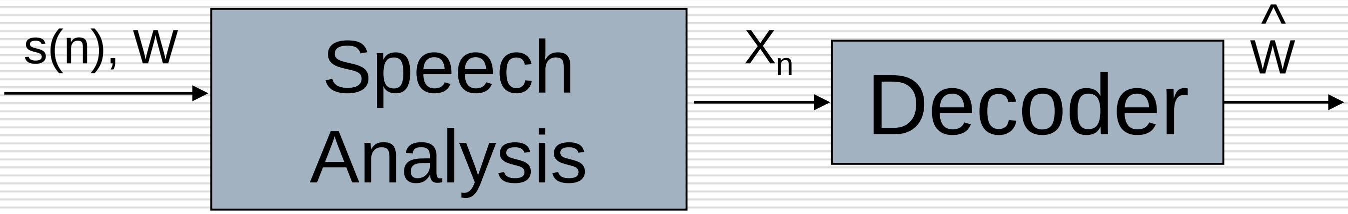
Acoustic Model & Hidden Markov Model (HMM)

Basic ASR Formulation



Basic ASR Formulation

The basic equation of Bayes rule-based speech recognition is

$$\hat{W} = \arg\max_{W} P(\mathbf{W} \mid \mathbf{X})$$

$$= \arg\max_{W} \frac{P(\mathbf{W})P(\mathbf{X} \mid \mathbf{W})}{P(\mathbf{X})}$$

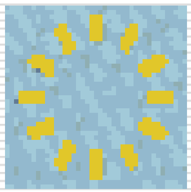
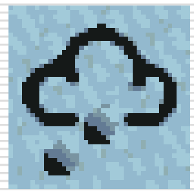
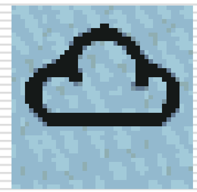
$$= \arg\max_{W} P(\mathbf{W})P(\mathbf{X} \mid \mathbf{W})$$

where $\mathbf{X}=X_1,X_2,\ldots,X_N$ is the acoustic observation (feature vector) sequence.

$$\hat{\mathbf{W}} = w_1 w_2 \ldots w_M$$

is the corresponding word sequence, $P(\mathbf{X}|\mathbf{W})$ is the acoustic model and $P(\mathbf{W})$ is the language model

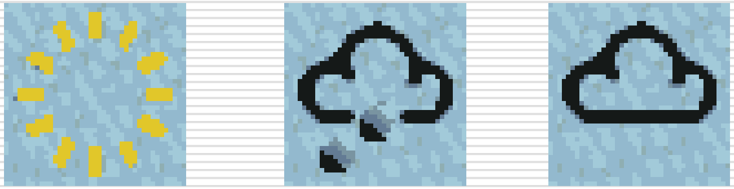
Non-Deterministic Patterns

- Suppose we want to model a weather forecast system.
- We cannot expect these three weather states to follow each other deterministically, but we might still hope to model the system that generates a weather pattern.
- We want to collect the following probabilities;

$$P\ (w_n \mid w_{n-1}, w_{n-2}, \ldots, w_1)$$

Non-Deterministic Patterns



- One way to do this is to assume that the state of the model depends only upon the previous states of the model.
- This is called the Markov assumption and simplifies problems greatly.
- When considering the weather, the Markov assumption presumes that today's weather can always be predicted solely given knowledge of the weather of the past few days

Markov Process

- A Markov process is a process which moves from state to state depending (only) on the previous *n* states.
- The process is called an *order n* model where *n* is the number of states affecting the choice of next state.
- The simplest Markov process is a first order process, where the choice of state is made purely on the basis of the previous state.

