CSC 384
Introduction to Artificial Intelligence

Uncertainty 5
Hidden Markov Models and Filtering

Alice Gao and Randy Hickey

Winter 2023

Learning Goals

By the end of this lecture, you should be able to

1. Explain the independence assumptions in a Hidden Markov Model.
2. Calculate the filtering probabilities for a time step in a Hidden Markov Model.

Outline

1. Introducing a Hidden Markov Model
2. Common Inference Tasks
3. Filtering

Inference in a Changing World

- So far, we can reason probabilistically in a static world.
- However, the world evolves over time. In an evolving world, we must reason about a sequence of events.
- Applications of sequential belief networks:
  - weather predictions
  - stock market predictions
  - patient monitoring
  - robot localization
  - speech and handwriting recognition

The Umbrella Story

You are a security guard stationed at a secret underground installation.

You want to know whether it is raining today.

Unfortunately, your only access to the outside world occurs each morning when you see the director coming in with, or without, an umbrella.

States and Observations

- There is a series of time steps.

- S_t denotes the hidden state at time t.

Umbrella story:

$S_t = 1$ means it rains on day t, and $S_t = 0$ otherwise.

- E_t denotes the observation at time t.

Umbrella story:

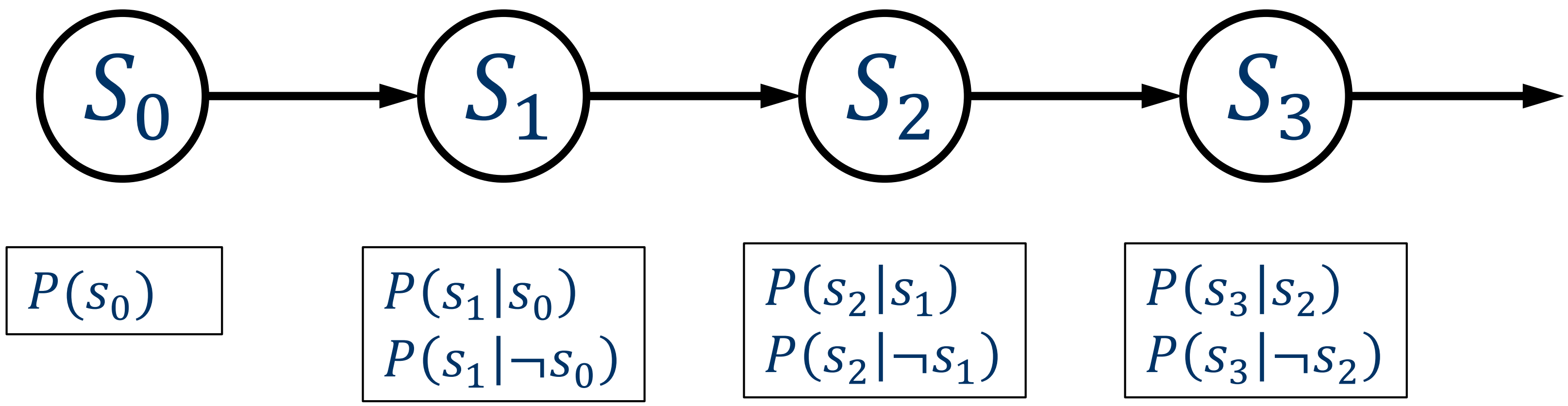
$E_t = 1$ means the director carries an umbrella on day t, and $E_t = 0$ otherwise.

Transition Model

The Markov assumption:

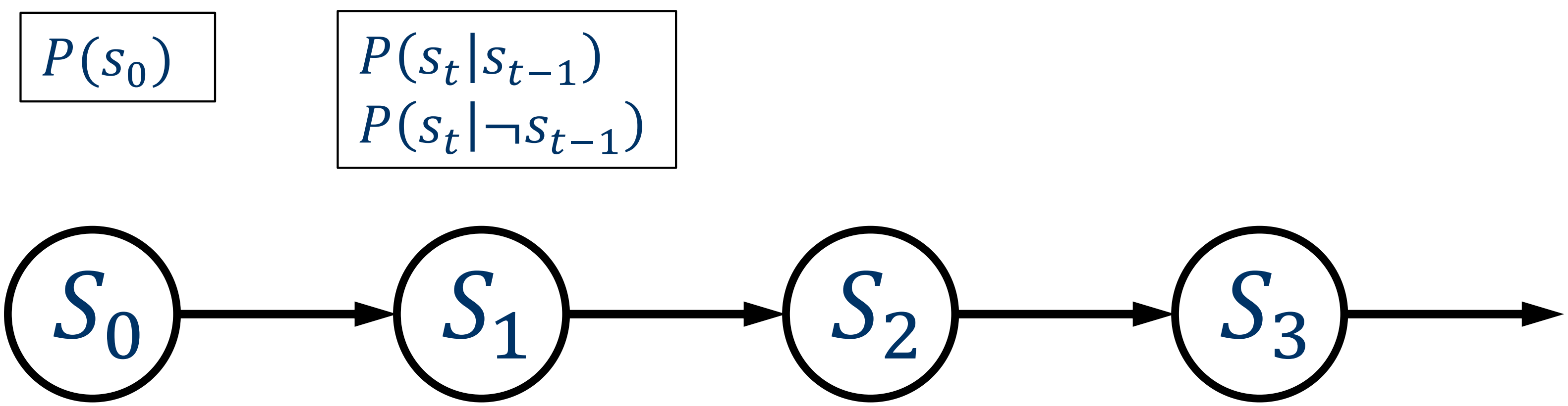
The future is independent of the past given the present.

$$P(S_t|S_0 \wedge \cdots \wedge S_{t-1}) = P(S_t|S_{t-1})$$



Stationary Process

- The dynamics does not change over time.
- $P(S_t|S_{t-1})$ is the same for every time step t.



