Lab11. Building Parse Trees

EXERCISE-1

In [1]:
```
import nltk,re,pprint
from nltk.tree import Tree
from nltk.tokenize import word_tokenize
from nltk.tag import pos_tag
from nltk.chunk import ne_chunk
import numpy as npt
```

In [2]:
```
np= nltk.Tree.fromstring('(NP (N Marge))')
np.pretty_print()
```

```
  NP
  |
  N
  |
Marge
```

In [3]:
```
vp= nltk.Tree.fromstring('(VP (V make) (NP (DET a) (N ham) (N sandwich)))')
vp.pretty_print()
```

```
          VP
  ________|___
 |            NP
 |     _______|_____
 V   DET      N     N
 |    |       |     |
make  a      ham sandwich
```

In [4]:
```
aux= nltk.Tree.fromstring('(AUX will)')
aux.pretty_print()
```

```
AUX
 |
will
```

Excercise-2

Exercise 2 Create a parse tree for the phrase old men and women. Is it well formed sentence or ambiguous sentence?. Steps:

1. Define the grammar (use fromstring() method)

2. Create sentence (as a list of words)

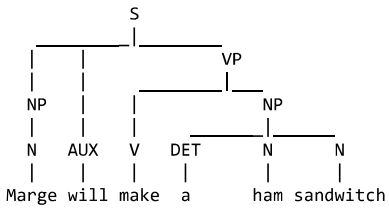
3. Create chart parser

4. Parse and print tree(s)

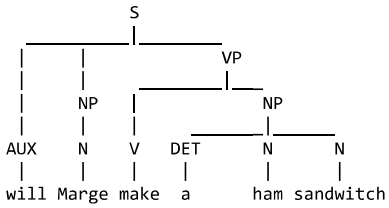
In [14]:
```
tree = nltk.Tree.fromstring('(NP (Adj old) (NP (N men) (Conj and) (N women)))')
tree.pretty_print()
```

```
         NP
  _______|___
 |           NP
 |     _______|_____
Adj  N       Conj   N
 |   |        |     |
old men      and  women
```

In [22]:
```
s1= nltk.Tree.fromstring('(S (NP (N Marge)) (AUX will) (VP (V make) (NP (DET a) (N ham) (N sandwitch))))')
s1.pretty_print()
```

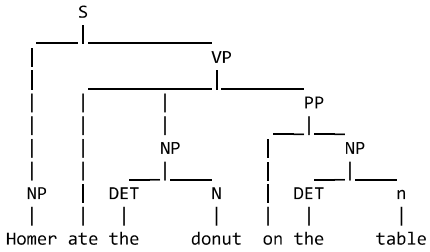


In [23]:
```
s2= nltk.Tree.fromstring('(S (AUX will) (NP (N Marge)) (VP (V make) (NP (DET a) (N ham) (N sandwitch))))')
s2.pretty_print()
```



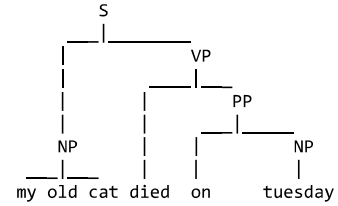
Exercise-4

In [25]:
```
s3= nltk.Tree.fromstring('(S (NP Homer) (VP ate (NP (DET the) (N donut)) (PP on (NP (DET the) (n table)))))')
s3.pretty_print()
```

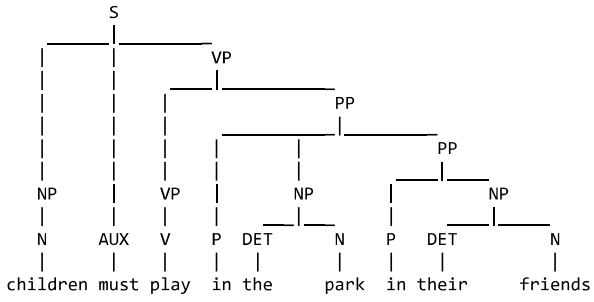


Exercise-5

In [26]:
```
s4= nltk.Tree.fromstring('(S (NP my old cat) (VP died (PP on (NP tuesday))))')
s4.pretty_print()
```



In [27]:
```
(S (NP (N children)) (AUX must) (VP (VP (V play)) (PP (P in) (NP (DET the) (N park)) (PP (P in) (NP (DET their) (N friends))))))')
```



In [28]:
```
print(vp)
#this is from exercise 1
```

```
(VP (V make) (NP (DET a) (N ham) (N sandwich)))
```

In [29]:
```
vp_rules= vp.productions() # list of all CF rules used in the tree
vp_rules
```

Out[29]:
```
[VP -> V NP,
 V -> 'make',
 NP -> DET N N,
 DET -> 'a',
 N -> 'ham',
 N -> 'sandwich']
```

In [30]:
```
vp_rules[0]
```

Out[30]:
```
VP -> V NP
```

In [31]:
```
vp_rules[1]
```

Out[31]:
```
V -> 'make'
```

In [32]:
```
vp_rules[0].is_lexical()
```

Out[32]:
```
False
```

In [33]:
```
vp_rules[1].is_lexical()
```

Out[33]:
```
True
```

Explore the CF rules of s5

In [34]:
```
print(s5)
```

```
(S
  (NP (N children))
  (AUX must)
  (VP
    (VP (V play))
    (PP
      (P in)
      (NP (DET the) (N park))
      (PP (P in) (NP (DET their) (N friends))))))
```

In [35]:
```
s5_rules= s5.productions()
s5_rules
```

Out[35]:
```
[S -> NP AUX VP,
 NP -> N,
 N -> 'children',
 AUX -> 'must',
 VP -> VP PP,
 VP -> V,
 V -> 'play',
 PP -> P NP PP,
 P -> 'in',
 NP -> DET N,
 DET -> 'the',
 N -> 'park',
 PP -> P NP,
 P -> 'in',
 NP -> DET N,
 DET -> 'their',
 N -> 'friends']
```

a. How many CF rules are used in s5?

In [36]:
```
print("How many CF values are used in s5 ",len(s5_rules))
```

```
How many CF values are used in s5  17
```

b. How many unique CF rules are used in s5?

In [37]:
```
x= npt.array(s5_rules)
print("How many unique CF rules are used in s5 ",len(npt.unique(x)))
```

```
How many unique CF rules are used in s5  15
```

c. How many of them are lexical?

In [39]:
```
n= 0
for x in s5_rules:
    if x.is_lexical():
        n= n+1
print("How many of them are lexical? ",n)
```

```
How many of them are lexical?  9
```

In [ ]: