CSCI 3202: Intro to Artificial Intelligence

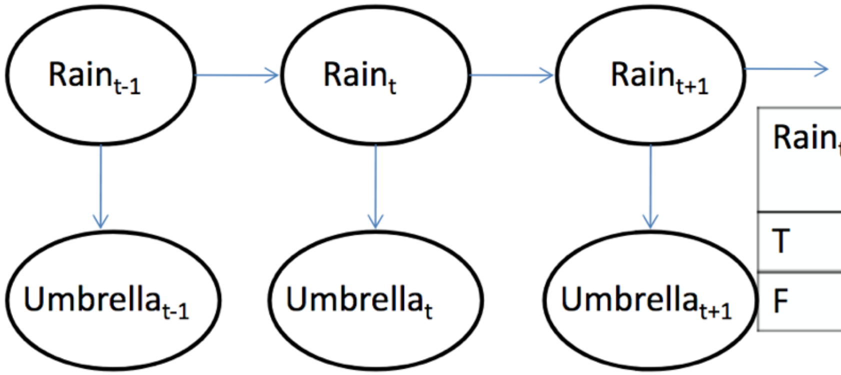
Lecture 29: HMM, Viterbi algorithm

Rhonda Hoenigman

Department of Computer Science

$Rain_{t-1}$	$P(Rain_t)$
T	0.7
F	0.3

$P(Rain_0 = T) = 0.5$



$Rain_t$	$P(Umbrella_t)$
T	0.9
F	0.2

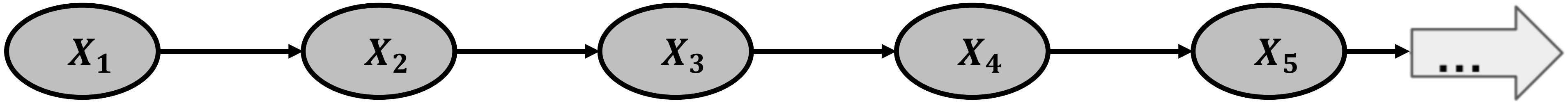
HMMs — Adapted from Russel and Norvig, Chapter 15.

Recap – Markov Models

A **Markov model** is a chain-structured Bayesian network.

- The value of X at a time t is the state at time t
- Stationary Markov model: All subsequent nodes have the same CPT (identically distributed)

Example: Is it raining? Let X_t denote the event that it is raining on day t

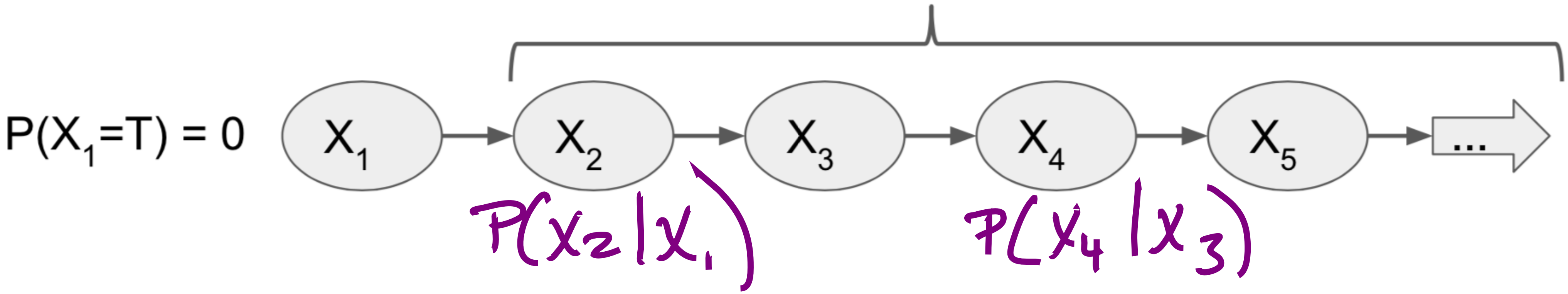


Markov Models

Recall: Causal chain Bayes net forms conditional independence

- Past and future are independent of the present.
- State at $t + 1$ only depends on state at t
- The CPTs give the transition probabilities from one state to another for this.

