Markov Chains

Markov chains with matrices

- Define a transition matrix for the chain:

$$T = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$

- Let v_t = a row vector representing the probability that the chain is in each state at time t.
- $v_t = v_{t-1} * T$

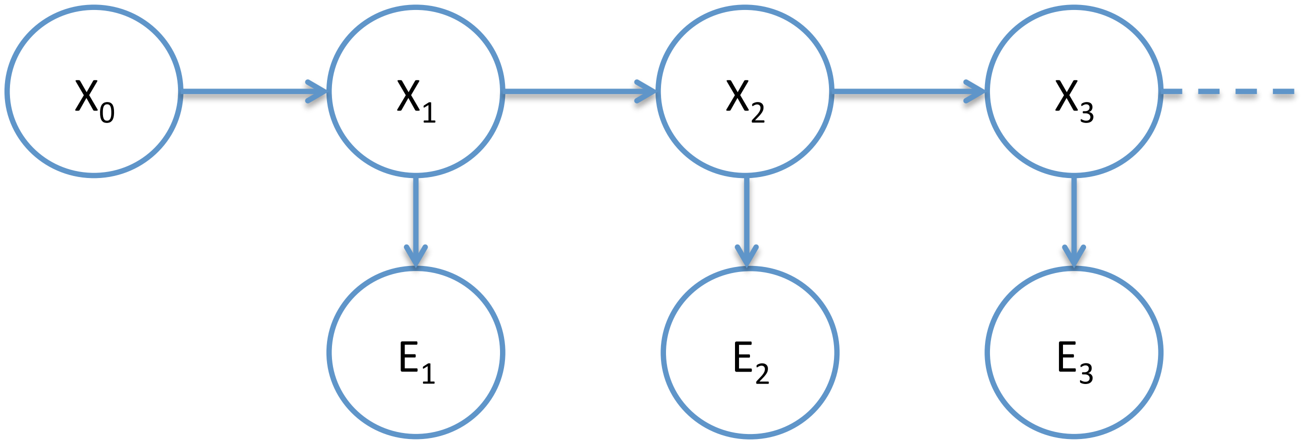
- Markov chains are pretty easy!
- But sometimes they aren't realistic…

- What if we can't directly know the states of the model, but we can see some indirect evidence resulting from the states?

Weather

- Regular Markov chain
  - Each day the weather is rainy or sunny.
  - P(Xt = rain | Xt-1 = rain) = 0.7
  - P(Xt = sunny| Xt-1 = sunny) = 0.9
- Twist:
  - Suppose you work in an office with no windows. All you can observe is weather your colleague brings her umbrella to work.

Hidden Markov Models

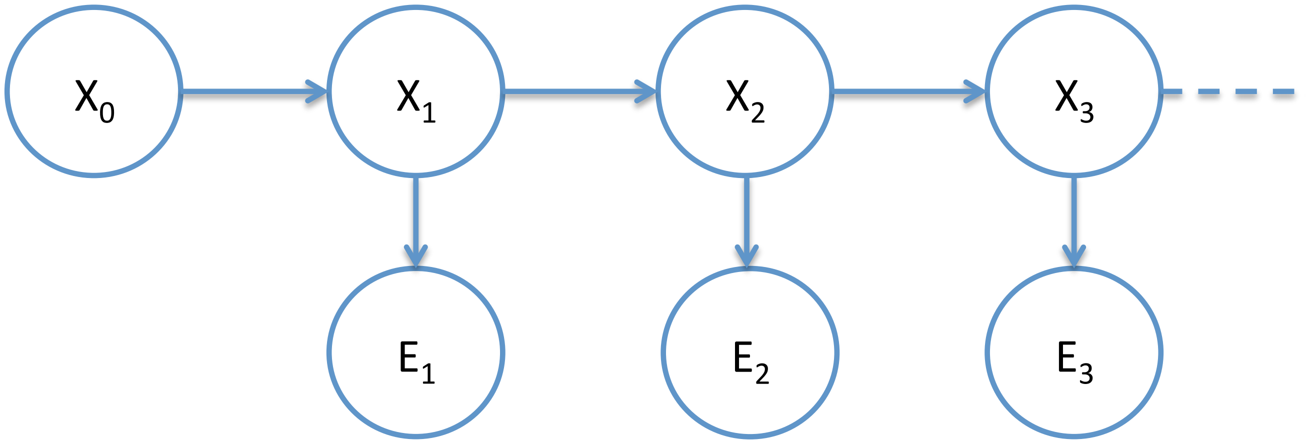


- The X's are the state variables (never directly observed).
- The E's are evidence variables.

