Question 1

The *init_wfst*() and *complete_wfst*() functions in *homework6_1.py* are modified from the given functions, so that the chart cells contain a set of non-terminals. The chart parser table when code isn't modified is:

```
WFST 1    2    3    4    5    6    7
0    NP   .    .    S    .    .    S
1    .    V    .    VP   .    .    VP
2    .    .   Det   NP   .    .    .
3    .    .    .    N    .    .    .
4    .    .    .    .    P    .    PP
5    .    .    .    .    .   Det   NP
6    .    .    .    .    .    .    N
```

The chart parser table when the code is modified is:

```
WFST 1    2    3    4    5    6    7
0   [NP]  .    .   [S]   .    .   [S]
1    .   [V]   .   [VP]  .    .   [VP]
2    .    .  [Det] [NP]  .    .    .
3    .    .    .   [N]   .    .    .
4    .    .    .    .   [P]   .   [PP]
5    .    .    .    .    .  [Det] [NP]
6    .    .    .    .    .    .    [N]
```

However, it is to be noted that for creating the chart-parser table, the grammar ought to be in CNF, which is not the case with the groucho grammar, as it has a production rule $NP \rightarrow Det\ N\ PP$, which has three non-terminals on the right hand side. This can be converted into CNF by adding a new rule $X \rightarrow N\ PP$, and replacing the original one by$NP \rightarrow Det\ X$. This modified grammar is saved as *goucho_grammar2* in the script. Upon running the *init_wfst*() and *complete_wfst*() functions on this modified grammar, we get the chart parser table in the following way:

```
WFST 1    2    3    4    5    6    7
0   [NP]  .    .   [S]   .    .   [S]
1    .   [V]   .   [VP]  .    .   [VP]
2    .    .  [Det] [NP]  .    .   [NP]
3    .    .    .   [N]   .    .   [X]
4    .    .    .    .   [P]   .   [PP]
5    .    .    .    .    .  [Det] [NP]
6    .    .    .    .    .    .   [N]
```

It can be observed that, now, the span for (2, 7) and (3, 7) also result in non-terminals, which wasn't captured earlier. Considering the sentence "I shot an elephant in my pajamas" which these charts represent, it can be observed that the original chart parser table wasn't

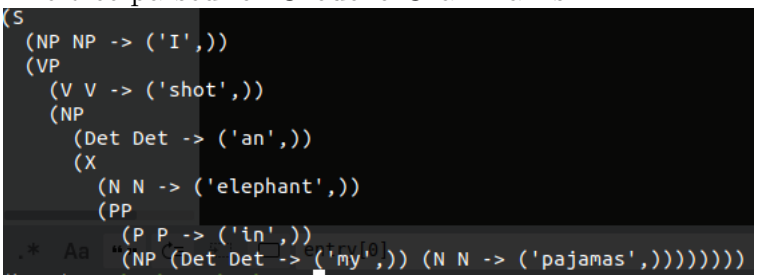
able to recognize the rule $NP \rightarrow Det\ N\ PP$, as can be seen from the fact that the *rhs* tuple for this production rule would be (Det, N, PP), while only tuples of length 2 were processed in the algorithm. Thus, the string 'an elephant in the pajamas' can not be broken down into two grammatical units by the original recognizer, which was done by converting the given grammar into CNF.

Question 2

The *update_init_chart*() and *update_complete_chart*() functions in *homework6_2.py* are used to modify the chart cells such that they contain tuples of the form $(Category, Rule, MidPoint)$. The chart parser table upon making the changes for the Groucho Grammar looks thus:

```
WFST 1    2    3    4    5    6    7
0    [(NP, "NP -> ('I',)", 0)] .    .    [(S, 'S -> NP VP', 3)] .
                                          .    [(S, 'S -> NP VP', 6)]
1    .    [(V, "V -> ('shot',)", 0)] .    [(VP, 'VP -> V NP', 3)] .
                                          .    [(VP, 'VP -> V NP', 6), (VP, 'VP -> VP PP', 6)]
2    .    .    [(Det, "Det -> ('an',)", 0)] [(NP, 'NP -> Det N', 3)] .
                                          .    [(NP, 'NP -> Det X', 6)]
3    .    .    .    [(N, "N -> ('elephant',)", 0)] .    .    [(X, 'X -> N PP', 6)]
4    .    .    .    .    [(P, "P -> ('in',)", 0)] .    [(PP, 'PP -> P NP', 6)]
5    .    .    .    .    .    [(Det, "Det -> ('my',)", 0)] [(NP, 'NP -> Det N', 6)]
6    .    .    .    .    .    .    [(N, "N -> ('pajamas',)", 0)]
```

The tree parsed for Groucho Grammar is:



For the L1 Grammar from Chapter 12 in Jufarsy and Martin SLP3, the following sentence was chosen "on the flight through Houston". The grammar for the same can be found in *question6_2b.py*