X_t	$P(X_{t+1}=T \mid X_t)$
T	0.3
F	0.1



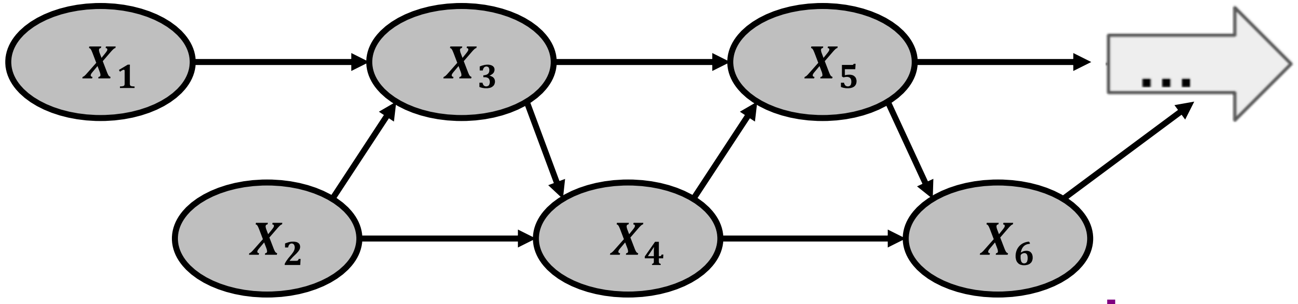
Markov Models

Recall: Causal chain Bayes net forms conditional independence

- Past and future are independent of the present.
- State at $t+1$ only depends on state at t
- The CPTs give the transition probabilities from one state to another for this.

Markov property
(first-order)

➢ We could also create models where the state at $t+1$ depends only on the state at t and $t-1$ (second-order Markov property), or higher …



$P(X_6 \mid X_5, X_4)$

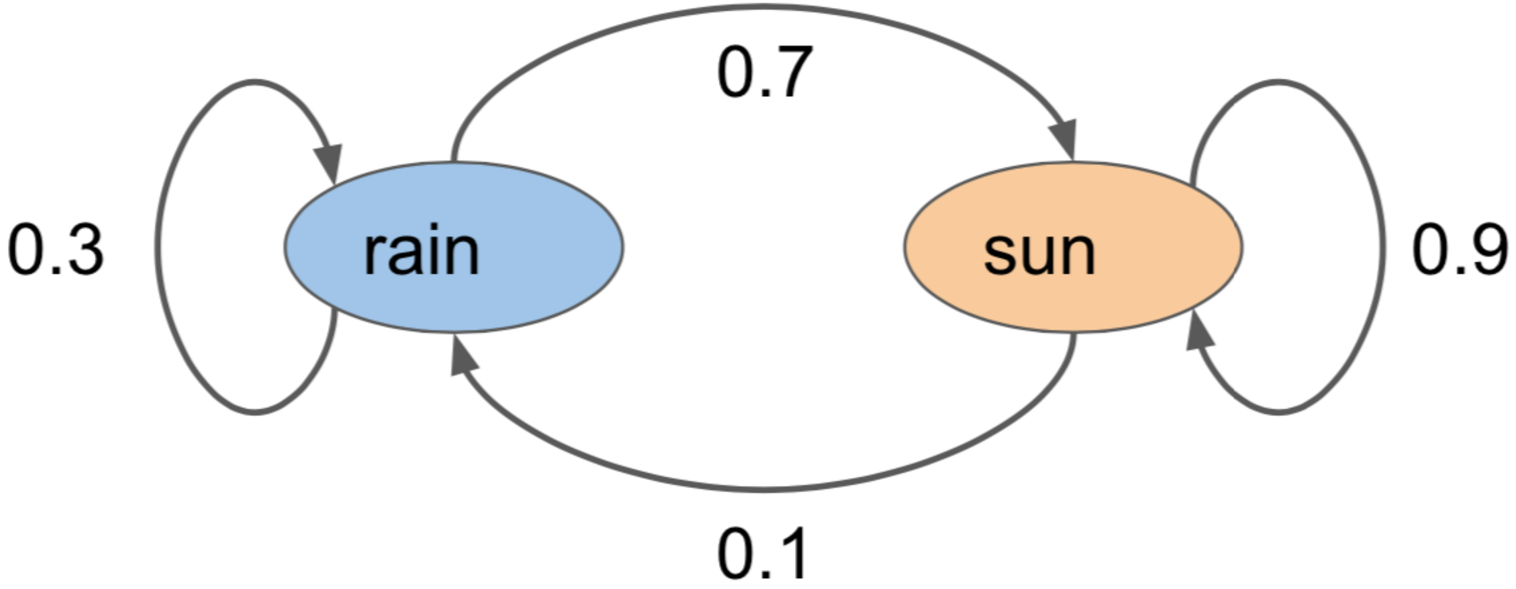
X_{t-1}	X_t	$P(X_{t+1}=T \mid X_t, X_{t-1})$
T	T	0.3
T	F	0.1
F	T	0.4
F	F	0.1