Common real-world uses

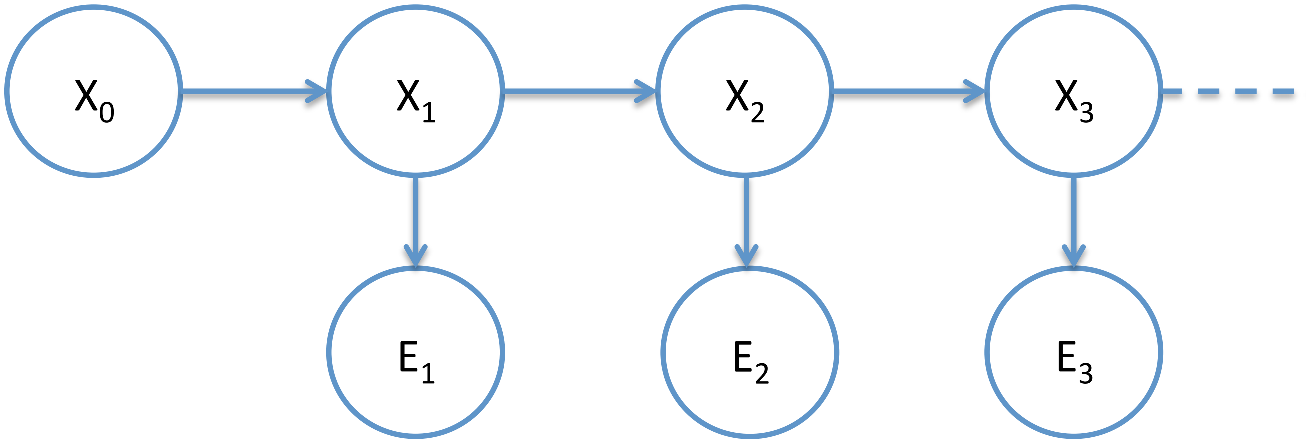
- Speech processing:
  - Observations are sounds, states are words.
- Localization:
  - Observations are inputs from video cameras or microphones, state is the actual location.
- Video processing (example):
  - Extracting a human walking from each video frame. Observations are the frames, states are the positions of the legs.

Hidden Markov Models



- $P(X_t \mid X_{t-1}, X_{t-2}, X_{t-3}, \ldots) = P(X_t \mid X_{t-1})$
- $P(X_t \mid X_{t-1}) = P(X_1 \mid X_0)$
- $P(E_t \mid X_{0:t}, E_{0:t-1}) = P(E_t \mid X_t)$
- $P(E_t \mid X_t) = P(E_1 \mid X_1)$

Hidden Markov Models



- What is $P(X_{0:t}, E_{1:t})$?

$$P(X_0)\prod_{i=1}^{t} P(X_i \mid X_{i-1})P(E_i \mid X_i)$$

Common questions

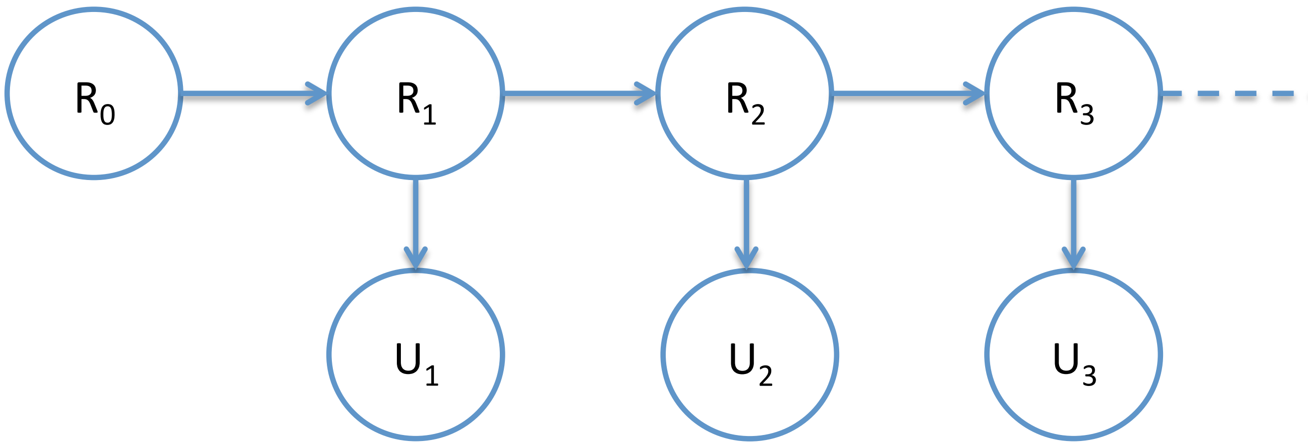
- **Filtering**: Given a sequence of observations, what is the most probable current state?
  – Compute $P(X_t \mid e_{1:t})$
- **Prediction**: Given a sequence of observations, what is the most probable future state?
  – Compute $P(X_{t+k} \mid e_{1:t})$ for some $k > 0$
- **Smoothing**: Given a sequence of observations, what is the most probable past state?
  – Compute $P(X_k \mid e_{1:t})$ for some $k < t$

Common questions

- **Most likely explanation:** Given a sequence of observations, what is the most probable sequence of states?
  - Compute $\underset{x_{1:t}}{\operatorname{argmax}} P(x_{1:t} \mid e_{1:t})$

- **Learning**: How can we estimate the transition and sensor models from real-world data? (Future machine learning class?)

Hidden Markov Models



- $P(R_t = yes \mid R_{t-1} = yes) = 0.7$
  $P(R_t = yes \mid R_{t-1} = no) = 0.1$
- $P(U_t = yes \mid R_t = yes) = 0.9$
  $P(U_t = yes \mid R_t = no) = 0.2$

Filtering

- Filtering is concerned with finding the most probable "current" state from a sequence of evidence.
- Let's compute this.

Forward algorithm

- Recursive computation of the probability distribution over current states.
- Say we have $P(X_t \mid e_{1:t})$

$$P(X_{t+1} \mid e_{1:t+1}) =$$

$$\alpha P(e_{t+1} \mid X_{t+1}) \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t \mid e_{1:t})$$

Forward algorithm

- Markov chain version:

$$P(X_{t+1}) = \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t)$$

- Hidden Markov model version:

$$P(X_{t+1} \mid e_{1:t+1}) =$$

$$\alpha P(e_{t+1} \mid X_{t+1}) \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t \mid e_{1:t})$$

Forward algorithm

- What is the probability of rain today given that the umbrella has been seen for the past two days?