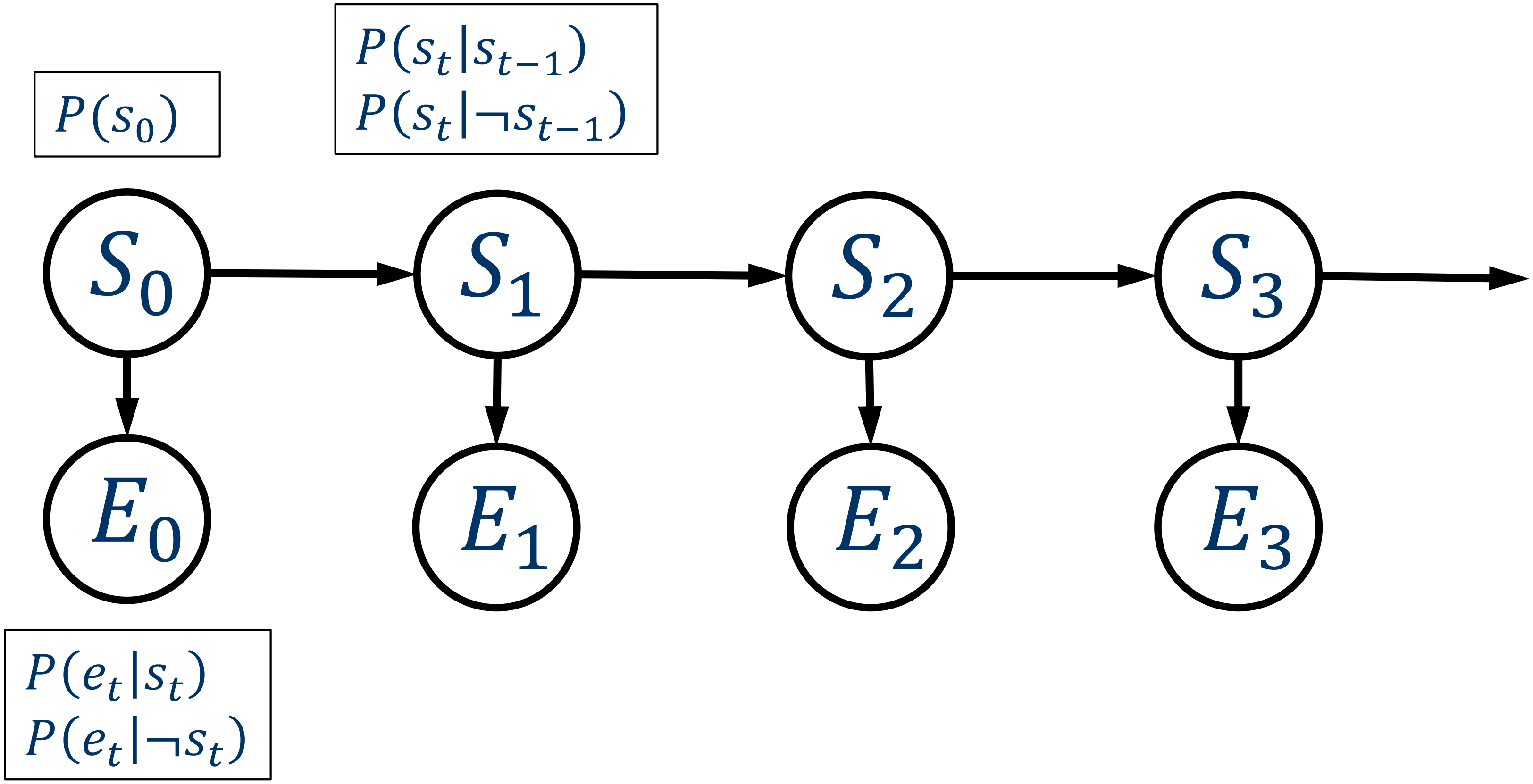
Advantages of a Stationary Process

- Simple to specify.
- Finite # of parameters to define an infinite network.

- In nature, the dynamics typically does not change.
- If the dynamics change, we can model it using another feature and incorporate it into the state.

