МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

ІНСТИТУТ КОМП’ЮТЕРНИХ НАУК ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

### *Кафедра “Системи автоматизованого проектування”*



Звіт

до лабораторної роботи №11

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

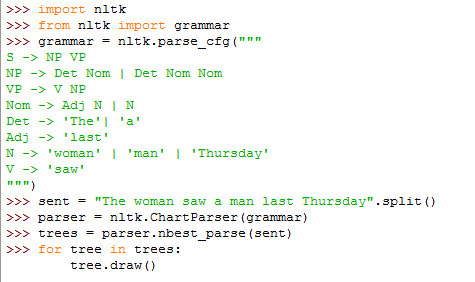
**Мета роботи:**

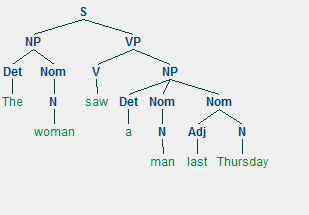
* Вивчення основ програмування на мові *Python*.
* Ознайомлення з автоматичним синтаксичним аналізом в NLTK.

**Тексти програм на мові Python**

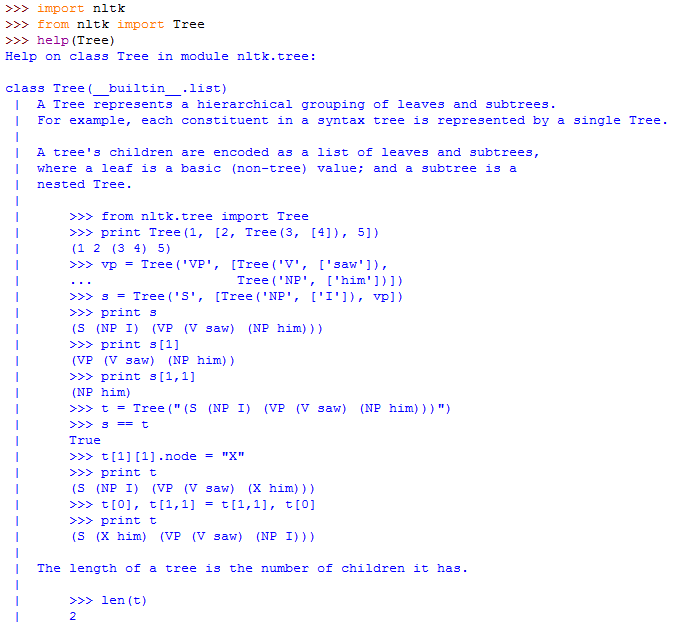
**Варіант 15**

1. Написати програму побудови дерев для речення  The woman saw a man last Thursday.



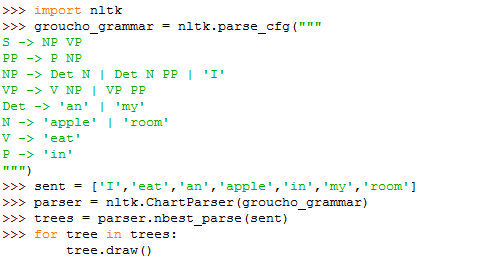


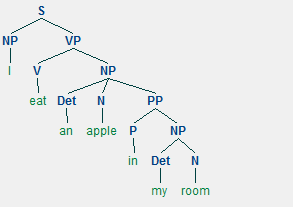
1. В класі Tree реалізовано різноманітні корисні методи. Переглянути файл допомоги  Tree з документації та описати основні з цих методів (import Tree,  help(Tree).

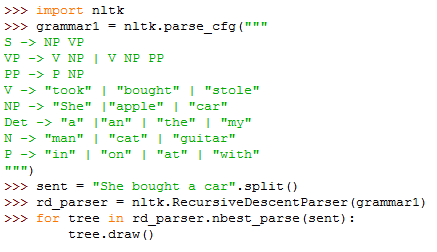


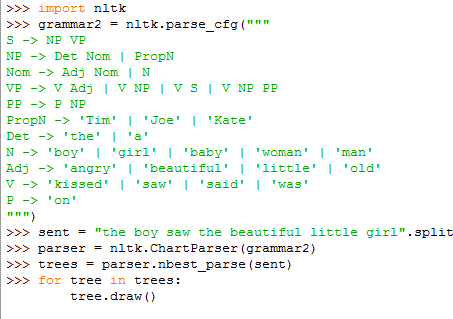
Файл допомоги пояснює принцип побудови дерев. Основним методом є Chomsky\_normal\_form, оскільки його інтерпретація зустрічається найчастіше.

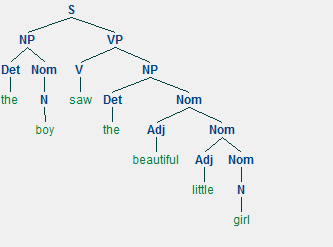
1. Перетворити всі дерева , які зустрічаються в методичних вказівка і зображені за допомогою дужок використовуючи  nltk.Tree() . Використовувати draw() для побудови графічного зображення дерева.



