

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЛЬВІВСЬКА ПОЛІТЕХНІКА»**

Кафедра “Захист інформації”



РЕГУЛЯРНІ ВИРАЗИ

**МЕТОДИЧНІ ВКАЗІВКИ
до лабораторної роботи №3**
з курсу «Програмування скриптовими мовами. Частина 2»
для студентів спеціальності
«Кібербезпека»

*Затверджено
на засіданні кафедри
"Захист інформації"
протокол № 01 від 27.08.2020 р.*

Львів – 2020

Регулярні вирази: Методичні вказівки до лабораторної роботи №3 з курсу «Програмування скриптовими мовами. Частина 2» для студентів спеціальності «Кібербезпека» / Укл. Я. Р. Совин – Львів: Національний університет "Львівська політехніка", 2020. – 7 с.

Укладач: Я. Р. Совин, канд. техн. наук, доцент

Відповідальний за випуск: В. Б. Дудикевич, д.т.н., професор

Рецензенти: А. Я. Горпенюк, канд. техн. наук, доцент
Ю. Я. Наконечний, канд. техн. наук, доцент

Мета роботи – навчитися використовувати регулярні вирази з модуля *re* для опрацювання текстової інформації.

1. ЗАВДАННЯ

2.1. Домашня підготовка до роботи

1. Вивчити теоретичний матеріал з лекції №5 “Регулярні вирази”.

2.2. Виконати в лабораторії

1. Написати і протестувати на декількох рядках регулярні вирази (РВ) для завдань з таблиці 1. Тестові рядки для перевірки регулярних виразів повинні показати, що РВ коректно працюють та відсіюють невірні варіанти.

Завдання

У всіх завданнях, якщо не вказане інше, вважати, що текст представлений символами англійського алфавіту. Слови можуть містити дефіс (-).

1. Написати РВ для знаходження всіх слів у тексті, які містять літеру 'n'.
2. Написати РВ для знаходження всіх слів у тексті, які починаються на літеру 'b'.
3. Написати РВ для знаходження всіх слів у тексті, які закінчуються на літеру 'e'.
4. Написати РВ для знаходження всіх слів у тексті, які містять сполучення літер 'th'.
5. Написати РВ для знаходження всіх слів у тексті, які не містять сполучення літер 'ou'.
6. Написати РВ для знаходження всіх слів у тексті, які містять літеру 'a' або 'r'.
7. Написати РВ для знаходження всіх слів у тексті, які починаються на літеру 'c' та закінчуються на літеру 'd'.
8. Написати РВ для знаходження всіх слів у тексті, які починаються та закінчуються на однакову літеру.
9. Написати РВ для знаходження всіх слів у тексті, які місять літеру 'e' за якою йдуть дві або більше літери 'ss'.
10. Написати РВ для знаходження всіх слів у тексті, які місять більше 5 приголосників.
11. Написати РВ для знаходження всіх слів у тексті, які місять не менше 3 голосників.
12. Написати РВ для знаходження всіх слів у тексті, які місять літеру 'r' та літеру 'i', яка повинна зустрічатися пізніше 't'.
13. Написати РВ для знаходження всіх слів у тексті, які не місять літеру 'k'.
14. Написати РВ для знаходження всіх слів у тексті, які не місять літеру 'h' та 'g'.
15. Написати РВ для знаходження всіх слів у тексті, які місять літеру 's' та не місять літеру 't'.
16. Написати РВ для знаходження всіх слів у тексті, які починаються на голосну.
17. Написати РВ для знаходження всіх слів у тексті, які починаються на приголосну.
18. Написати РВ для знаходження всіх слів у тексті довжиною рівно 6 символів.
19. Написати РВ для знаходження всіх слів у тексті довжиною 6 і більше символів.
20. Написати РВ для знаходження всіх слів у тексті з двома одинаковими літерами підряд.
21. Написати РВ для знаходження всіх слів у тексті довжиною менше 6 символів.
22. Написати РВ для знаходження всіх слів у тексті з 3-ма голосними підряд.
23. Написати РВ для знаходження всіх слів у тексті з 5-ма приголосними підряд.
24. Написати РВ для знаходження всіх слів у тексті у яких всі літери різні.
25. Написати РВ для знаходження всіх слів у тексті, які місять літеру 'd', але не на початку чи в кінці слова.
26. Написати РВ для знаходження всіх цілих чисел у тексті, які складаються від 1 до 3 цифр (0-9).