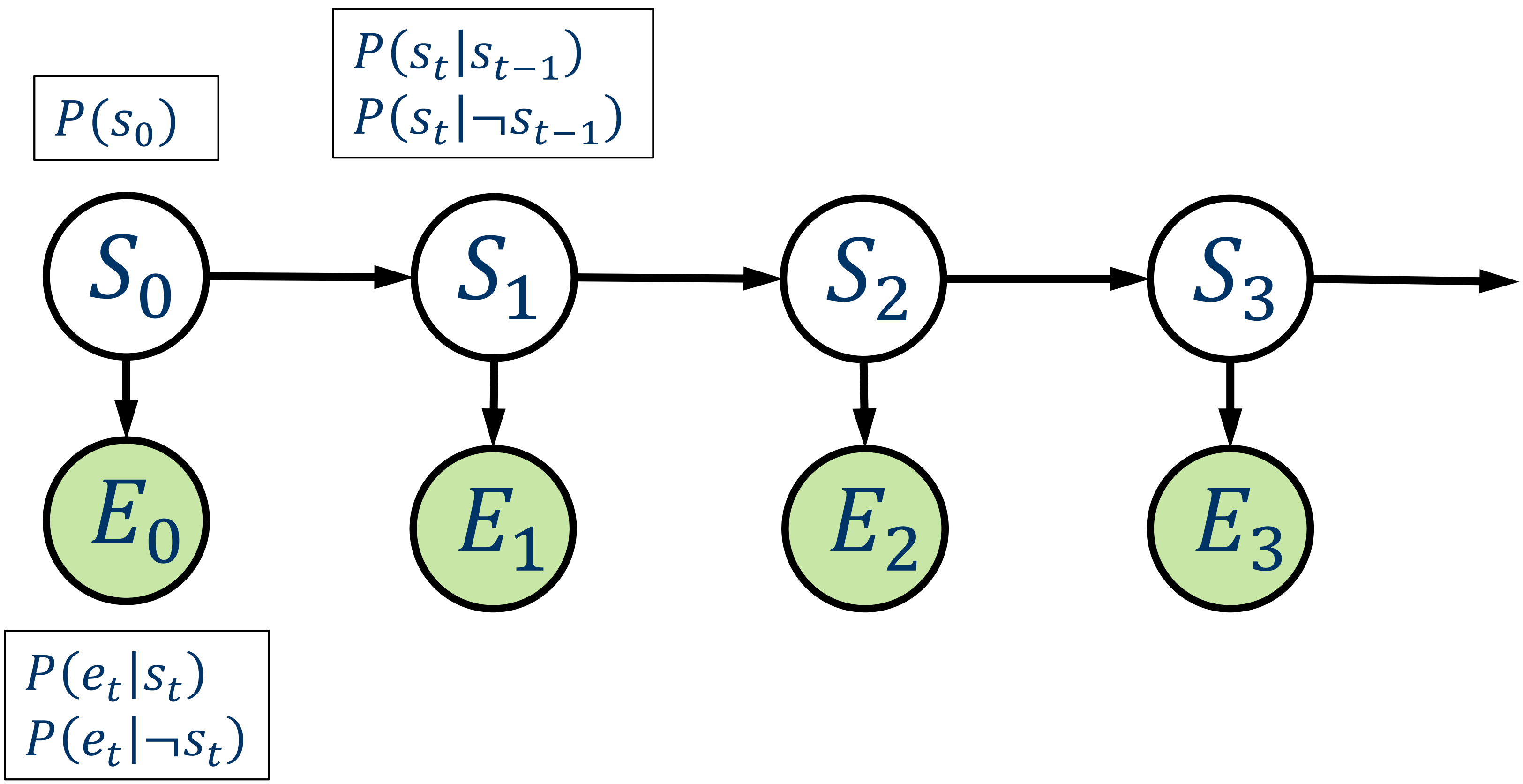
Observation Model

Each state is sufficient to generate its observation.



$$P(E_t|S_0 \wedge E_0 \ldots \wedge S_{t-1} \wedge E_{t-1} \wedge S_t) = P(E_t|S_t)$$

A Hidden Markov Model



COMMON INFERENCE TASKS

Common Inference Tasks

- Filtering: $P(S_t | e_0 \wedge \cdots \wedge e_t)$
  - Which hidden state am I in right now?

- Prediction: $P(S_{t+k} | e_0 \wedge \cdots \wedge e_t)$
  - Which hidden state will I be in on a future date?

- Smoothing: $P(S_{t-k} | e_0 \wedge \cdots \wedge e_t)$
  - Which hidden state was I in on a past date?

- Most likely explanation:
  - Which sequence of hidden state is most likely to have generated the observations so far?

Common Inference Tasks

- Filtering: $P(S_t | e_0 \wedge \cdots \wedge e_t)$
  - Which hidden state am I in right now?
- Prediction: $P(S_{t+k} | e_0 \wedge \cdots \wedge e_t)$
  - Which hidden state will I be in on a future date?
- Smoothing: $P(S_{t-k} | e_0 \wedge \cdots \wedge e_t)$
  - Which hidden state was I in on a past date?
- Most likely explanation:
  - Which sequence of hidden state is most likely to have generated the observations so far?

Performing Inference Tasks

- An HMM is a Bayesian network.
- Calculate any probability using the variable elimination algorithm!

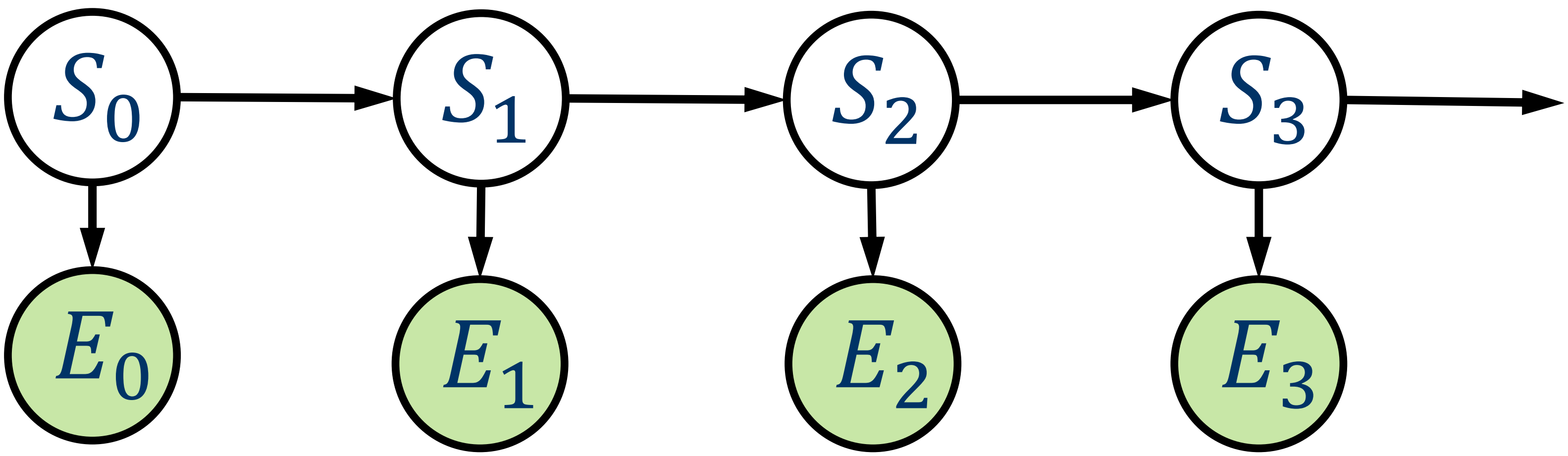
- But there are more efficient algorithms for HMM.
  - The forward-backward algorithm: filtering and smoothing.
  - The Viterbi algorithm: most likely explanation.

