Homework 5
Due Tuesday, 03/11/08

Please turn in a print-out of your homework solutions at the beginning of class. If a program is required for a problem then you should provide sample input and output. (If the input is data from an on-line source then you may just indicate where the file can be found.) Programming solutions to problems should be placed in separate files whose names indicate the problem number (e.g. prob1.py). These separate files should be put into a folder whose name includes the assignment number and your name (e.g., Hmwk1-yourname). This folder should be dragged into the class dropoff folder at

`/common/cs/cs181/dropbox`

1. Please do problem 1 in Exercise 11.3.4 on page 275 of Bird et al.

2. In this exercise i want you to annotate a grammar in order to make sure that only correct sentences can be generated. You will start with the grammar in file `decGrammar.fcfg`, which contains a grammar to generate simple declarative sentences. It is a simple extension of that shown on page 265 of Bird et al.

   Part a below is a stand-alone exercise. Both parts a and b should build on the original grammar in `decGrammar.fcfg`. However, part c should build on b, and part d on c (though a detailed grammar is not required for d)

   (a) Please expand the grammar to handle the pronouns 'I', 'we', 'you', 'he', 'she', 'they', as well as 'me', 'us', 'him', 'her', and 'them'. this will involve adding new lexical rules (see the starter grammar) as well as new features to handle the person and whether it appears in the subject or object position. (See problem 16.2 on page 42 of Jurafskey & Martin for examples.)

   Test your annotated grammar with commands like the following (after you have imported nltk):

   ```
   >>> cp = load_earley('file:impGrammar.fcfg',trace=2)
   >>> tokens = 'I like the dog'.split()
   >>> trees = cp.nbest_parse(tokens)
   ```

   Provide test output that shows that your grammar accepts and rejects enough of the appropriate sentences that I can have confidence that it works correctly. (You can use trace=0 on the output for me once you are convinced your grammar works.)

   (b) Please expand the grammar to handle auxiliary verbs like 'is', 'are', 'do', and 'does'. Your new grammar should recognize sentences like "Jody does like the dog." and "The children do like the dog".

   This will involve adding productions like:

   ```
   S -> NP Aux VP
   Aux -> 'do' | 'does'
   ```

   although you will also have to add appropriate annotation of features. See the provided grammar for the form for writing the annotated rules. Write your grammar with features in the same fashion. You may add new features to the grammar if necessary.

   Make sure your new grammar only accepts grammatical sentences. In particular, you must worry about getting the right forms for the verbs in each case. For example, the following are clearly not right.

- Jody do likes the bike.
- The girls does ride the bike.

Again provide sufficient test data to convince me that your grammar works.

(c) Please expand the grammar to handle questions and commands. That is, add the rules:

```
S -> Aux NP VP
S -> VP
```

Again provide sufficient test data to convince me that your program works.

(d) Handling "wh-questions" can be challenging because of the distance between words like 'who', 'what', 'where', etc. and the lexical items that they must agree with.

Discuss how you could force agreement for sentences like:

```
What does Jody like?
Who likes the cars?
```

You need not provide detailed grammar rules.