Markov Models

Mini-forward algorithm: incremental belief updating. Calculate probability distribution at time t using prior distributions

Suppose that x_1 is known.

Then for $t = 2, 3, 4, \ldots P(x_t) = \sum_{x_{t-1}} P(x_t|x_{t-1})P(x_{t-1})$

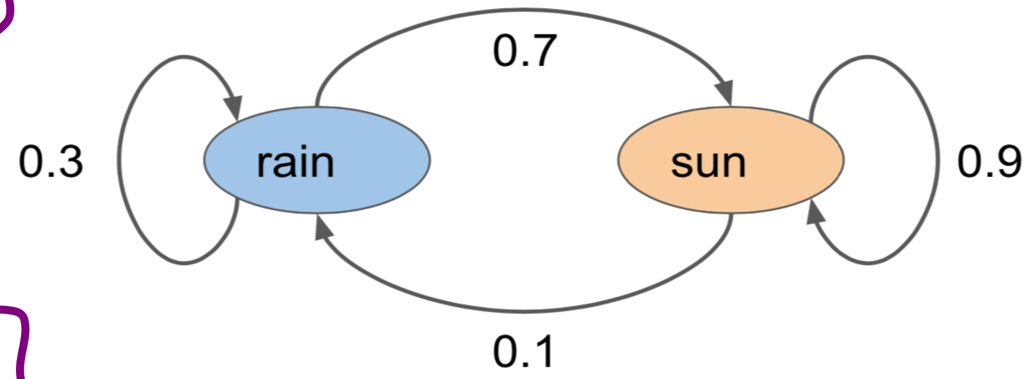


Mini_forward Algorithm

Example: Using the mini-forward algorithm, find $P(X_3 = s)$

Assume $P(X_1 = \text{sun})$

$$P(X_t) = \sum_{t-1} P(X_t \mid X_{t-1}) P(X_{t-1})$$



$$P(X_2) = P(X_2 \mid X_1 = s) P(X_1 = s) + P(X_2 \mid X_1 = r) P(X_1 = r)$$

$$= P(X_2 \mid X_1 = s) = \begin{array}{c} \begin{bmatrix} .9 & .1 \end{bmatrix} \\ \begin{matrix} s & \;\; r \end{matrix} \end{array}$$

