Kingdom of Saudi Arabia
Ministry of Education
Shaqra University
College of Computing & Information Technology

المملكة العربية السعودية
وزارة التعليم
جامعة شقراء
كلية الحاسب الالي و تقنية المعلومات

LAB – III
Compiler Design

Regular Expressions

(Basics of RE Package in Python)

MAR 2022 – JUNE 2023

DEPARTMENT OF COMPUTER SCIENCE

COLLEGE OF COMPUTING AND INFORMATION TECHNOLOGY

SHAQRA UNIVERSITY

P.O. BOX 15572

Dr. Nayyar Ahmed Khan / Dr. Nouf Altamami

EMAIL:

nayyar@su.edu.sa ---- naltamami@su.edu.sa

Regular Expressions (RE) in Python Basics

A **Reg**ular **Ex**pression (RegEx) is a sequence of characters that defines a search pattern. For example,

^a...s$

The above code defines a RegEx pattern.

The pattern is: **any five-letter string starting with a and ending with s**.

A pattern defined using RegEx can be used to match against a string.

Python has a module named "re" to work with RegEx. Here's an example:

```
import re
pattern = '^a...s$'
#test_string = 'albatross'
#Test with both values of test_string
test_string = 'abyss'
result = re.match(pattern, test_string)
if result:
  print("Search successful.")
else:
  print("Search unsuccessful.")
```

Here, we used re.match() function to search pattern within the test_string. The method returns a match object if the search is successful. If not, it returns None.

Python RegEx

Python has a module named re to work with regular expressions. To use it, we need to import the module.

Exercise I:

re.findall()

The re.findall() method returns a list of strings containing all matches.

```
# Program to extract numbers from a string
import re
string = 'hello 22 hi 98. Howdy 16'
```

```
pattern = '\d+'
result = re.findall(pattern, string)
print(result)
```

re.split()

The re.split method splits the string where there is a match and returns a list of strings where the splits have occurred.

```
import re
string = 'Twelve:12 Eighty nine:89.'
pattern = '\d+'
result = re.split(pattern, string)
print(result)
```

If the pattern is not found, re.split() returns a list containing the original string.

re.sub()

re.sub(pattern, replace, string)

The method returns a string where matched occurrences are replaced with the content of replace variable.

```
# Program to remove all whitespaces
import re
# multiline string
string = 'abc 12\
de 23 \n f45 6'
# matches all whitespace characters
pattern = '\s+'
# empty string
replace = ''
new_string = re.sub(pattern, replace, string)
print(new_string)
```

re.subn()

The re.subn() is similar to re.sub() except it returns a tuple of 2 items containing the new string and the number of substitutions made.

```
# Program to remove all whitespaces
import re
# multiline string
string = 'abc 12\
de 23 \n f45 6'
# matches all whitespace characters
pattern = '\s+'
# empty string
replace = ''
new_string = re.subn(pattern, replace, string)
print(new_string)
```

re.search()

The re.search() method takes two arguments: a pattern and a string.

The method looks for the first location where the RegEx pattern produces a match with the string.

If the search is successful, re.search() returns a match object; if not, it returns None.

match = re.search(pattern, str)

```
import re
string = "Python is fun"
# check if 'Python' is at the beginning
match = re.search('\APython', string)
if match:
  print("pattern found inside the string")
else:
  print("pattern not found")
```

match.group()

The group() method returns the part of the string where there is a match.

```
import re
string = '39801 356, 2102 1111'
# Three digit number followed by space followed by two digit
number
```

```
pattern = '(\d{3}) (\d{2})'
# match variable contains a Match object.
match = re.search(pattern, string)
if match:
  print(match.group())
else:
  print("pattern not found")
```

1. What is Regular Expression?

.....

.....

.....

.....

.....

.....

2. Can you Write a Program to Check the Regular Expression for a Password Verification entered by a User in the Sign In Page? Give a Demo Code in GitHub Account

.....

.....

.....

.....

.....

.....