FILTERING

HMM with Some Numbers

$P(s_0) = 0.5$

$$P(s_t | s_{t-1}) = 0.7$$
$$P(s_t | \neg s_{t-1}) = 0.4$$



$$P(e_t | s_t) = 0.9$$
$$P(e_t | \neg s_t) = 0.2$$

Filtering

- Goal is to calculate $P(S_t | e_0 \wedge \cdots \wedge e_t)$ for every t.
- Recursive algorithm

- Base case: calculate $P(S_0 | e_0)$.
- Recursive case:

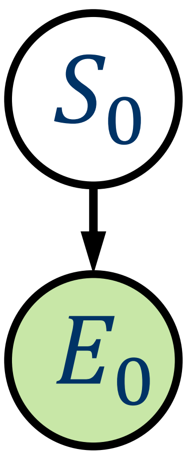
  Use $P(S_t | e_0 \wedge \cdots \wedge e_t)$ to calculate $P(S_{t+1} | e_0 \wedge \cdots \wedge e_{t+1})$

FILTERING BASE CASE CALCULATIONS

Filtering for $t = 0$ (base case)

Calculate $P(S_0|e_0)$.

$P(s_0) = 0.5$

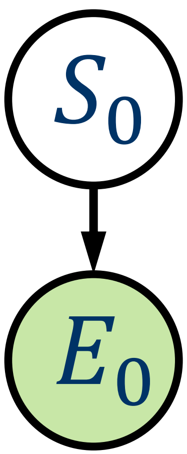


$P(e_t|s_t) = 0.9$
$P(e_t|\neg s_t) = 0.2$

Filtering for $t = 0$ (base case)

Calculate $P(S_0|e_0)$.

$$P(S_0|e_0) = \frac{P(S_0)P(e_0|S_0)}{P(e_0)} = \alpha P(S_0)P(e_0|S_0)$$



$P(s_0) = 0.5$

$P(e_t|s_t) = 0.9$
$P(e_t|\neg s_t) = 0.2$

Filtering for $t = 0$ (base case) calculations

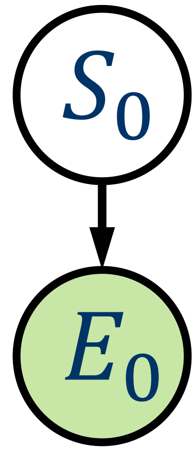
Calculate $P(S_0|e_0)$ (i.e., the director brought an umbrella on day 0)

$P(S_0|e_0) = \alpha P(S_0)P(e_0|S_0)$

Filtering for $t = 0$ (base case) calculations

Calculate $P(S_0|e_0)$ (i.e., the director brought an umbrella on day 0)

$P(S_0|e_0) = \alpha P(S_0)P(e_0|S_0)$



$P(s_0) = 0.5$

$P(e_t|s_t) = 0.9$
$P(e_t|\neg s_t) = 0.2$

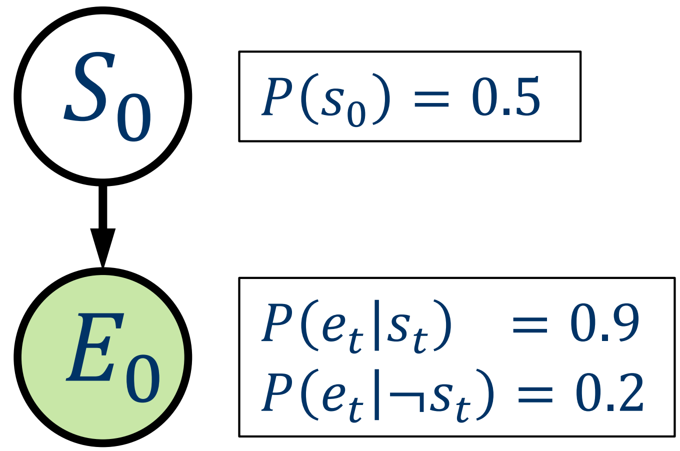
1. Multiply the probabilities.

$$P(s_0)P(e_0|s_0) = 0.5 * 0.9 = 0.45$$
$$P(\neg s_0)P(e_0|\neg s_0) = 0.5 * 0.2 = 0.1$$

Filtering for $t = 0$ (base case) calculations

Calculate $P(S_0|e_0)$ (i.e., the director brought an umbrella on day 0)

$P(S_0|e_0) = \alpha P(S_0)P(e_0|S_0)$



1. Multiply the probabilities.

$$P(s_0)P(e_0|s_0) = 0.5 * 0.9 = 0.45$$
$$P(\neg s_0)P(e_0|\neg s_0) = 0.5 * 0.2 = 0.1$$

2. Normalize the probabilities.

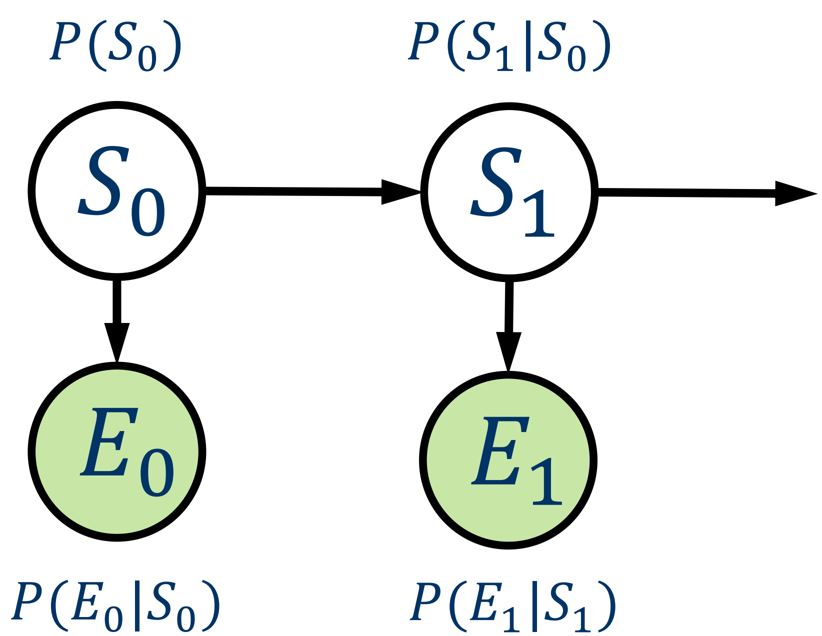
$$P(s_0|e_0) = \alpha P(s_0)P(e_0|s_0) = \frac{0.45}{0.45+0.1} = 0.818$$
$$P(\neg s_0|e_0) = \alpha P(\neg s_0)P(e_0|\neg s_0) = 1 - 0.818 = 0.182$$