$$P(X_3) = \sum_{x_2} P(X_3 \mid X_2) P(X_2)$$

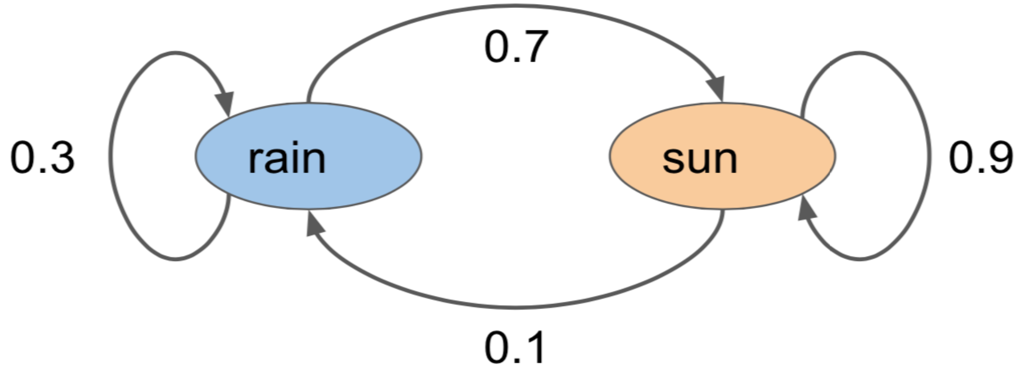
$$= P(X_3 = s \mid X_2 = s) P(X_2 = s) + P(X_3 = s \mid X_2 = r) P(X_2 = r)$$

$$= [.9 \times .9 + .7 \times .1] = .88$$

Markov Models

Define the transition matrix T such that T_{ij} = probability of moving from state i to j

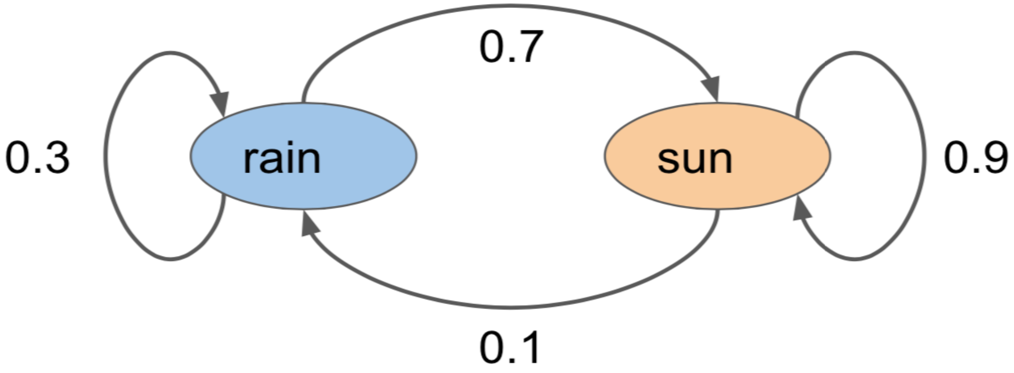
Say X_1 = sun, X_2 = rain. Then: $T = \begin{bmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{bmatrix}$



Markov Models

It turns out that the transition matrix also gives us a short-cut for calculating multi-step transition probabilities.

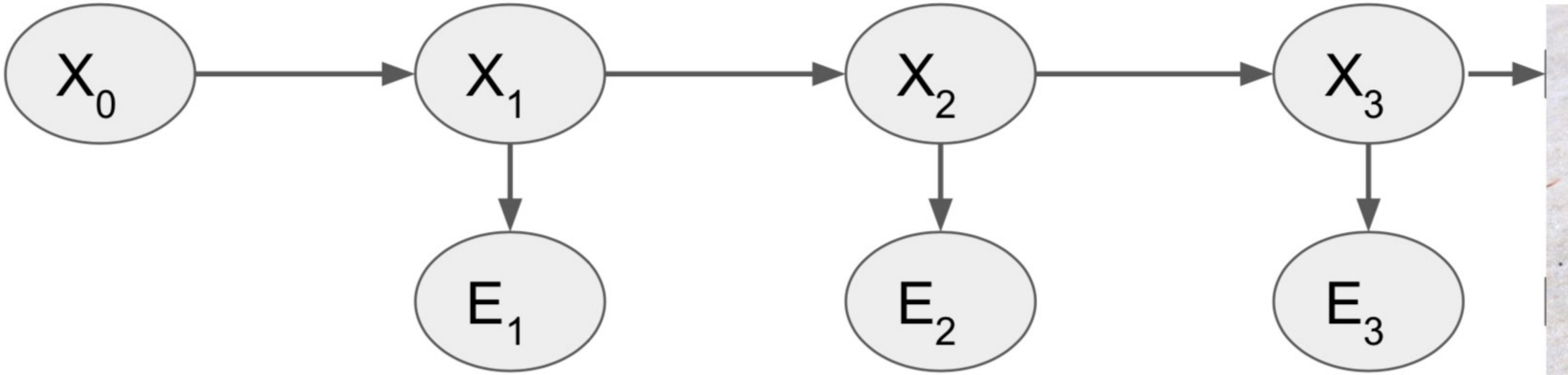
Example: It is sunny today. What is the probability that it will be sunny 2 days from now?



