return

Group 2 : \d\d\d\d

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 Backslash \ 🡪 \( \) \.

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

In the absence of group, String is returned if they are there then a list of string tuples will be returned.

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

OR

9. In regular expressions, what does the character stand for? *Question is missing*

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

+ used to add one or more occurrences

[a-zA-Z] + [\d] : any string of alphabets succeeded by Numbers

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

{4}: match must be of 4 occurrences

{4,5}: match must of between 4 to 5 occurrences

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

\d : numbers or digits [0-9]

\w : word can be a letter ( lower and uppercase included ), number or underscore [a-zA-Z0-9\_]

\s : whitespace character

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

\D : non digit [^0-9]

\W : Not a word. [^a-zA-Z0-9\_]

\S : Not a whitespace chacter

14. What is the difference between .\*? and .\*? *Question not understood*

.

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

[w]

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

[a-zA-Z0-9\_]: here in it doesn’t matter if the characters are from lower or uppercase

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

. character matches the single characters

re.DOTALL will also match newline characters if passed as 2nd argument in re.complie()

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

X drummers, X pipers, five rings, X hens

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

To ignore comments and whitespaces inside the regular expression.

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)

^\d{1,3}(?:,\d{3})\*$

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)

(r'^[A-Z]\w+ 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.'

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