Midterm

0) Language Models

Assume you are given the following corpus of text:

<s> I love NLP <e>

<s> You love programming <e>

a) Provide all the bigram probabilities for this corpus

$P(\text{I} \mid \langle s \rangle) = \frac{1}{2}$ $\quad$ $P(\text{You} \mid \langle s \rangle) = \frac{1}{2}$

$P(\text{love} \mid \text{I}) = 1$ $\quad$ $P(\text{love} \mid \text{You}) = 1$

$P(\text{NLP} \mid \text{love}) = \frac{1}{2}$ $\quad$ $P(\text{programming} \mid \text{love}) = \frac{1}{2}$

$P(\langle e \rangle \mid \text{NLP}) = 1$ $\quad$ $P(\langle e \rangle \mid \text{programming}) = 1$

b) Calculate the probability for the sentence

<s> You love NLP <e>

$\left(\frac{1}{2}\right)\ (1)\ \left(\frac{1}{2}\right)\ (1) = \frac{1}{4}$

c) What is the perplexity of the corpus?

$|V| = 8$

$$ppl = 2^{-\frac{1}{|w|}\sum_{i=1}^{N} \log_2 P(w_i)} = 2^{-\frac{1}{8}(4\log_2(\frac{1}{2}) + 4\log_2(1))}$$

$$= 2^{-\frac{1}{8}(4(-1) + 4(0))} = 2^{\frac{1}{2}} = 1.41$$

d) What is the probability of the sentence <s> I like NLP <e>? What is the probability if we adjust all the bigrams with add-one smoothing?

$P(\text{like} \mid \text{I}) = 0$, so $P(\text{sentence}) = 0$

$P(\text{I} \mid \langle s \rangle) = \frac{1+1}{2+8} = \frac{2}{10}$ $\quad$ $P(\text{sentence}) = \frac{8}{8100}$

$P(\text{love} \mid \text{I}) = \frac{1+1}{1+8} = \frac{2}{9}$

$P(\text{NLP} \mid \text{love}) = \frac{1+1}{2+8} = \frac{2}{10}$

$P(\langle e \rangle \mid \text{NLP}) = \frac{1+1}{1+8} = \frac{2}{9}$

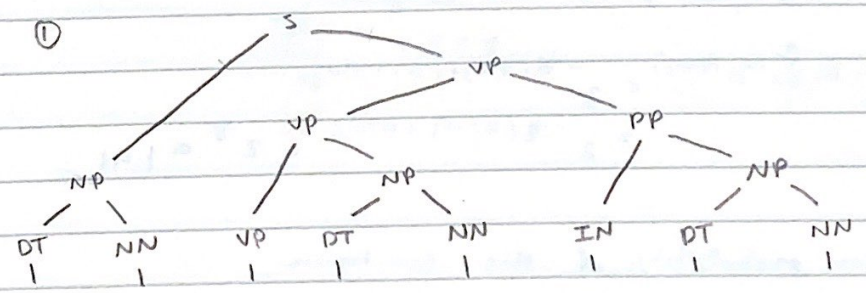
2) A Probabilistic Context-Free Grammar (PCFG)

Consider the following PCFG, where the top half of the table is the grammar, and below is the lexicon

Rule	P(Rule)
S → NP VP	1.0
NP → DT NN	0.9
NP → NP PP	0.1
VP → VP PP	0.2
VP → VP NP	0.5
PP → IN NP	1.0
IN → with	1.0
DT → the	1.0
VP → saw	0.3
NN → woman	0.4
NN → professor	0.4
NN → telescope	0.2

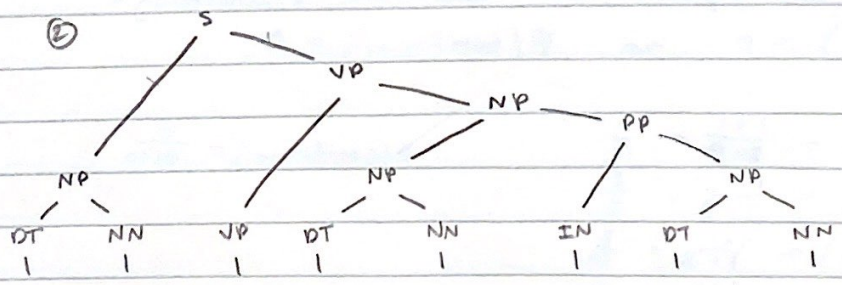
a) For the sentence "the woman saw the professor with the telescope", how many valid parses are there? Draw the parse trees.

①



the woman saw the professor with the telescope

②



the woman saw the professor with the telescope

b) What are the probabilities? If more than one, which is largest and what is the meaning or interpretation of that parse tree?

① $(1.0)(0.9)(0.5)(0.2)(0.9)(1.0)(0.9) = 0.0729$

$(0.0729)(1.0)(0.4)(0.3)(1.0)(0.4)(1.0)(1.0)(0.2) = 0.00070$

② $(1.0)(0.9)(0.5)(0.1)(0.9)(1.0)(0.9) = 0.03645$

$(0.03645)(1.0)(0.4)(0.3)(1.0)(0.4)(1.0)(1.0)(0.2) = 0.00035$

The more probable parse tree implies that the woman uses the telescope to see the professor

Bias

What is one example of bias in NLP, and how would you approach addressing it?

An example of bias in NLP is the way in which people of color are depicted in models such as the example shown of GPT-3. To fix this, I would include data which shows an even representation of all backgrounds.

I affirm that I will not give or receive any unauthorized help on this exam, and that all work will be my own

0. The written portion took me about 3-4 hours in total. I haven't completed the programming section, but I'm assuming it will take around the same time.

1. I liked that I was able to find examples from class which helped solve the problems