Hidden Markov Models (HMMs)

Example: Suppose you are a graduate student in a basement office.
You are curious if it is raining, and the only contact you have with the outside world is through your advisor. If it is raining, she brings her umbrella 90% of the time, and has it just in case on 20% of sunny days. You know that historically, 40% of rainy days were followed by another rainy day, and 30% of sunny days were followed by a rainy day.
Evidence: advisor carrying an umbrella
State: Rain or sun

Can you predict the state X_t from the evidence E_t?

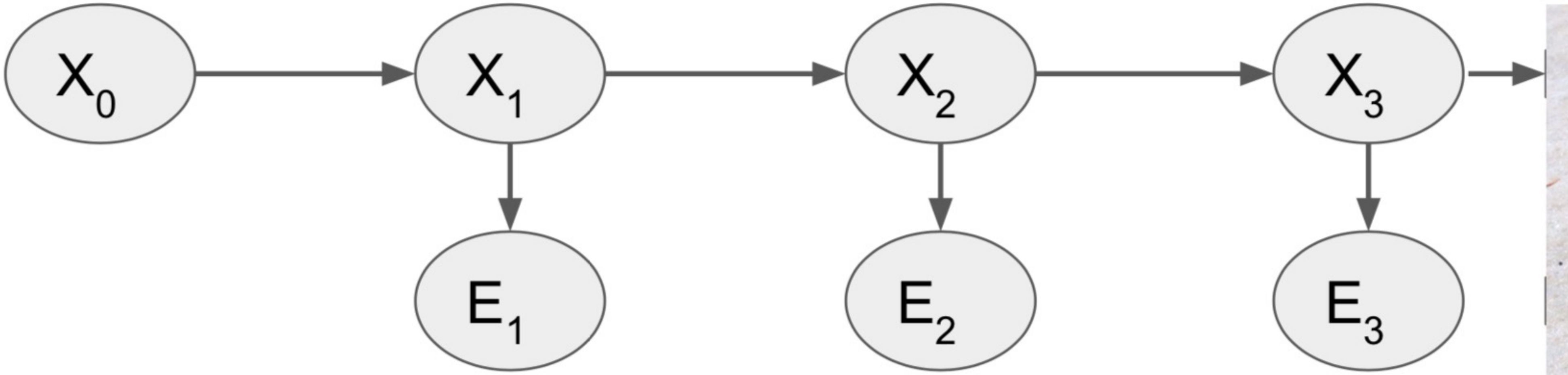


Hidden Markov Models (HMMs)

Example: Suppose you are a graduate student in a basement office. You are writing your dissertation, so you don't get to leave very often.

Evidence: advisor carrying an umbrella
State: Rain or sun

Can you predict the state X_t from the evidence E_t?



Hidden Markov Models（HMMs）

Example: People in a small village are either Healthy or have Fever. A doctor diagnoses Fever by asking people how they feel, and they say either normal, cold, or dizzy.

A patient visits the doctor two days in a row. What is P(Fever) on each day if they report feeling normal on day 1 and dizzy on day 2?

Evidence: How the patient feels.

