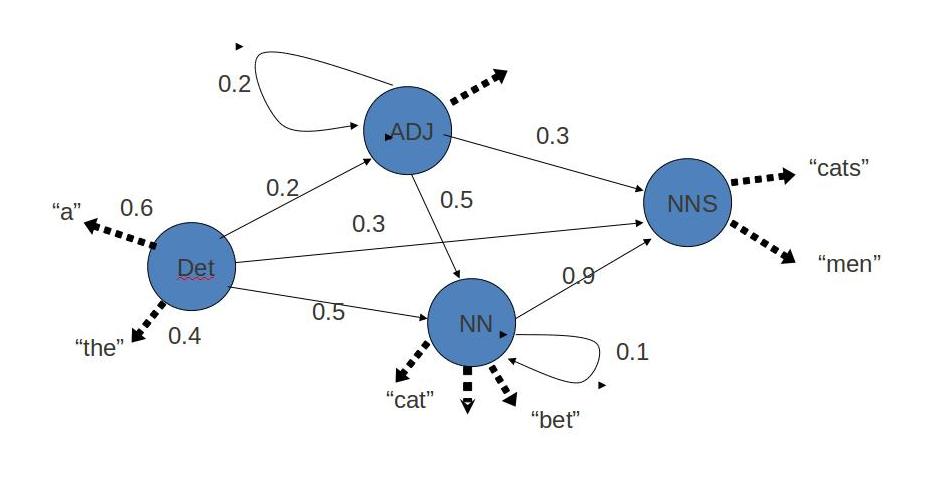
**Experiment 4**

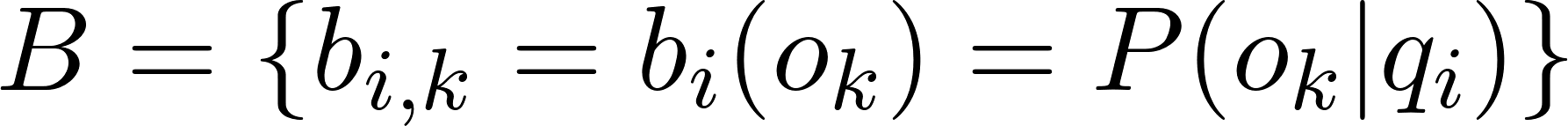
**Aim**  
To study and implement Hidden Markov Models (HMM) to calculate the probability of a sequence of tags using NLTK

**Theory**

A Hidden Markov Model (HMM) is a statistical Markov model in which the system being modelled is assumed to be a Markov process with unobserved (hidden) states. In a regular Markov model, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible.



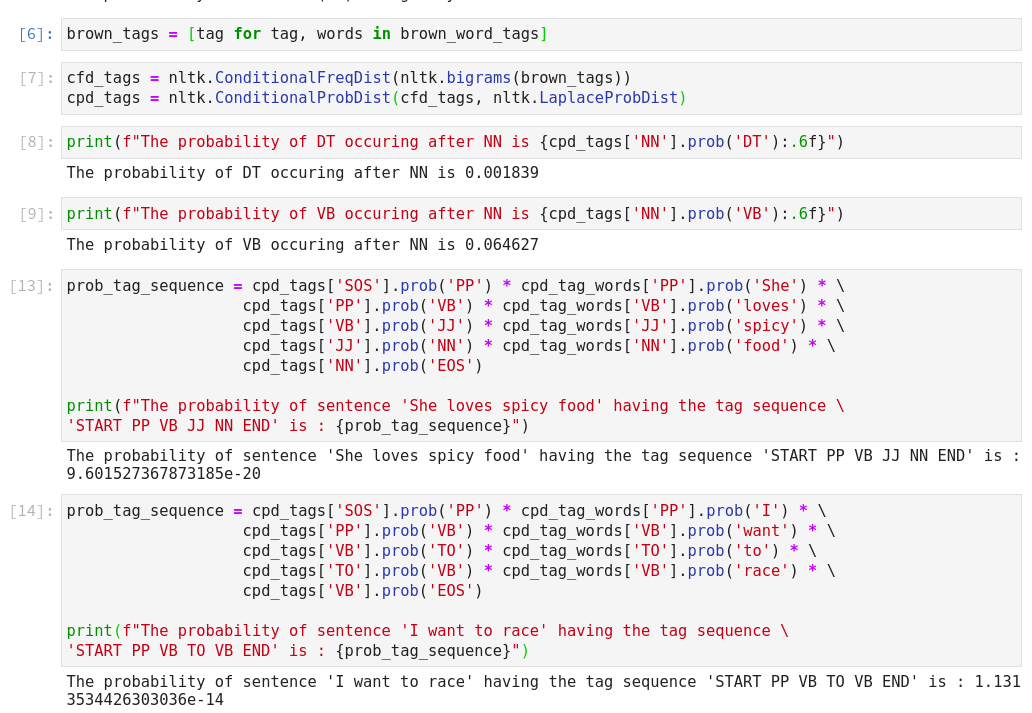
Hidden Markov Model has two important components:

1. Transition Probabilities: The one-step transition probability is the probability of transitioning from one state to another in a single step.
2. Emission Probabilities: The output probabilities for an observation from state. Emission probabilities [](https://www.codecogs.com/eqnedit.php?latex=B%20%3D%20%5C%7B%20b_%7Bi%2Ck%7D%20%3D%20b_i(o_k)%20%3D%20P(o_k%20%7C%20q_i)%20%5C%7D#0), where okis an Observation. Informally, B is the probability that the output is [](https://www.codecogs.com/eqnedit.php?latex=o_k#0) given that the current state is [](https://www.codecogs.com/eqnedit.php?latex=q_i#0)

For POS tagging; it is assumed that POS are generated as random processes, and each process randomly generates a word. Hence, transition matrix denotes the transition probability from one POS to another and emission matrix denotes the probability that a given word can have a particular POS.

**Code**





**Conclusion**Thus, studied the Hidden Markov Model and computed the probability tag sequence using HMMs.