МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

ІНСТИТУТ КОМП’ЮТЕРНИХ НАУК ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

### *Кафедра “Системи автоматизованого проектування”*



Звіт

до лабораторної роботи №12

на тему: «Вивчення бібліотеки прикладних програм NLTK для опрацювання текстів природною мовою. Автоматичний синтаксичний аналіз (частина2)»

з дисципліни “Комп’ютерна лінгвістика”

Виконала:

студентка групи ПРЛм-11

Свіляк О.Ю.

Прийняв:

викладач

Дупак Б.П.

Львів-2015

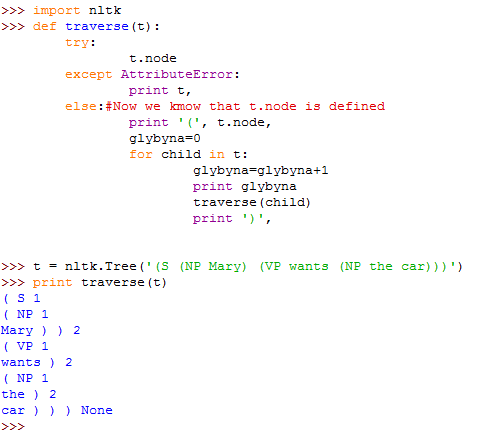
**Мета роботи:**

* Вивчення основ програмування на мові *Python*.
* Ознайомлення з автоматичним синтаксичним аналізом в NLTK.

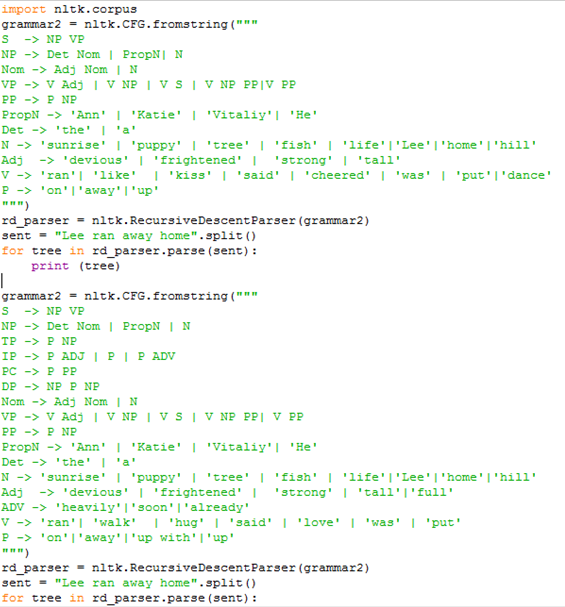
**Тексти програм на мові Python**

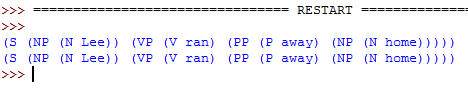
**Варіант 15**

1.Написати рекурсивну функцію для перегляду дерева, яка визначає його глибину. Дерево з одного вузла має глибину рівну нулю. (глибина піддерева це максимальна глибина його дітей плюс один)



4.Розширити граматику grammar2 з попередньої лабораторної роботи правилами які розділяють прийменники як перехідні, неперехідні та такі що вимагають PP доповнення. На основі цих правил побудуйте дерево розбору для речення Lee ran away home, використовуючи аналізатор рекурсивного спуску.



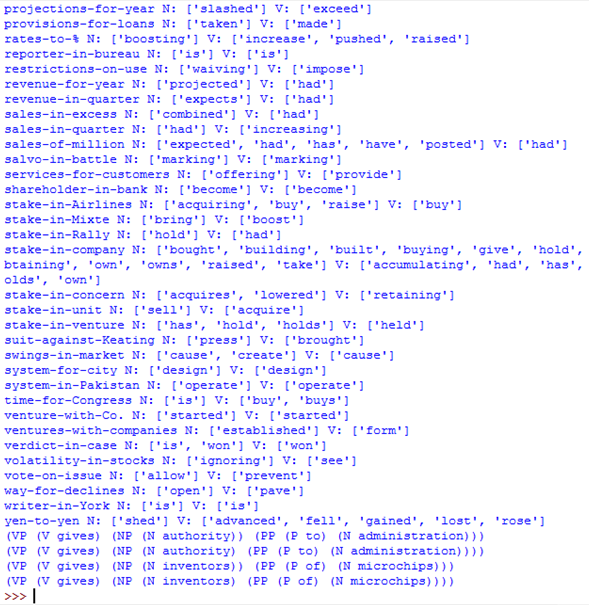


5.Вибрати декілька (2) загальних дієслова та напишіть програми для вирішення наступних задач:

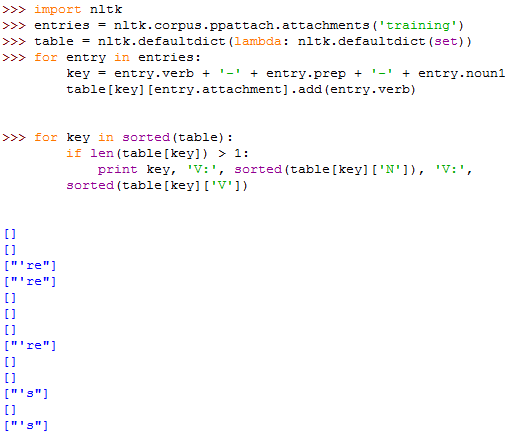
Пошук дієслів в корпусі Prepositional Phrase Attachment Corpus nltk.corpus.ppattach. Пошук всіх випадків вживання дієслова з двома різними РР в яких перший іменник, або другий іменник або прийменник залишаються незмінними

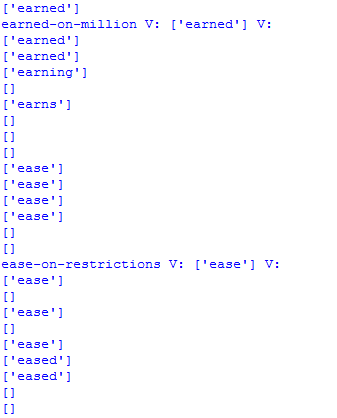
Розробити правила CFG граматики для врахування цих випадків.



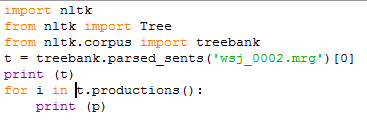


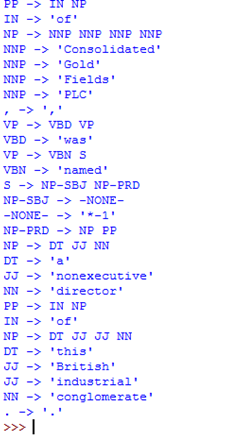
8.Здійснити аналіз корпуса Prepositional Phrase Attachment Corpus та спробувати знайти фактори, які впливають на місце приєднання PP.





12.Розробити програму обробки дерев корпуса Treebank nltk.corpus.treebank , яка вилучить всі правила з кожного з дерев за допомогою Tree.productions(). Правилами, які зустрічаються тільки один раз можна знехтувати. Правила з однаковими лівими частинами та подібними правими частинами об’єднати для отримання еквівалентного але більш компактного набору правил.





**Висновок:** на цій лабораторній роботі я ознайомилася з автоматичним синтаксичним аналізатором NLTK.