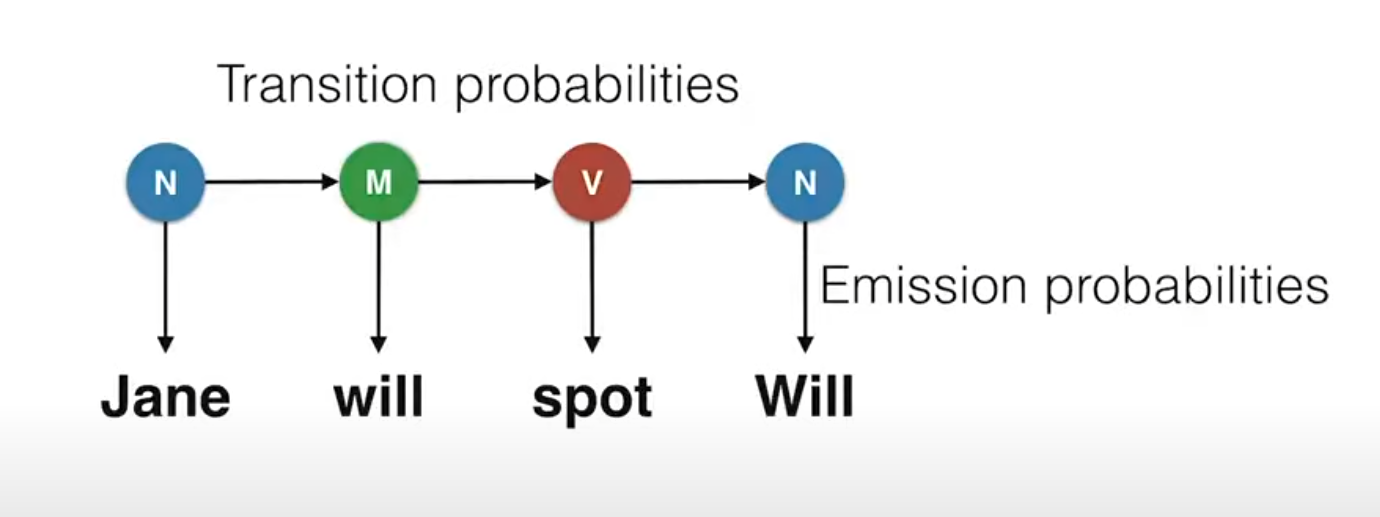
**CSC 417 Unit 3 Day 2 Outline**

1. Natural Language Processing
   1. Formal Grammars
      1. Definition
         1. Scheme for specifying the sentences allowed in a language
         2. Includes structural rules for combining words into well-formed clauses
      2. As rules become *more* restrictive, language becomes *simpler*/less expressive
         1. Human use of language often exploits ambiguity (e.g. poetry/creative writing)
   2. Hidden Markov Models
      1. Markov Chain
         1. System that transitions from one state to another according to probabilistic rules
         2. Probability of a given transition is not affected by any previous transition (depends only on current state)
         3. Allow us to calculate the probability of a specific sequence of transitions
      2. Hidden Markov Model
         1. Markov chain is *invisible* (cannot be directly observed)
            1. Transition probabilities are known, but we cannot see the states themselves
         2. State generates observations (e.g. an individual’s mood may be affected by the weather)
         3. NLP application: parts of speech tagging
            1. Transition probabilities

Probability of one part of speech following another

* + - * 1. Emission probabilities

Probability that a part of a speech is a particular word



* + - 1. Python implementations
         1. hmmlearn
         2. Pomegranate