FILTERING RECURSIVE CASE DERIVATION

Filtering for $t = 1$ (recursive case)

Calculate $P(S_1|e_0 \wedge e_1)$ using variable elimination algorithm.

$$P(S_1|e_0 \wedge e_1) = \overset{4}{\alpha}\, \underset{1}{P(e_1|S_1)} \overset{3}{} \overset{2b}{\sum_{S_0}} P(S_0) \underset{1}{\overset{2a}{P(e_0|S_0)}}\, P(S_1|S_0)$$



1) Restrict

2) Eliminate hidden variables

2a) Multiply

2b) Sum out

3) Multiply

4) Normalize

Filtering for $t = 1$ (recursive case)

Calculate $P(S_1|e_0 \wedge e_1)$ using $P(S_0|e_0)$.

Prob of rain on day 1 given that we saw umbrella on days 0 & 1

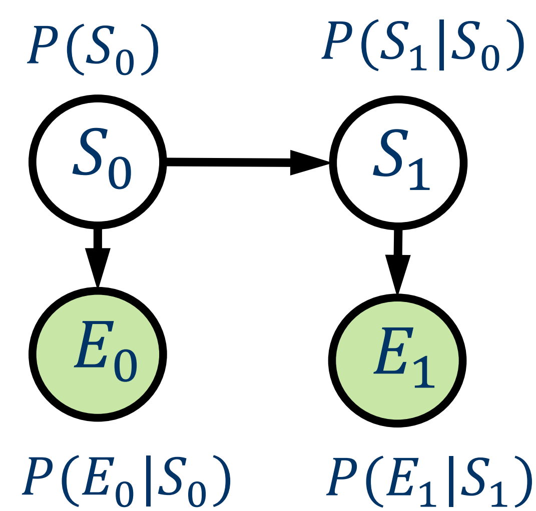
$$P(S_1|e_0 \wedge e_1) = \alpha_1 P(e_1|S_1) \sum_{S_0} P(S_0)P(e_0|S_0) P(S_1|S_0)$$

$$\frac{\alpha_1}{\alpha_0}$$

Prob of rain on day 0

given that we saw umbrella on day 0

$$P(S_0|e_0) = \alpha_0 \underbrace{P(S_0)P(e_0|S_0)}_{= \frac{P(S_0|e_0)}{\alpha_0}}$$



Filtering for $t = 1$ (recursive case)

Calculate $P(S_1|e_0 \wedge e_1)$ using $P(S_0|e_0)$.

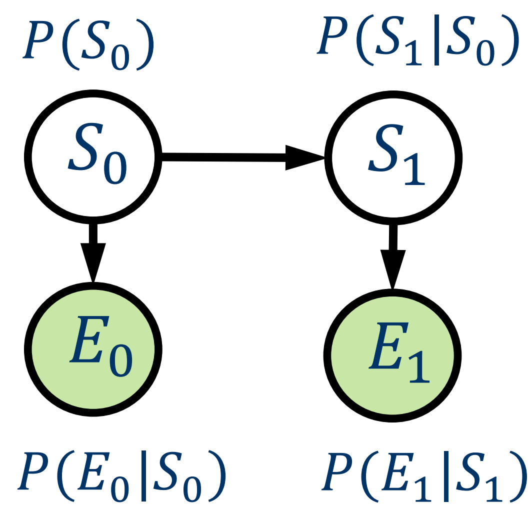
Prob of rain on day 1 given that we saw umbrella on days 0 & 1