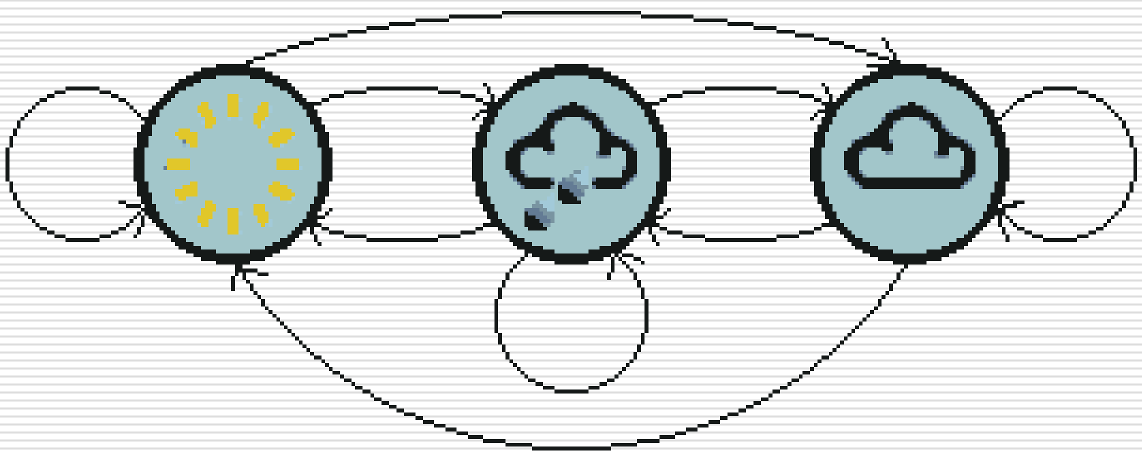
Markov Process

- The simplest Markov process is a first order process, which can be expressed as;

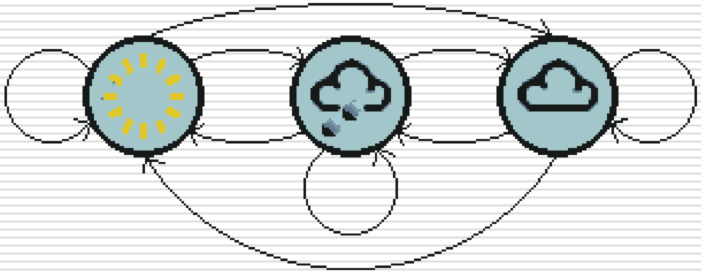
$$P\ (w_n \mid w_{n-1}, w_{n-2}, \ldots, w_1) \approx P\ (w_n \mid w_{n-1})$$

Markov Process

- The figure below shows all possible first order transitions between the states of the weather example.



Markov Process



- Notice that for a first order process with M states, there are M^2 transitions between states since it is possible for any one state to follow another.

Markov Process

- Associated with each transition is a probability called the state transition probability - this is the probability of moving from one state to another.
- These M^2 probabilities may be collected together in an obvious way into a state transition matrix. Notice that these probabilities do not vary in time. This is an important (if often unrealistic) assumption.

The joint probability using the Markov Assumption:

$$P(W1, \ldots, Wn) = \prod_{i=1}^{n} P(Wi \mid Wi-1)$$

Markov Process

Transition Matrix

$$A = \{a_{ij}\} = \begin{array}{c|ccc} & S_1 & S_2 & S_3 \\ \hline S_1 & 0 & 0 & 1 \\ S_2 & 1/2 & 1/2 & 0 \\ S_3 & 1/3 & 2/3 & 0 \end{array}$$

Question: What is the probability of states sequence of $\{S_3, S_2, S_1\}$

$$P(S_3, S_2, S_1) = P(S_3)P(S_2 \mid S_3)P(S_1 \mid S_2) = \frac{1}{3} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{9}$$

Question

Given that the weather on day 1 is sunny (state 3),
What is the Probability that
The weather for eight consecutive days is "sun-sun-sun-rain-rain-sun-cloudy-sun"?

	rain	cloud	sun
rain	0.4	0.3	0.3
cloud	0.2	0.6	0.2
sun	0.1	0.1	0.8

Solution: **O** = sun sun sun rain rain sun cloudy sun

3 3 3 1 1 3 2 3

$$P(\bar{O}|Model) = P[3]P[3|3]P[3|3]P[1|3]P[1|1]P[3|1]P[2|3]P[3|2]$$

$$= \pi_3 a_{33} a_{33} a_{31} a_{11} a_{13} a_{32} a_{23} = 1 \cdot (0.8)(0.8)(0.1)(0.4)(0.3)(0.1)(0.2)$$

$$= 1.536 \times 10^{-4}$$

Markov Process

- We have now defined a first order Markov process consisting of :

