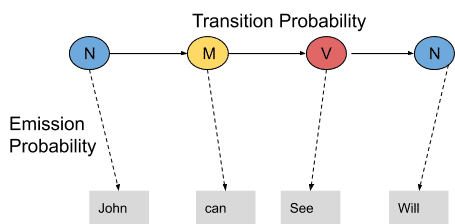
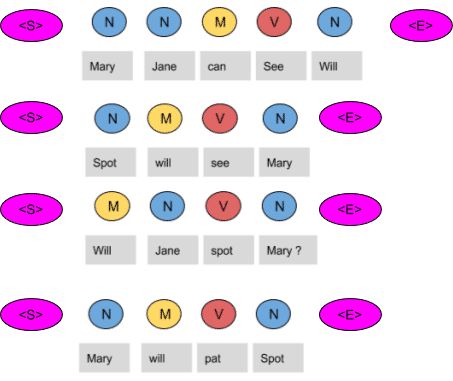
HMM





Emission Probability Table Sate Transition Table

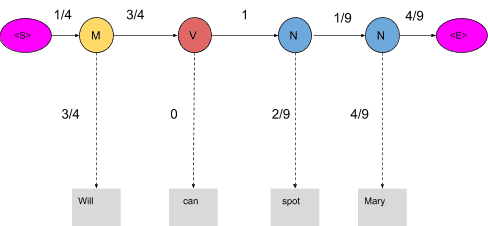
|  |  |  |  |
| --- | --- | --- | --- |
| Words | Noun | Model | Verb |
| Mary | 4/9 | 0 | 0 |
| Jane | 2/9 | 0 | 0 |
| Will | 1/9 | 3/4 | 0 |
| Spot | 2/9 | 0 | 1/4 |
| Can | 0 | 1/4 | 0 |
| See | 0 | 0 | 2/4 |
| pat | 0 | 0 | 1 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | N | M | V | <E> |
| <S> | 3/4 | 1/4 | 0 | 0 |
| N | 1/9 | 3/9 | 1/9 | 4/9 |
| M | 1/4 | 0 | 3/4 | 0 |
| V | 4/4 | 0 | 0 | 0 |

Take a new sentence and tag them with wrong tags. Let the sentence, ‘ Will can spot Mary’  be tagged as-

* Will as a  model
* Can as a verb
* Spot as a noun
* Mary as a noun

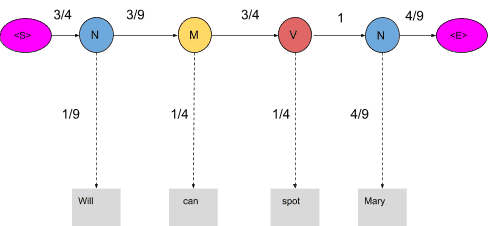
Now calculate the probability of this sequence being correct in the following manner.



Since the tags are not correct, the product is zero.

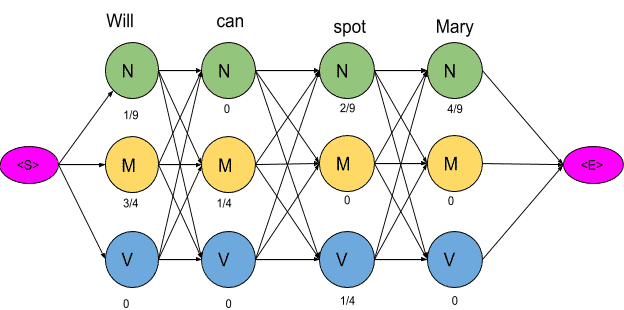
**1/4\*3/4\*3/4\*0\*1\*2/9\*1/9\*4/9\*4/9=0**

When these words are correctly tagged, we get a probability greater than zero as shown below

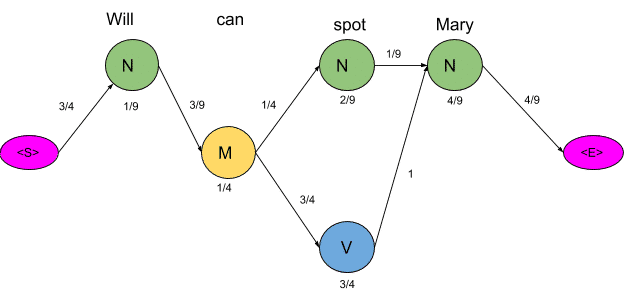


Calculating  the product of these terms we get,

**3/4\*1/9\*3/9\*1/4\*3/4\*1/4\*1\*4/9\*4/9=0.00025720164**



The next step is to delete all the vertices and edges with probability zero, also the vertices which do not lead to the endpoint are removed.



<S> to M (3/4) is deleted because M (3/4) has one non zero path and that is to M (1/4). But if you look up in the state transition table you will notice that M to M transition is 0.

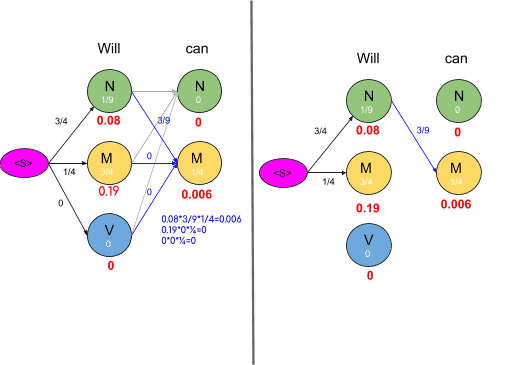
Now there are only two paths that lead to the end, let us calculate the probability associated with each path.

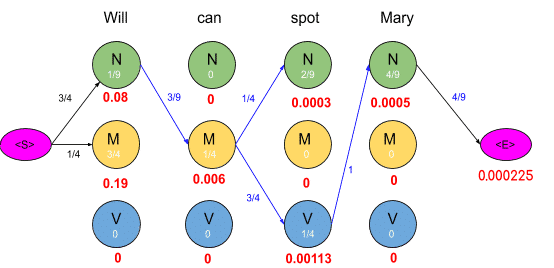
<S>→N→M→N→N→<E> =**3/4\*1/9\*3/9\*1/4\*1/4\*2/9\*1/9\*4/9\*4/9=0.00000846754**

<S>→N→M→V→N→<E>=**3/4\*1/9\*3/9\*1/4\*3/4\*1/4\*1\*4/9\*4/9=0.00025720164**

Clearly, the probability of the second sequence is much higher and hence the HMM is going to tag each word in the sentence according to this sequence.

**Optimizing HMM with Viterbi Algorithm**





To get an optimal path, we start from the end and trace backward, since each state has only one incoming edge, This gives us a path as shown below

