# Markov Models

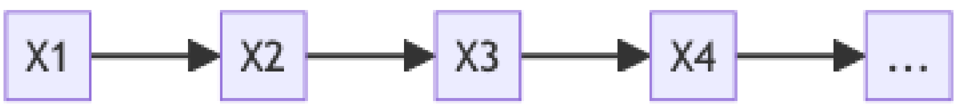
- Want to reason about a sequence of observation
- Need to introduce time (or space) into our models and update beliefs based on
  - Getting more evidence (done with BNs)
  - World changing over time/space (Markov models)

## Probability Recap

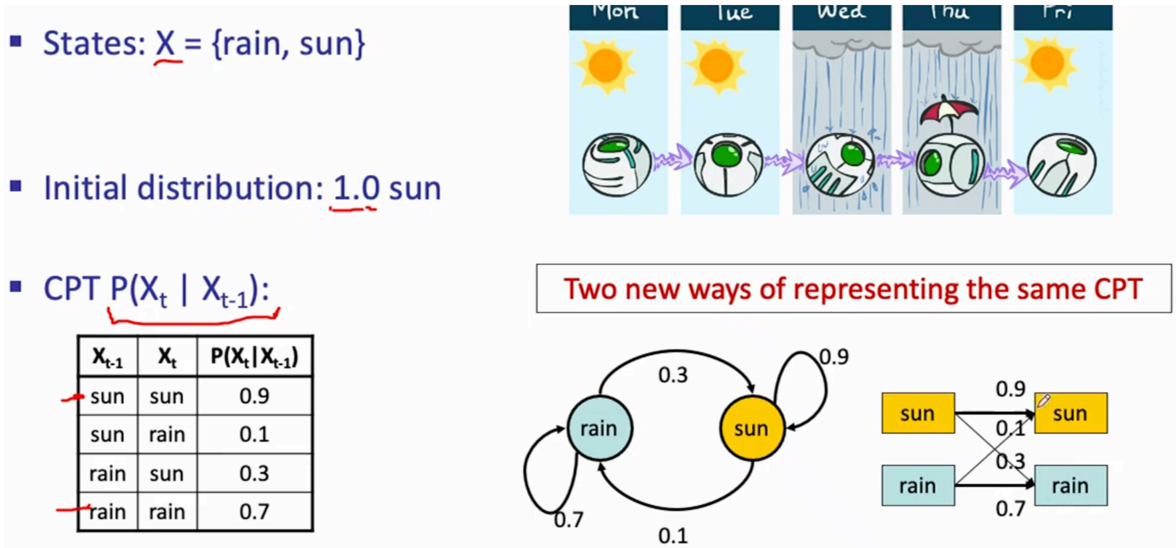
- Conditional probability := $\mathbb{P}(x|y) = \frac{\mathbb{P}(x,y)}{\mathbb{P}(y)}$
- Marginal probability := $\mathbb{P}(x) = \sum_y \mathbb{P}(x, y)$
- Product rule := $\mathbb{P}(x, y) = \mathbb{P}(x|y)\mathbb{P}(y)$
- Chain rule := $\mathbb{P}(X_1, X_2, X_n) = \prod_{i=1}^{n} \mathbb{P}(X_i|X_1, ...X_{i-1})$
- $X$ independent with $Y \iff \forall x, y : \mathbb{P}(x, y) = \mathbb{P}(x)\mathbb{P}(y)$
- $X$ and $Y$ conditionally independent given $Z \iff \forall x, y, z : \mathbb{P}(x, y|z) = \mathbb{P}(x|z)\mathbb{P}(y|z)$
  - Basically independent, but tagging on given $z$ to every probability function
- Proportionality $\triangleq$ $\mathbb{P}(X) \propto_x f(X) \iff \mathbb{P}(X) = kf(X)|_{k:=\text{constant not dependent on x}}$
  - Equivalent to $\mathbb{P}(X) = \frac{f(X)}{\sum_x f(x)}$

## Markov Chains

- How beliefs about state change with passage of time
- The value of $X$ at a given time is called the state
  - Each $X$ includes the full distribution over all possible states




- Parameter := transition probabilities or dynamics; specify how the state evolves over time (also, initial state probabilities)
  - Initial state probabilities := $\mathbb{P}(X_1)$
  - Transition probabilities := $\mathbb{P}(X_t|X_{t-1})$
- Stationary assumption := transition probabilities are the same at all times
  - $\mathbb{P}(X_1|X_0) = \mathbb{P}(X_2|X_1)$
  - The directed edges are the same probability for all transitions
- Same as MDP transition model, but with no choice of action
- A ‘growable’ BN (can always use BN methods if we truncate to fixed length)
- Past and future are independent given the present
  - Each time step only depends on the previous
  - Called the (first order) Markov property




- eg:
  - $\mathbb{P}(X_2 = \text{sun}) = \sum_{x_1} \mathbb{P}(x_1, X_2 = \text{sun}) = \sum_{x_1} \mathbb{P}(X_2 = \text{sun}|x_1)\mathbb{P}(x_1)$

Mini-Forward Algorithm

- Given $\mathbb{P}(X_1)$ and $\mathbb{P}(X_t|X_{t-1})$
- Alg: $\mathbb{P}(x_t) = \sum_{x_{t-1}} \mathbb{P}(x_{t-1}, x_t) = \sum_{x_{t-1}} \mathbb{P}(x_t|x_{t-1})\mathbb{P}(x_{t-1})$
  - Have to build up $\mathbb{P}(x_{t-1})$ recursively
- Converges to some fixed (stationary) distribution
  - Same when extended for any arbitrary state in the simulation
  - Starting state does not matter

Stationary Distributions

- Distribution converges
- Denoted with $\mathbb{P}_\infty$
- $\mathbb{P}_\infty(X) = \mathbb{P}_{\infty+1}(X) = \sum_x \mathbb{P}(X|x)P_\infty(x)$
- Note that the sum of all probabilities at their stationary distribution sum to 1
  - Can give an extra equation if needed

## Hidden Markov Models (HMMs) formulation

- How beliefs change with passage of time and evidence
- Underlying Markov chain over states $X$
- You observe outputs (effects) at each time step
- Effects / emissions at each step $:= \mathbb{P}(E_t|X_t)$
  - Just another table like the transition table given the state
- Have two important independence properties
  1. Markov hidden process $\triangleq$ future depends on past via the present
  2. Current observation independent of all else given current state
- Evidence variables are not necessarily independent (need some given state between them to make them independent)

## Filtering with HMMs

- How to infer beliefs from evidence
- Belief state $:= B_t(X) = \mathbb{P}_t(X_t|e_1, ..., e_t)$
  - Belief of $X$ at time $t$

- Start with a guess $B_1(X)$ (usually uniform)
  - Update with more evidence gained over time
- Idea: start with $\mathbb{P}(X_1)$ and try to derive $B_t$ in terms of $B_{t-1}$
  - $\mathbb{P}(X_{t+1}|e_{1:t}) = \sum_{x_t} \mathbb{P}(X_{t+1}, x_t|e_{1:t}) = \sum_{x_t} \mathbb{P}(X_{t+1}, x_t|e_{1:t})\mathbb{P}(x_t|e_{1:t})$
    * The final factor of the last equation is the previous belief
    * Can ignore the evidence in the first half of the summation yielding $\sum_{x_t} \mathbb{P}(X_{t+1}|x_t)\mathbb{P}(x_t|e_{1:t})$
  - $B'(X_{t+1}) = \sum_{x_t} \mathbb{P}(X'|x_t)B(x_t)$
  - Beliefs get 'pushed' through the transitions