$$P(S_1|e_0 \wedge e_1) = \alpha\, P(e_1|S_1) \sum_{S_0} P(S_0|e_0) P(S_1|S_0)$$

Prob of rain on day 0

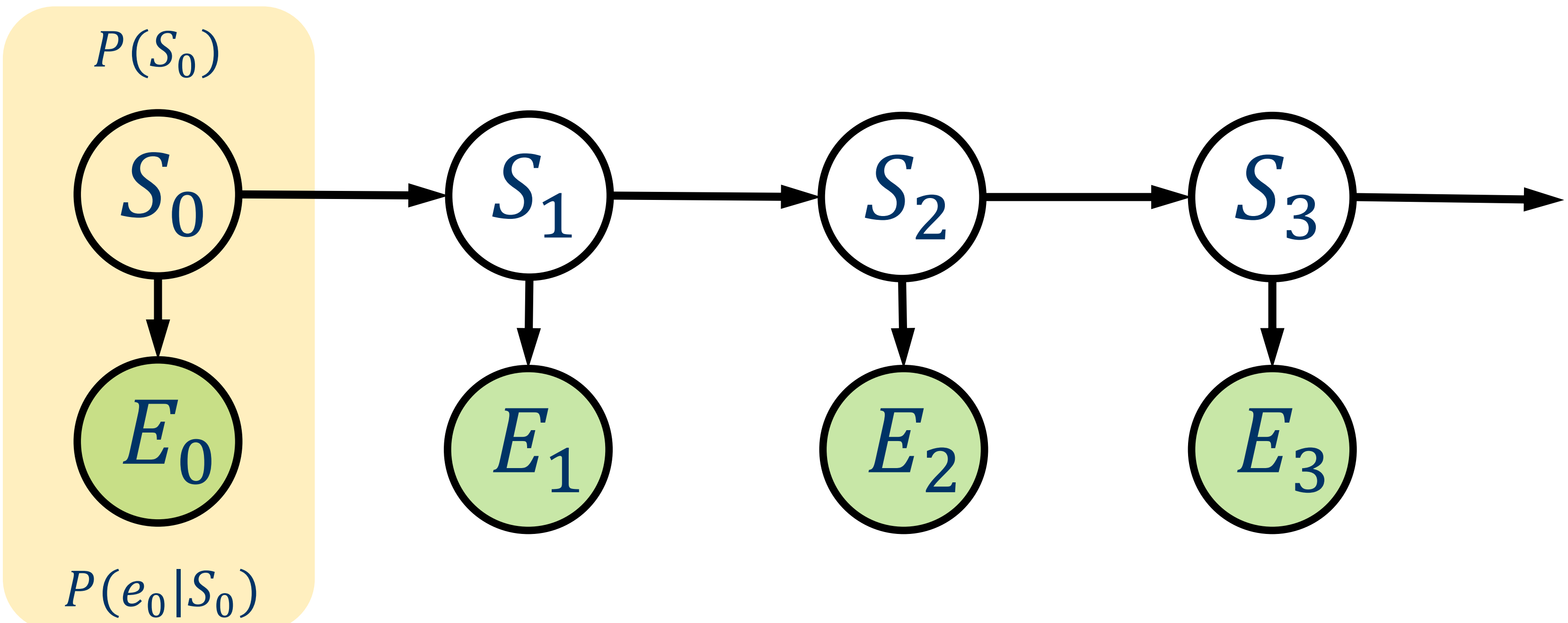
given that we saw umbrella on day 0

$$P(S_0|e_0) = \alpha P(S_0) P(e_0|S_0)$$



Filtering $t = 0$

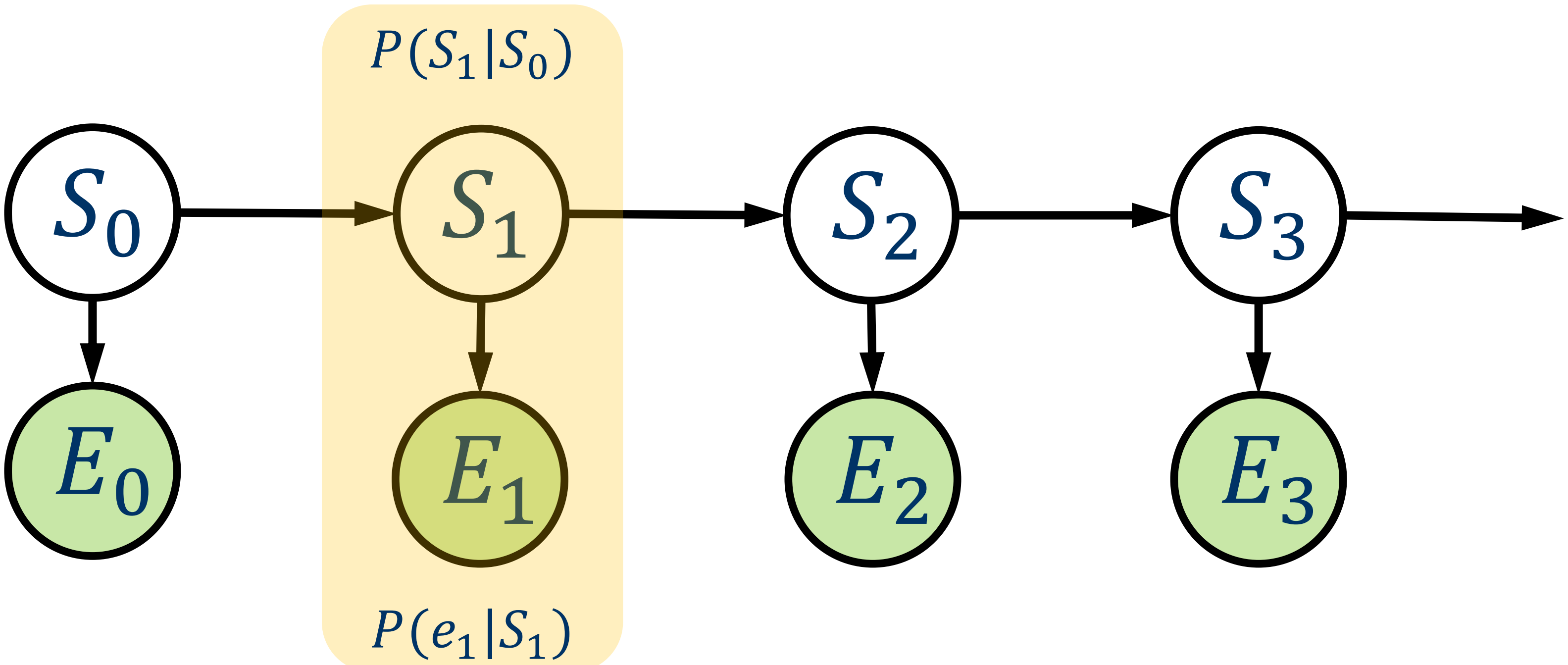
$$\boldsymbol{P(S_0|e_0)}$$
$$\boldsymbol{= \alpha\, P(S_0)P(e_0|S_0)}$$



Filtering $t = 0 \rightarrow t = 1$

$$\boldsymbol{P(S_1 | e_0 \wedge e_1)}$$
$$\boldsymbol{= \alpha\, P(e_1 | S_1) \sum_{S_0} P(S_0 | e_0) P(S_1 | S_0)}$$

