ECS 189G HW 3: The PCFG Model

Task 1

In this part, the first command to be run is `perl ./cfgparse.pl grammar2 lexicon < examples.sen`. I read the output and see that each line's third entry is 0.5, which is the probability of the parse T that is given according to the given sentence S. Furthermore, I also inspected both the `grammar2` and the `lexicon` files, and saw that each entry in grammar2 and lexicon has a weight of 1. This means that each rule is given a uniform weight. Therefore grammar2 is implementing a **uniform language model**.

Task 2

The mostly noticeable difference is that the first command, `perl ./cfgparse.pl grammar1 lexicon < examples.sen`, produced a myriad of failures; the second command, `perl ./cfgparse.pl grammar1 grammar2 lexicon < examples.sen`, still produce expected results. What I noticed from these two grammar files is that the grammar1 file does not have any rules that handle the Misc non-terminals, while grammar2 does; hence any phrases with non-terminal miscs that appear in the examples will fail to be processed by grammar 1.

However, when grammar1 and grammar2 are merged, I noticed that the parser chose grammar1 for those cases that grammar1 does not fail. This implies that grammar1 is a better model than grammar2.

Task 3

I chose N=100 and ran the following commands:

```
perl ./cfggen.pl --text 100 grammar1 lexicon > g1gen
perl ./cfggen.pl --text 100 grammar2 lexicon > g2gen
perl ./cfggen.pl --text 100 grammar1 grammar2 lexicon > g12gen
```

I found strikingly different patterns in the sentences created by grammar1 and grammar2. Overall, grammar1 produces significantly better sentences than grammar2; moreover, grammar1 combined with grammar2 also produces better results than grammar1 alone. This result follows from the difference in the weight of rules "ROOT -> S1 / S2" in each grammar: in grammar1 "ROOT -> S1" has a weight of 99, with "S1 -> NP VP" having a weight of 1; in grammar2 "ROOT -> S2" only has a weight of 1, followed by other rules all with weights of 1. This leads to grammar1 generating mostly "NP VP" sentences, which correspond to the basic structure of English sentences; on the other hand, grammar2 generates sentences that are much more varied, and most of them are not grammatically correct in English.

Considering the reasoning above plus the answers for the previous two tasks, I believe that it is not very reasonable to let each grammar rule have uniform weight. This is because whether in the examples provided or in our real lives, some grammar rules are more commonly occuring than others. Grammar1 does better job than grammar2 mostly due to assigning higher weights to more commonly seen grammar rules.

Grammar1 and Grammar2 combined is performing even better, as it works similarly to the backoff model I learned in the spelling correction homework (homework 1). If there are any occurences of miscs, the work is handled by grammar2. But, since the weights given by grammar1 are more realistic, it is a better language model.

Task 4

For this task I started from what is given grammar1 and grammar2, and tried to assign different weights. For the grammatical rules that are more likely to happen, I assigned stronger weights to those rules that happen more in our real lives. I manually adjusted the relative weights between each rules, and attempted to run cfggen.pl after each modification.

Task 5

I have got a total of 19 / 20 grammatically correct sentences, although many of them do not really make sense from a human's perspective.