МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЛЬВІВСЬКА ПОЛІТЕХНІКА»

іНСТИТУТ КОМП’ютерних НАУК та ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

### *Кафедра “Системи автоматизованого проектування”*

Звіт

До лабораторної роботи №1

на тему: «Вивчення бібліотеки прикладних програм NLTK, для

опрацювання текстів природною мовою.»

Виконала:   
студентка групи ПРЛм-11  
Свіляк О.Ю.

Прийняв:  
викладач  
 Дупак Б.П.

Львів 2015

**Мета роботи:** Вивчення основ програмування на мові Python.

**Теоретичні відомості**

Python - це проста і потужна об’єктно-орієнтована мова програмування високого рівня з чудовими можливостями для обробки лінгвістичних даних.

Natural Language Toolk (NLTK) – набір Python бібліотек, які призначені для аналізу текстів природною мовою. NLTK дозволяє здійснювати символьний та статистичний аналіз текстів, створювати графічні звіти та містить детальну документацію і використовується в проектах з лінгвістики, штучного інтелекту, машинного навчання, автоматизації документообігу. Його можна застосовувати як начальний комплекс, готовий аналітичний інструмент або платформу для створення прикладних систем опрацювання текстів. NLTK вільно розповсюджується (http://www.nltk.org) і всі бажаючі можуть його встановити згідно інструкції розробників.

Python – інтерпретаційна мова, яка дозволяє зекономити час, що витрачається на компіляцію. Інтерпретатор можна використовувати інтерактивно, що дозволяє експериментувати з можливостями мови і створювати фрагменти програм або тестувати окремі функції.

Інтерпретатор – це програма яка виконує Python програми.

При запуску інтерпретатора ми бачимо інформацію про його версію, додаткову інформацію і запрошення >>> вводити оператори Python. У випадку використання Interactive DeveLopment Environment (IDLE) нам доступні додаткові зручності, зокрема у відображенні тексту програми на екрані.

*Представлення тексту.*

Текст або частини тексту в програмах на Python представляються за допомогою стрічок (string) і повинен відділятися від решти програми лапками (одинарними, подвійними або потрійними).

*Змінні. Операції присвоювання.*

Для того щоб використовувати, в подальшому, значення нам потрібно їх зберегти розмістивши в певному місці в пам’яті комп’ютера. Місце де зберігають ці значення називається змінна і їй присвоюється якесь ім’я. Значення цій змінній присвоюються операцією присвоювання. Оператор присвоювання це знак дорівнює.

*Доступ до окремих символів стрічки.*

Змінна яка називається стрічкою (має тип змінної - стрічка) є дуже важливою при опрацюванні текстів. Всі позиції в стрічці є пронумеровані починаючи з нуля. Для доступу до окремої позиції в стрічці ми записуємо цю позицію в квадратних дужках.

Така операція називається індексуванням. Позиція записана в квадратних дужках називається індексом.

У випадку коли потрібно отримати доступ до більш ніж одного символу необхідно в квадратних дужках вказати індекс першого символу і останнього розділивши їх двокрапкою. Цей процес називається slicing (зріз).

*Списки.*

Lists(списки) – тип даних для опису послідовності значень. В Python списки представляються як послідовність записана через кому і у квадратних дужках.

Значення у списках можна змінювати на відміну від стрічок де значення змінити неможливо. Вміст списків можна міняти в будь-який момент часу і тому списки підтримують набір операцій або методів.

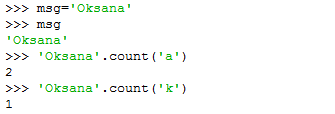
*Кортежі.*

Tuples(кортежі) – подібний до списків тип даних але подібно до стрічок кортежі також не можемо змінювати.

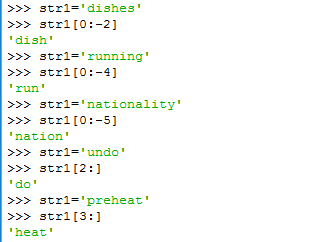
**Тексти програм на мові Python**

**Варіант 15**

1.3. Здійснити арифметичні операції зі стрічкою msg.



1.7. Використовуючи зрізи видаліть афікси у наступних словоформах: dish-es, run-ning,nation-ality, un-do, pre-heat.



1.10. Поясніть результат виконання msg[::-1].

Безымянный.png

У результаті виконання даної операції стрічку виводить у зворотньому порядку.

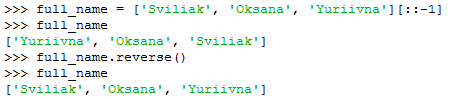
1.17. Створіть файл test.py, який містить стрічку msg . Використайте наступні оператори і поясніть отримані результати.

>>> from test import msg

>>> msg

Безымянный.png

1.13. Представити прізвище, ім’я та по батькові як список стрічок. Використовуючи метод.reverse() та зріз [::-1] змінити стрічку. Результати пояснити.

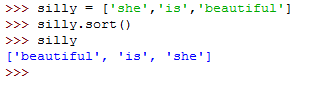


Метод reverse() міняє місцями списки стрічок без зміни у них символів з кінця на початок, а зріз [::-1] повертає списки у початкове положення.

1.22. Напишіть програму перетворення списку стрічок в одну стрічку.

Безымянный.png

1.23.Напишіть програму, яка надрукує слова із стрічки silly за абеткою.



**Висновок:** на цій лабораторній роботі я розглянула та засвоїла основи програмування на мові Python