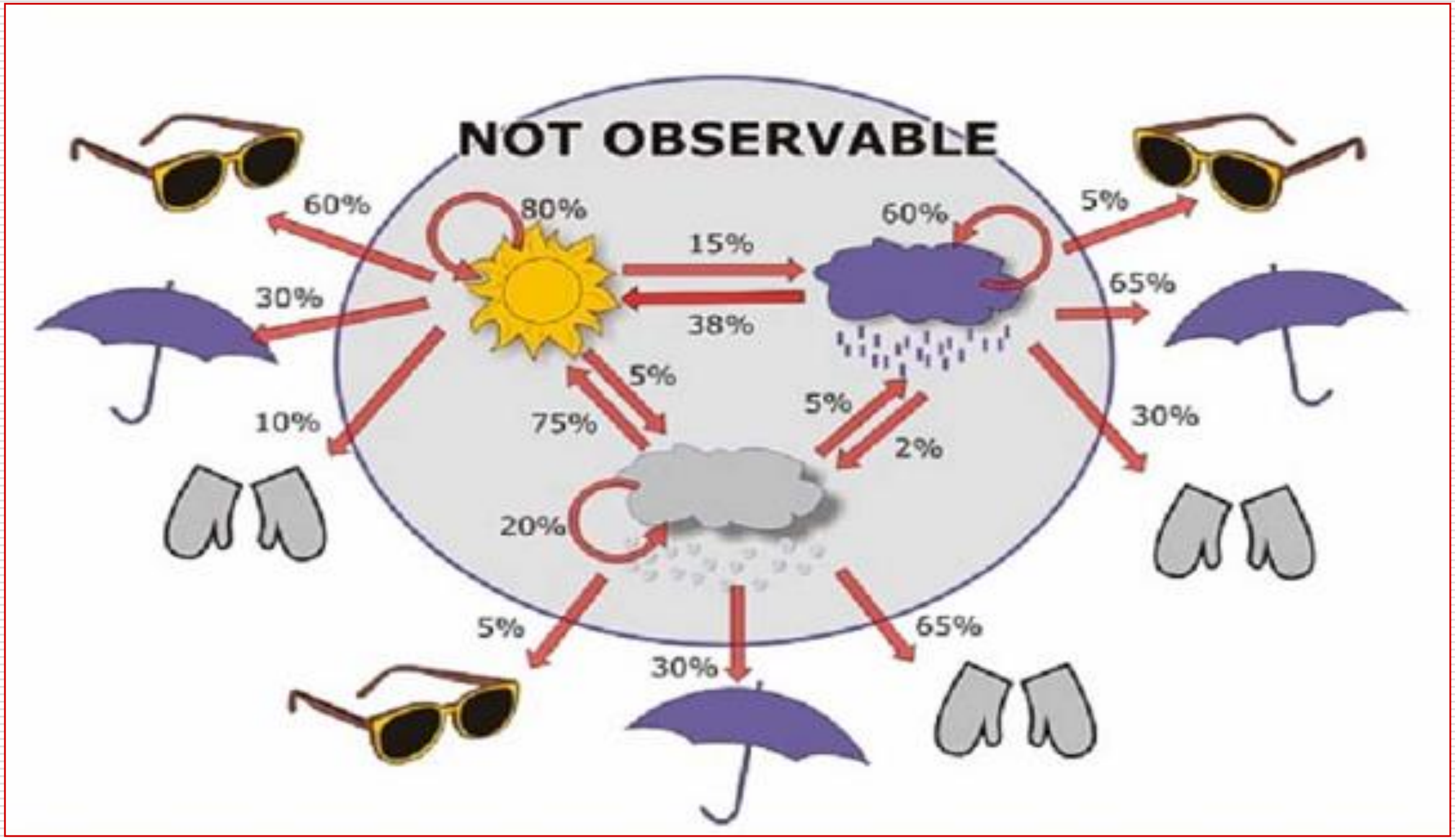
1. **states :** Three states - sunny, cloudy, rainy.
2. π **vector** : Defining the probability of the system being in each of the states at time 0.
3. **state transition matrix** : The probability of the weather given the previous day's weather.

- Any system that can be described in this manner is a Markov process.

From Markov To Hidden Markov

- **The previous model assumes that each state can be uniquely associated with an observable event**
  - Once an observation is made, the state of the system is then trivially retrieved
  - This model, however, is too restrictive to be of practical use for most realistic problems
- **To make the model more flexible, we will assume that the outcomes or observations of the model are a probabilistic function of each state**
  - Each state can produce a number of outputs according to a probability distribution, and each distinct output can potentially be generated at any state
  - These are known a **Hidden Markov Models (HMM)**, because the state sequence is not directly observable, it can only be approximated from the sequence of observations produced by the system

From Markov To Hidden Markov



Patterns generated by a hidden process

- A more realistic problem is that of recognizing speech; the sound that we hear is the product of the vocal chords, size of throat, position of tongue and several other things.
- Each of these factors interact to produce the sound of a word, and the sounds that a speech recognition system detects are the changing sound generated from the internal physical changes in the person speaking.

Patterns generated by a hidden process

- In addition to the probabilities defining the Markov process, we therefore have another matrix, termed the confusion matrix, which contains the probabilities of the observable states given a particular hidden state.