Keeyon Ebrahimi
HW3

1. (a)

Start	Moo	Hello	End
Cow	$1.0 * 0.9$	$0.9 * 0.5 * 0.1$	$.045 * 0.2$
Duck	0	$0.9 * 0.3 * 0.4$	$0.108 * 0.2$

This shows that the most probably way to get 'Moo Hello End' is when you go (Start) to (Cow = Moo) to (Duck = Hello) to (End)

The probability of this happening is $0.108 * .2 = 0.0216 = 2.16\%$

Solution: 2.16%

(b) We can also emit 'Moo Hello End' by going (Start) to (Cow = Moo) to (Cow = Hello) to (End)

The probability of *just this state sequence* is $0.45 * 0.2 = .009 = 0.9\%$
To find total probability of this sentence, we add the probabilities of all the state sequences that result in our correct sentence. This means that we do $0.9\% + 2.16\% = 3.06\%$
Solution: 3.06%

2. **Correct Example:**

```
---------
Sentence:                                Dogs always count.
Annotate Dogs == <constit cat="NNS">
Annotate always == <constit cat="RB">
Annotate count == <constit cat="VB">
```

This analysis goes like this:

```
Sate start
Arc to NNS
Emit Dogs
Arc to RB
Emit always
Arc to VBN
Emit count
Arc to .
```

To find the relative probability, lets see the probability from 'RB - always' to 'VBN - count' We need to multiple $P(transition) * P(Emit)$
Probability of transition is probability from RB to VBN.

Probability of Transition from RB to VBN $\frac{2982}{37153} = 8.03\%$

Now we need to find the probability of Emitting count in this state. This probability is not there, meaning it is only one, so probability of emitting count in VB is $\frac{1}{24107}$

This shows that the majority influencing factor is the previous state being RB, and how we usually have a VBN after that.
Relative probability becomes $.0803 * 0.00004148172$, which is very low. Now lets look at the incorrect example.

Incorrect Example:

```
---------
Sentence:                                  Dogs count.
Annotate Dogs == <constit cat="JJ">
Annotate count == <constit cat="NN">
Annotate . == <constit cat=".">
```

This analysis goes like this:

```
Sate start
Arc to JJ
Emit Dogs
Arc to NN
Emit count
Arc to .
```

As we can see, we analyzed the word bark wrong. Lets see why.

So from going from JJ to (NN emit count) we get:
Probability(Arc to NN) * Probability(Emit count in NN). This is

Probability Arc to NN: $\frac{32961}{73230} = 0.4501 = 45.01\%$
Wow that is high! Now, lets check the probability to emit count in NN.
Probability to Emit 'count' in NN: $\frac{1}{159394}$
Not high at all, but when we multiply the two we get
$0.4501 * \frac{1}{159394} = 2.823 * 10^{-6}$

Now, lets see what the probability of the correct analysis is in order to see why this sentence was analyzed incorrectly.

Correct analysis would have gone like this:

```
Sate start
Arc to JJ
Emit Dogs
Arc to VB
Emit count
Arc to .
```

Probability (Arc to VB) * Probability (Emit count in VB)
Probability (Arc to VB) = $\frac{8}{73230}$
Probability (Emit count in VB) = $\frac{1}{31555}$

Relative probability = $\frac{8}{73230} * \frac{1}{31555} = 3.46 * 10^{-9}$

This is substantially lower than the incorrect analysis, which makes sense why this was analyzed incorrectly.