In [1]:
```
import nltk
from nltk import CFG
tokenizer = nltk.tokenize.TreebankWordTokenizer()
```

In [2]:
```
import nltk


cfg = nltk.CFG.fromstring("""
S -> NP VP
VP -> V NP | VP PP
PP -> P NP | P S
NP -> 'I' | Det N | Det N PP
V -> 'saw'
N -> 'man' | 'hill' | 'telescope'
Det -> 'a' | 'A'
P -> 'on' | 'with'
""")

input_sentence = "I saw a man with a telescope"

input_sentence_token=tokenizer.tokenize(input_sentence)
input_sentence_token
```

Out[2]:
```
['I', 'saw', 'a', 'man', 'with', 'a', 'telescope']
```

In [3]:
```
chart_parser = nltk.ChartParser(cfg)
```

In [4]:
```
for tree in chart_parser.parse(input_sentence_token):
    print(tree)
```

```
(S
  (NP I)
  (VP
    (VP (V saw) (NP (Det a) (N man)))
    (PP (P with) (NP (Det a) (N telescope)))))
(S
  (NP I)
  (VP
    (V saw)
    (NP (Det a) (N man) (PP (P with) (NP (Det a) (N telescope))))))
```

In [5]:
```
#The grammmer is ambiguous so that there can be two parse trees we can have two part t
#the grammer is ambiguous
```

In [6]:
```
cfg1 = nltk.CFG.fromstring("""
S -> NP VP
VP -> V NP
PP -> P NP | P S
NP -> 'I' | Det N | Det N PP
V -> 'saw'
N -> 'man' | 'hill' | 'telescope'
Det -> 'a' | 'A'
P -> 'on' | 'with'
""")

input_sentence1 = "I saw a man with a telescope"
```

```
input_sentence_token1=tokenizer.tokenize(input_sentence1)
input_sentence_token1
chart_parser = nltk.ChartParser(cfg1)
for tree in chart_parser.parse(input_sentence_token):
    print(tree)
```

```
(S
  (NP I)
  (VP
    (V saw)
    (NP (Det a) (N man) (PP (P with) (NP (Det a) (N telescope))))))
```

In [7]:
```
#here the ambiguity is removed and there's only one pharse tree
```

In [16]:
```
cfg2=nltk.CFG.fromstring("""
S -> NP VP
NP -> N | D N | A N | A N P | D N P | D A N P | N P | N N
NP -> Pn | Pr
VP -> V | NP V | Advp V | Advp Advp V | Advp Adv Advp V
Advp -> Adv | Deg Adv | N N N | N N | N | Pg N N N
N -> "බල්ලා"|"මිනිසා"|"මේසය"|"වන්න"|"බත්"|"ඊයේ"|"පෙරේදා"|"සඳුදා"|"බදාදා"|"ඉරිදා"|
D -> "ඒ"|"මේ"|"අර"|"ඔය"|"සමහර"
A -> "උස"|"හොඳ"|"නරක"|"කොට" "ලස්සන"
P -> "උඩ"|"යට"|"මත"
Pn -> "අමල්"|"කමල්"|"සමන්"|"සචිනි"|"මාලා"|"නාමල්"
Pr -> "ඇය" "ඔවුන්"|"අපි"/"මම"|"උඹ"|"උන්"|"ඔබ"|"ඔබලා"
V -> "කෑවා"|"බිව්වා"|"දැක්කා"|"ගියා"|"දිව්වා"|"වා"|"නැටුවා"| "කමු"
Adv -> "ලස්සනට" | "කැනට" | "ඉක්මනට" | "වේගයෙන්" | "හෙමින්" | "පහසුවෙන්" |"පන
Deg -> "බොහොම" | "හරිම" | "නොම" | "මාර"
PA -> "කන" | "බොන" | "යන" | "එන" | "බලන" | "කියන" | "ලියන" | "නටන" | "කා
""")
```

In [28]:
```
input_sentence2 ="අපි බත් කමු"
input_sentence3 ="කමල් ගිය අවුරුද්දේ ජනවාරි මාසේ ඉස්කෝලෙදි බොහොම ලස්සනට නැටුවි
input_sentence4 = "සමන් හරිම වේගයෙන් දිව්වා"

input_sentence_token=tokenizer.tokenize(input_sentence4)
input_sentence_token
chart_parser = nltk.ChartParser(cfg2)
for tree in chart_parser.parse(input_sentence_token):
    print(tree)
```

```
(S (NP (Pn සමන්)) (VP (Advp (Deg හරිම) (Adv වේගයෙන්)) (V දිව්වා)))
```

In [ ]:
```
#sentence 2 and 3 doesn't have a parse tree
```