27. Написати РВ який би заміняв у тексті два і більше пробіли одним та вивести результат.
28. Написати РВ для знаходження перших слів в кожному рядку тексту.
29. Написати РВ для знаходження перших двох літер кожного слова в тексті.
30. Написати РВ для конвертації дати в тексті з формату YYYY-MM-DD у формат DD-MM-YYYY.
31. Написати РВ який би знаходив у тексті цілі десяткові числа і повертає та виводив їх. Ціле десяткове число складається з 1 або більше цифр, може починатися з мінуса (-), не містить інших символів крім цифр. Приклад цілих чисел: "5000", "-999", "00". Приклад не цілих чисел: "", " 5", "+999", "0.0".
32. Написати РВ, який би знаходив у заданому повному шляху до файлу його розширення і повертає та виводив його. Наприклад: 'archive.zip' → 'zip', 'image.jpeg' → 'jpeg', 'archive.tar.gz' → 'gz'.
33. Написати РВ, який би формував для заданої фрази абревіатуру, повертає та виводив її. Наприклад: 'Graphics Interchange Format' → 'GIF', 'frequently asked questions' → 'FAQ', 'Joint Photographic Experts Group' → 'JPEG', 'content management system' → 'CMS', 'HyperText Markup Language' → 'HTML'.
34. Написати РВ який би знаходив у тексті дати у форматі YYYY-MM-DD і повертає та виводив їх. Приклад валідних дат: "2016-01-02", "1900-01-01". Приклад невалідних дат: "2016-02-99", "20-02-20", "1980-30-05".
35. Написати РВ який би знаходив у тексті дати у форматі YYYY-MM-DD і повертає тільки рік YYYY. Приклад валідних дат: "2016-01-02", "1900-01-01". Приклад невалідних дат: "2016-02-99", "20-02-20", "1980-30-05".
36. Написати РВ який би знаходив у тексті слова, що містять дві пари однакових літер з одною літерою між ними і повертає та виводив їх. Приклад слів: freebee, assessed, voodoo.
37. Написати РВ який би всі розширення JPEG у імені файлу заміняв на "jpg" та виводив результат. Наприклад 'avatar.jpeg' → 'avatar.jpg', 'Avatar.JPG' → 'Avatar.jpg', 'AVATAR.Jpg' → 'AVATAR.jpg'.
38. Написати РВ який би видаляв у IP-адресі нулі на початку октету та вивести результат. Наприклад: "216.08.094.196" → 216.8.94.196.
39. Написати РВ який би знаходив у тексті всі коректні MAC-адреси та вивести результат. MAC-адреса містить шість пар hex-символів розділених символами : або -. Наприклад: "01:23:45:67:89:ab" або "ff-23-45-67-89-ab".
40. Написати РВ який би знаходив у тексті всі номера стільникових телефонів у форматі xxx-xxx-xxxx та вивести результат. Наприклад: "097-567-8901".
41. Написати РВ який би знаходив у тексті всі номера стільникових телефонів у форматі xxx-xxx-xxxx та вивести код оператора. Наприклад: "097-567-8901" - "097".
42. Написати РВ який би знаходив у тексті домени в адресах електронної пошти, повертає та виводив їх. Приклад рядок 'abc.test@gmail.com, xyz@test.in, test.first@analyticsvidhya.com, first.test@rest.biz' → [gmail.com', 'test.in', 'analyticsvidhya.com', 'test.biz']
43. Написати РВ який би знаходив у тексті домени верхнього рівня в адресах електронної пошти, повертає та виводив їх. Приклад рядок 'abc.test@gmail.com, xyz@test.in, test.first@analyticsvidhya.com, first.test@rest.biz' → ['com', 'in', 'com', 'biz'].
44. Написати РВ, який би видаляв у тексті два однакових слова підряд (слово, один або декілька пробілів і знову те саме слово). Приклад: 'Два однакових слова підряд – досить поширенна помилка' → 'Два однакових слова підряд – досить поширенна помилка'.
45. Написати РВ який би перевіряв чи відповідає введений пароль вказаним вимогам політики безпеки, а саме: пароль від 7 до 15 символів без пробілів, обов'язково хоча б одна цифра,

обов'язково хоча б одна велика літера, обов'язково хоча б один спеціальний символ з _, ?, *, \$, @, -.

46. Написати РВ який би перевіряв чи відповідає введений пароль вказаним вимогам політики безпеки, а саме: пароль від 8 до 20 символів без пробілів, обов'язково хоча б дві цифри, обов'язково хоча б одна велика літера, обов'язково хоча б два спеціальних символи з _, ?, *, \$, @, -, !, &, #.
47. Написати РВ який би перевіряв чи відповідає введений пароль вказаним вимогам політики безпеки, а саме: пароль від 10 до 20 символів без пробілів, обов'язково хоча б три цифри, але не більше 5, обов'язково хоча б дві великі літери, обов'язково хоча б один спеціальний символ з _, ?, *, \$, @, -, !, &, #.
48. Написати РВ який би перевіряв чи відповідає введений пароль вказаним вимогам політики безпеки, а саме: пароль від 10 до 20 символів без пробілів, обов'язково хоча б 2 цифри, але не більше 4 і всі цифри різні, обов'язково хоча б дві великі літери, обов'язково хоча б один спеціальний символ з _, ?, *, \$, @, -, !, &, #, але не більше двох.
49. Написати РВ який би перевіряв чи відповідає введений пароль вказаним вимогам політики безпеки, а саме: пароль від 12 до 20 символів без пробілів, обов'язково хоча б дві цифри, але не більше 5, обов'язково хоча б дві великі літери, обов'язково хоча б один спеціальний символ з _, ?, *, \$, @, -, !, &, #, але не більше трьох.
50. Написати РВ який би перевіряв чи є рядок коректним українським реєстраційним автомобільним номером. Вважати, що номер складається з цифр (D) і літер А, В, Е, І, К, М, Н, О, Р, С, Т, Х (L) і має формат LL DDDD LL (кількість пробілів може бути довільна). Перші дві літери – номер області, допустимі значення задані у таблиці.

Регіон	Код
АР Крим	АК
місто Київ	АА
Вінницька область	АВ
Волинська область	АС
Дніпропетровська область	АЕ
Донецька область	АН
Київська область	АІ
Житомирська область	АМ
Закарпатська область	АО
Запорізька область	АР
Івано-Франківська область	АТ
Харківська область	АХ
Кіровоградська область	ВА
Луганська область	ВВ
Львівська область	ВС
Миколаївська область	ВЕ
Одеська область	ВН
Полтавська область	ВІ
Рівненська область	ВК
Сумська область	ВМ
Тернопільська область	ВО
Херсонська область	ВТ
Хмельницька область	ВХ
Черкаська область	СА

Чернігівська область	СВ
Чернівецька область	СЕ
місто Севастополь	СН

Табл. 1

Варіанти завдань

Варіант	Номера завдань
1.	11, 23, 29, 34, 38
2.	7, 16, 26, 41, 50
3.	18, 24, 27, 41, 48
4.	3, 15, 20, 31, 41
5.	1, 10, 16, 42, 49
6.	3, 6, 16, 32, 40
7.	3, 14, 15, 41, 43
8.	9, 19, 29, 34, 38
9.	5, 20, 23, 42, 50
10.	18, 19, 24, 35, 46
11.	20, 26, 29, 32, 38
12.	2, 14, 16, 43, 47
13.	3, 24, 25, 34, 46
14.	4, 9, 30, 36, 45
15.	1, 5, 21, 31, 39
16.	8, 9, 19, 37, 41
17.	4, 12, 20, 39, 40
18.	3, 20, 30, 35, 37
19.	3, 6, 9, 38, 42
20.	1, 11, 13, 39, 40
21.	15, 18, 20, 39, 44
22.	9, 20, 30, 43, 47
23.	10, 14, 27, 31, 46
24.	24, 25, 30, 34, 37
25.	9, 21, 22, 37, 43
26.	6, 7, 27, 32, 47
27.	14, 27, 28, 37, 49
28.	9, 14, 17, 33, 44
29.	2, 24, 29, 39, 42
30.	23, 24, 27, 41, 42

2. ЗМІСТ ЗВІТУ

1. Мета роботи.
2. Лістинг програми.
3. Результати роботи програм (текст і скріншоти).
4. Висновок.

3. КОНТРОЛЬНІ ЗАПИТАННЯ

1. Яким чином у РВ задається повторення?
2. Які є функції для пошуку шаблону в рядку і в чому відмінність між ними?
3. Які є функції для модифікації рядка?
4. Що задають якірні прив'язки?
5. Які є прапорці та їх призначення?

4. СПИСОК ЛІТЕРАТУРИ

1. Learn to Program with Python 3. A Step-by-Step Guide to Programming, Second Edition / Irv Kalb. – Mountain View: Apress, 2018. – 361 p.
2. Python Pocket Reference, Fifth Edition / Mark Lutz. – Sebastopol: O'Reilly Media, Inc., 2014. – 264 p.
3. The Python 3 Standard Library by Example / Doug Hellmann. – Boston: Addison-Wesley, 2018. – 5261 p.

НАВЧАЛЬНЕ ВИДАННЯ

РЕГУЛЯРНІ ВИРАЗИ

МЕТОДИЧНІ ВКАЗІВКИ

до лабораторної роботи №3

з курсу «Програмування скриптовими мовами. Частина 2»

для студентів спеціальності

«Кібербезпека»

Укладач:

Я. Р. Совин, канд. техн. наук, доцент