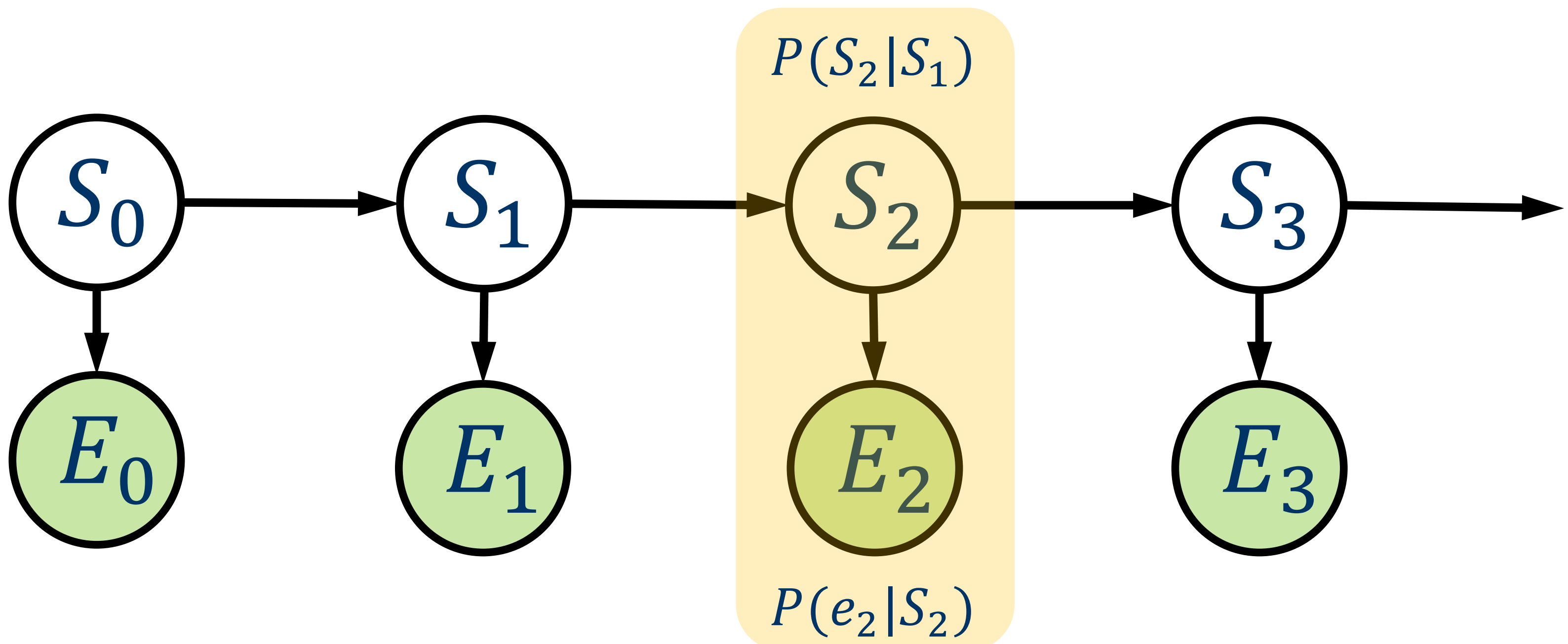
$\boldsymbol{P(S_0 | e_0)}$



Filtering $t = 1 \rightarrow t = 2$

$$\boldsymbol{P(S_2|e_0 \wedge e_1 \wedge e_2)}$$
$$\boldsymbol{= \alpha\, P(e_2|S_2) \textstyle\sum_{S_1} P(S_1|e_0 \wedge e_1) P(S_2|S_1)}$$

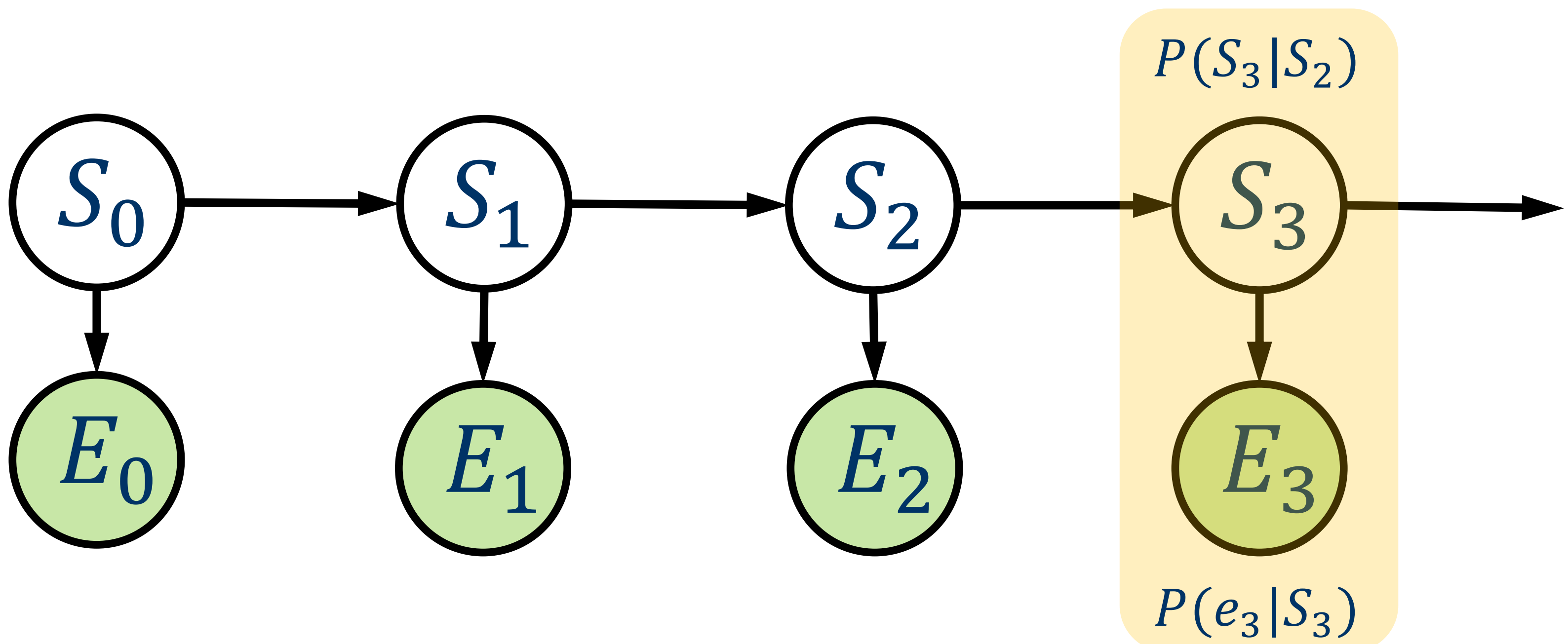
$\boldsymbol{P(S_0|e_0)}$ $\boldsymbol{P(S_1|e_0 \wedge e_1)}$



