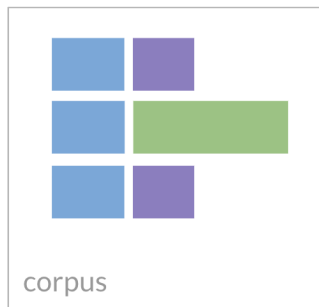


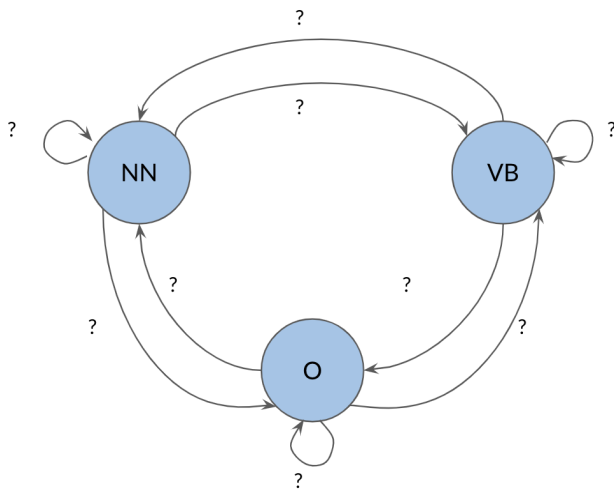
Calculating Probabilities

Here is a visual representation on how to calculate the probabilities:



transition probability: $\text{blue} + \text{purple} = 2/3$

The number of times that blue is followed by purple is 2 out of 3. We will use the same logic to populate our transition and emission matrices. In the transition matrix we will count the number of times tag $t_{(i-1)}$, $t_{(i)}$ show up near each other and divide by the total number of times $t_{(i-1)}$ shows up (which is the same as the number of times it shows up followed by anything else).



1. Count occurrences of tag pairs

$$C(t_{i-1}, t_i)$$

2. Calculate probabilities using the counts

$$P(t_i | t_{i-1}) = \frac{C(t_{i-1}, t_i)}{\sum_{j=1}^N C(t_{i-1}, t_j)}$$

$C(t_{i-1}, t_i)$ is the count of times $tag (i-1)$ shows up before $tag i$. From this you can compute the probability that a tag shows up after another tag.