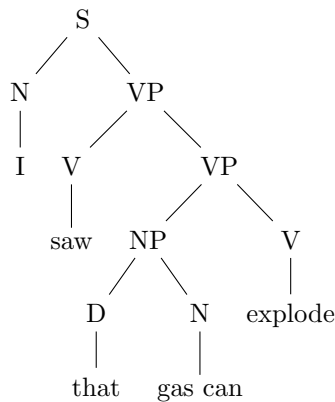Kekoa Riggin
LING 566
HW 3

Question 1

Prior to applying to CLMS, my experience had been in translation and localization. During my work in these fields, I realized that the industries depended entirely upon the technology that is standard in localization workflows. My current opinion is that there will be a $\approx 95\%$ decrease in translation work due to automation within my lifetime. My original career goal was to contribute to that decrease as much as possible, but now I see that there is even more demand for NLPers in the fields of speech recognition and virtual assistants. I consider all the doors to be open to me, but there is much work ahead.
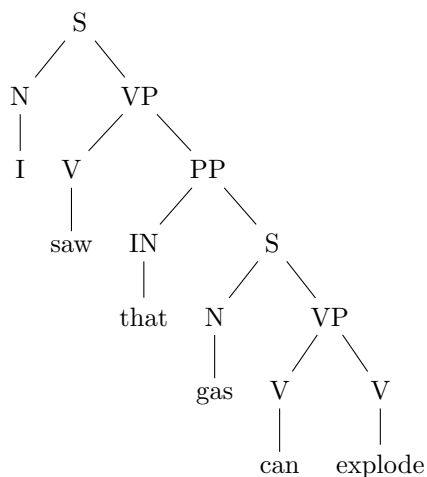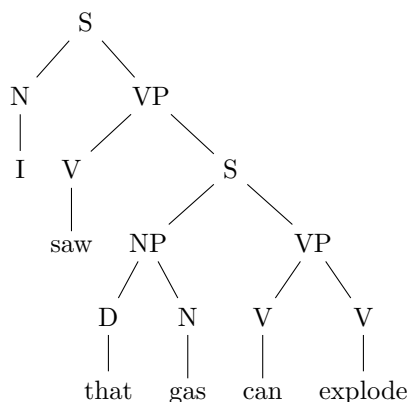
Question 2

Part a

1. What was see was a gas can, which exploded.

```
            S
          /   \
        N       VP
        |      /   \
        I    V       VP
             |      /    \
            saw   NP      V
                 /  \     |
                D    N   explode
                |    |
              that  gas can
```

2. What was seen was the ability that gas has to explode

```
            S
          /   \
        N       VP
        |      /   \
        I    V       PP
             |      /   \
            saw   IN      S
                  |     /   \
                that   N      VP
                       |     /  \
                      gas   V     V
                            |     |
                           can  explode
```

3. What was seen was a particular gas's ability to explode. Here, there is an ommitted "that": *I saw "that" that gas can explode.*

```
              S
            /    \
          N       VP
          |      /   \
          I    V       S
               |      /  \
             saw   NP      VP
                  /  \    /  \
                 D    N  V     V
                 |    |  |     |
               that  gas can  explode
```

Part b

1. (S   (N I)
        (VP saw
             (NP that
                  gas can)
             (VP explode)))

2. (S   (NP I)
        (VP saw
             (PP that
                  (S gas
                       (VP can
                            explode)))))

3. (S   (NP I)
        (VP saw
             (VP (NP that
                      gas)
                 (VP can
                      explode))))

Question 3

For all of the orderings of the letters of the alphabet, we use the function 26^6. But because we do not want words that have only consonants or only vowels, we need to subtract the number of words that can be formed with only those symbols. So we subtract 21^6 for consonants and 5^6 for vowels. Thus our equation is:
$26^6 - 21^6 - 5^6 = 223134030$

Question 4

Because repetition is permitted, but we want to find distinct tuples, we find the factorial of the total number of characters divided by the product of the factorials of the numbers of repetitions.
There are nine characters and characters 1-4 and 5-6 are idential. Therefore, our formula will be:
$\frac{9!}{4! \times 2!} = 362832$

Question 5

Part a

This is a classic n choose k situation because we cannot perform a pairwise comparison on two copies of the same document and comparing documents a, b is the same as comparing documents b, a.
Thus, the formula we will use is:
$\frac{n!}{k!(n-k)!}$
And our answer will be the sum of n choose k for each document type.
So for conference proceedings, journal articles, and workshop abstracts, we use the following formulas respctively:
$\frac{7!}{2! \times 5!} = 21$ $\frac{9!}{2! \times 7!} = 36$ $\frac{3!}{2! \times 1!} = 3$
Which makes the total number of pairwise comparisons between documents of the same type 60.

Part b

This is also a n choose k situation; however, it is more complex. In these cases, n will be the sum of the number of documents per document type. Then, we will subtract the sum of the number of pairwise comparisons of the same document type.
$\frac{(7+9+3)!}{2! \times (7+9+3-2)!} = \frac{19!}{2! \times 17!} = 171$
And taking the 60 total pairwise comparisons from documents of the same type from above:
$171 - 60 = 111$

Extra Credit

$\frac{(n+1)!}{k!((n+1)-k)!}$