State: Healthy or Fever

Hidden Markov Models (HMMs) – inference

- **Filtering:** Type of inference. Compute the belief state given the evidence observed so far.
  - $P(X_t \mid e_{1:t})$

  - Examples:
    - What is probability of rain today given the observation of the umbrella every day, including today?
    - What is the probability of Fever given the observation of cold, dizzy, cold evidence?

HMM example

Example: People in a small village are either Healthy or have Fever. A patient visits three days in a row and reports feeling normal, dizzy, then normal. What is the P(Fever) on each day?

States: Healthy, Fever

Evidence: normal, dizzy, cold

Initial distribution Healthy = 0.6
Fever = 0.4

Transition CPT

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
h	h	.7
h	f	.3
f	h	.4
f	f	.6

Emission CPT

Healthy	$P(E_t \mid \text{Healthy})$
dizzy	0.1
cold	0.4
normal	0.5

Fever	$P(E_t \mid \text{Fever})$
dizzy	0.6
cold	0.3
normal	0.1

HMM Example

Track current state estimate and update it as new evidence becomes available

We want:

$$P(X_t | E_{1:t}) = \underbrace{\sum_{X_{t-1}} P(X_t | X_{t-1}) P(X_{t-1})}_{\text{Incorporates evidence up through } X_{t-1} \text{ and transition to } X_t} \; \underbrace{\alpha P(E_t | X_t) P(X_t)}_{\text{Incorporates } E_t}$$

HMM Example – Day 1

Initial dist: $[0.6, 0.4]$

$$P(X_1) = \sum_{x_0} P(X_1 | X_0) P(X_0)$$

$$= [h \to h + f \to h, h \to f + f \to f]$$

$$= [0.7 + 0.6 + 0.4 \times 0.4, 0.3 \times 0.6 + 0.6 \times 0.4]$$

$$= [.58, .42]$$

Incorporate evidence

$$P(X_1 | normal_1) = \alpha P(normal_1 | X_1) P(X_1)$$

$$= \alpha [.5 \times .58, .1 \times .42]$$

$$= \alpha [.29, .042] \qquad \alpha = .29 + .042$$

$$= [.87, .13]$$

HMM Example – Day 2

Patient feels dizzy

Our belief about the state of the patient influences our belief on day 2.

$P(X_2 \mid normal_1, dizzy_2)$

$$P(X_2 \mid normal_1) = \sum_{X_1} P(X_2 \mid X_1) P(X_1 \mid normal_1)$$

$$= [.7 \times .87 + .4 \times .13,\ .3 \times .87 + .6 \times .13]$$

$$= [.661,\ .339]$$

<u>Evidence</u>

$$P(X_2 \mid normal_1, dizzy_2) = \alpha P(dizzy_2 \mid X_2) P(X_2 \mid normal_1)$$

$$= \alpha [.1 \times .661,\ .6 \times .339]$$

$$= [.245,\ .755]$$

HMM Example – Day 3

Hidden Markov Models (HMMs) – most likely explanation

- Given a sequence of observations, find the most likely state sequence that generated those observations.

  - $argmax_{\mathbf{x}1:t}\, P(\mathbf{X}_{1:t} \mid \mathbf{e}_{1:t})$

  - Examples:
    - Umbrella appears three days in a row and then doesn't appear on the fourth day. The most likely explanation is that it rained on the first three days and then not on the fourth day.

    - Fever, Healthy example

HMM Example – Viterbi algorithm

Example: People in a small village are either Healthy or have Fever. A patient visits three days in a row and reports feeling normal, dizzy, then normal. What is the most likely sequence that generates these observations?

- Viterbi algorithm: Given a sequence of observations, find the most likely state sequence that generated those observations.
  - *Generate trellis of state transitions and values, then work backwards to get path*
  - $max(P(e_t|X_t)P(X_t|X_{t-1})P(X_{t-1}))$
  - $argmax_{\mathbf{x}1:t}\, P(\mathbf{X}_{1:t} \mid \mathbf{e}_{1:t})$

HMM Example – Day 1 – with Viterbi

HMM Example – Day 2 – Viterbi

HMM Example – Day 3 – Viterbi

HMM Example – Viterbi