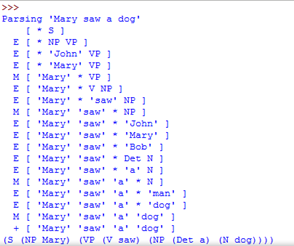


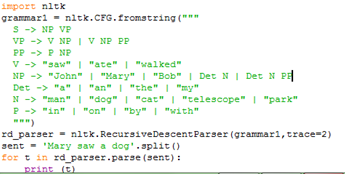






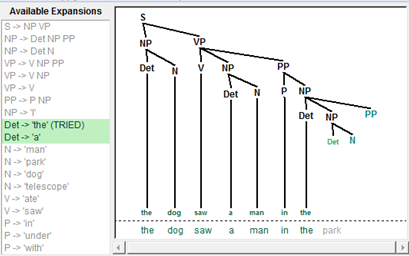
11. Здійснити модифікацію граматики в демонстраційній програмі синтаксичного аналізу згідно алгоритму рекурсивного спуску . Змінити перше правило  NP -> Det N PP, на NP -> NP PP. Використовуючи  Step кнопку спробувати побудувати дерево розбору. Що відбувається?



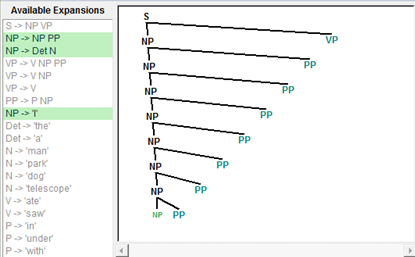


Програма виводить етапи розбору речення. Якщо замінити NP -> Det N PP, на NP -> NP PP то програма не закінчує розбір.

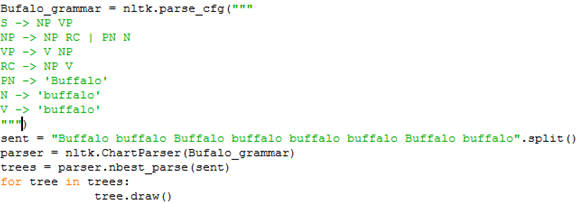
До:

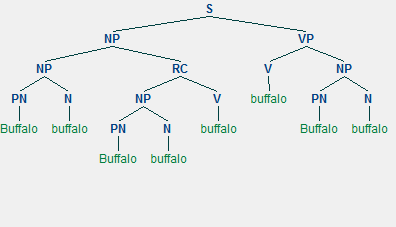


Після:

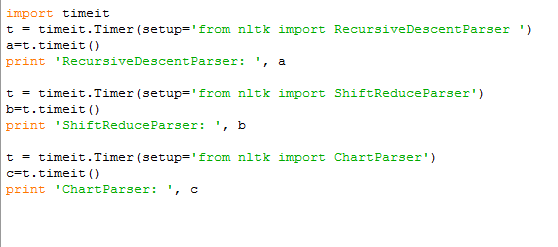


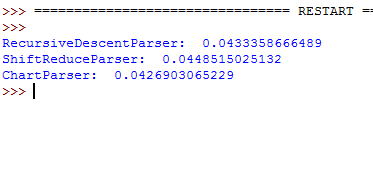
1. Здійснити аналіз послідовності слів: Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo. Оскільки, згідно з http://en.wikipedia.org/wiki/Buffalo\_buffalo\_Buffalo\_buffalo\_buffalo\_buffalo\_Buffalo\_buffalo це граматично правильне речення, напишіть контексно-вільну граматику на основі дерева наведеного на цій сторінці з Інтернету. Здійсніть нормалізацію слів (lowercase), для моделювання ситуації коли слухач сприймає це речення на слух. Скільки дерев розбору може мати це дерево в такому випадку?



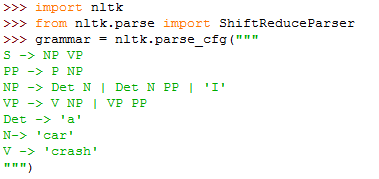


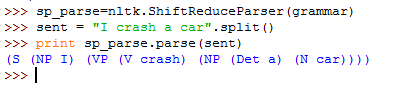
1. Написати програму порівняння швидкодії всіх аналізаторів, які згадувалися в методичних. Використовувати timeit функцію для визначення часу синтаксичного аналізу одного і того самого речення різними аналізаторами.





1. Використовуючи ручку і папір, здійснити всі кроки роботи аналізаторів рекурсивного спуску та переміщення згортання для довільного речення використовуючи просту контексно-вільну граматику.





**Висновок:** на цій лабораторній роботі я ознайомилася з автоматичним синтаксичним аналізатором NLTK, здійснювала модифікацію граматики в демонстраційній програмі синтаксичного аналізу згідно алгоритму рекурсивного спуску.