Computational Linguistics I — *CMSC723, Fall 2010*

HW03: HMMs and CFGs

Hand in at: `http://www.cs.utah.edu/~hal/handin.pl?course=cmsc723`. Remember that only PDF submissions are accepted. We encourage using LaTeX to produce your writeups. See `hw00.tex` for an example of how to do so. You can make a `.pdf` out of the `.tex` by running "`pdflatex hw00.tex`".

1 Alice and the Crazy Coke Machine

Alice is up late studying for her algorithms final exam and needs to stay hydrated (and caffienated!). Unfortunately, Clarence was the one who bought the soda machine in the lab. He went for the cheapest model. All you can do with this soda machine is put money in and hope it gives you the type of drink you want. It carries three types of soda: Coke (C), Diet Coke (D) and Sprite (S).

Alice has been monitoring the soda machine for a while and has figured out that it behaves as an HMM. It has two internal states (call them A and B). When asked for a soda from state A, it gives a coke with probability 0.4, a diet coke with probability 0.35 and a sprite with probability 0.25. On the other hand, when asked for a soda in state B, it gives a diet coke with probability 0.5, a sprite with probability 0.3 and a coke with probability 0.2. Furthermore, it transitions from state A to A with probability 0.7 and from state B to B with probability 0.6.

The machine formally works as follows. It is in some state s. Someone puts in money and it dispenses a soda according to the probability rules set out above, being in state s. It then (randomly) transitions to a new state s' according to the transition probabilities above.

1. Draw a state space lattice for this soda machine (as in Figure 5.18 on p148 of J+M) for three time steps.
2. Suppose that Alice doesn't know what state the machine is in currently (specifically, she believes it's equally likely to be in either state), but puts money in and gets a Sprite out. What is the probability distribution over states that it was in when Alice put her money in? What is the probability distribution over states that it is in now?
3. Suppose Alice comes back the next day (so again she doesn't know what state the machine is in) and really wants a diet coke. Unfortunately, the machine isn't being particularly nice to her and it produces the following series of emissions upon taking money from Alice: C S S D. What is the most likely sequence of states the machine went through in this process?

2 Context Free Grammars

Consider the following context free grammar:

```
% grammar
S    -> NP VP         % declarative sentence
S    -> VP            % command
NP   -> Det Noun      % simple NP
NP   -> Noun          % indeterminate NP
NP   -> Noun Noun     % compound noun
VP   -> Verb NP       % simple VP
VP   -> VP PP         % VP with an adjunct
```

```
PP   -> Prep NP       % simple PP

% lexicon
Det  -> an
Noun -> flies
Noun -> arrow
Verb -> time
Verb -> flies
Verb -> like
Prep -> like
```

1. Given this grammar, find *as many parses as possible* for the sentence “time flies like an arrow.” Draw them as trees. (Hint: if you’re using LaTeX, you can use the qtree package `http://www.ling.upenn.edu/advice/latex/qtree/` to create nice-looking trees without much effort!)
2. Does this CFG describe a regular language? If so, why? If not, why not?
3. Convert this grammar to Chomsky normal form.