

## Homework 6: chart parsing

**Q1** Modify the functions `init_wfst()` and `complete_wfst()` so that chart cells contain a *set* of nonterminals (exercise 26 of NLTK book chapter 8). Turn in diagnostic output from your modified code along with a writeup that cogently explains how that chart encodes both analyses of "I shot an elephant in my pajamas" on the groucho grammar.

**Q2 (mandatory for grads, extra credit for ugrads)** Upgrade your system from a recognizer to a parser by changing the elements in a chart entry so that they are triples of the form

( Category, Rule, MidPoint )

Write a recursive function that returns a collection of trees, given a chart. Here is an idea to get you started:

```
def treesOfChart(chart, category, span):
    consider all chart entries in span (i,j) having the desired "category"
    if entry uses a terminal rule
        then return a list consisting of just the height-1 tree
        that is rooted in "category" and whose single leaf is the appropriate word
    if entry uses a binary rule,
        then return a list [ nltk.tree.Tree(category, [t1,t2]) |
                            for t1 in treesOfChart(chart, rule RHS1, (i, MidPoint))
                            for t2 in treesOfChart(chart, rule RHS2, (MidPoint, j)) ]
    otherwise, give up with an error message
```

Note that you only have to consider grammars in Chomsky Normal Form. Demonstrate that your parser works correctly on the L1 grammar from chapter 12 of Jurafsky and Martin SLP3. Make it into a generator if you want.