Filtering $t = 2 \to t = 3$

$$\boldsymbol{P(S_3 | e_0 \wedge \cdots \wedge e_3)}$$
$$\boldsymbol{= \alpha\, P(e_3 | S_3) \textstyle\sum_{S_2} P(S_2 | e_0 \wedge \cdots \wedge e_2) P(S_3 | S_2)}$$

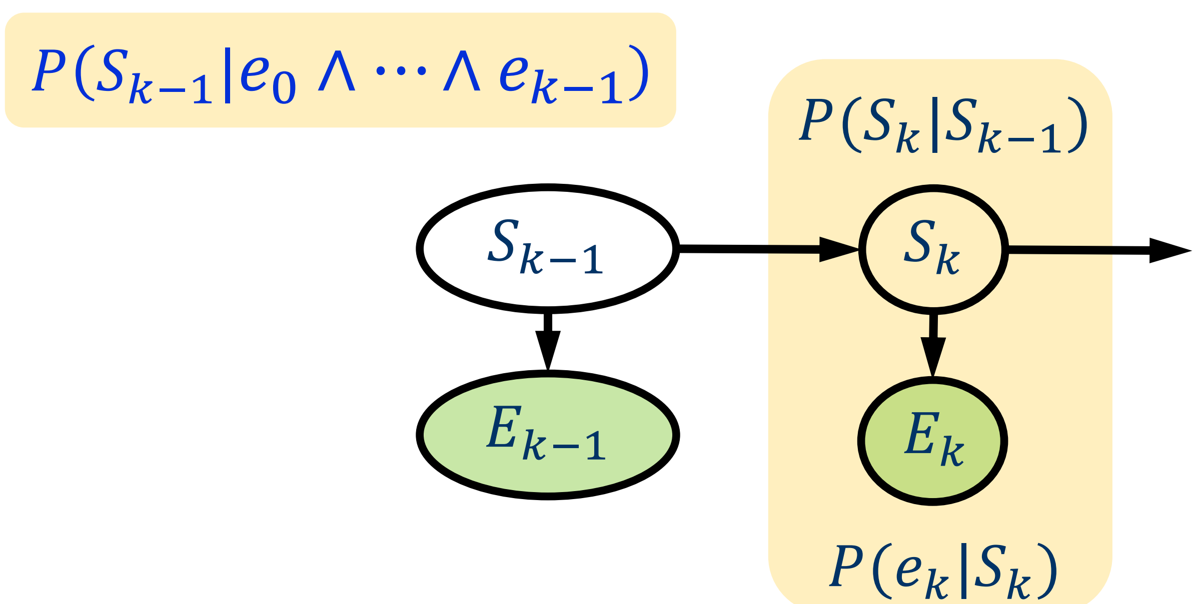
$\boldsymbol{P(S_0 | e_0)}$ $\boldsymbol{P(S_1 | e_0 \wedge e_1)}$ $\boldsymbol{P(S_2 | e_0 \wedge e_1 \wedge e_2)}$



Filtering $t = k - 1 \rightarrow t = k$

Calculate $P(S_k | e_0 \wedge \cdots \wedge e_k)$ given $P(S_{k-1} | e_0 \ldots \wedge e_{k-1})$

$$P(S_k | e_0 \wedge \cdots \wedge e_k)$$
$$= \alpha P(e_k | S_k) \sum_{S_{k-1}} P(S_{k-1} | e_0 \wedge \cdots \wedge e_{k-1}) P(S_k | S_{k-1})$$



Filtering Summary

Goal is to calculate $P(S_t|e_0 \wedge \cdots \wedge e_t)$ for every t.

Base case $(t = 0)$:

$$P(S_0|e_0) = \alpha P(S_0)P(e_0|S_0)$$

Recursive case $(t = k - 1 \rightarrow t = k)$:

$$P(S_k|e_0 \wedge \cdots \wedge e_k)$$
$$= \alpha P(e_k|S_k) \sum_{S_{k-1}} P(S_{k-1}|e_0 \wedge \cdots \wedge e_{k-1}) P(S_k|S_{k-1})$$

FILTERING RECURSIVE CASE CALCULATIONS

Filtering for $t = 1$ calculations

Calculate $P(S_1|e_0 \wedge e_1)$ (brought umbrella on days 0 & 1)

Given that $P(S_0|e_0) = (0.818, 0.182)$

$$\boldsymbol{P(S_1|e_0 \wedge e_1) = \alpha\, P(e_1|S_1) \sum_{S_0} P(S_0|e_0) P(S_1|S_0)}$$

1. $P(S_1|e_0) = \sum_{S_0} P(S_0|e_0) P(S_1|S_0)$
2. $P(S_1|e_0 \wedge e_1) = \alpha P(e_1|S_1) P(S_1|e_0)$

$P(s_0) = 0.5$

$P(s_t|s_{t-1}) = 0.7$
$P(s_t|\neg s_{t-1}) = 0.4$

$P(e_t|s_t) = 0.9$
$P(e_t|\neg s_t) = 0.2$

Filtering for $t = 1$ calculations

Calculate $P(S_1 | e_0 \wedge e_1)$ (brought umbrella on days 0 & 1)
Given that $P(S_0 | e_0) = (0.818, 0.182)$

1. Calculate $P(S_1 | e_0) = \sum_{S_0} P(S_0 | e_0) P(S_1 | S_0)$.

$$P(s_1 | e_0) = P(s_0 | e_0) P(s_1 | s_0) + P(\neg s_0 | e_0) P(s_1 | \neg s_0)$$
$$= 0.818 * 0.7 + 0.182 * 0.4 = 0.6454$$

$$P(\neg s_1 | e_0) = P(s_0 | e_0) P(\neg s_1 | s_0) + P(\neg s_0 | e_0) P(\neg s_1 | \neg s_0)$$
$$= 0.818 * 0.3 + 0.182 * 0.6 = 0.3546$$

$P(s_0) = 0.5$

$P(s_t | s_{t-1}) = 0.7$
$P(s_t | \neg s_{t-1}) = 0.4$

$P(e_t | s_t) = 0.9$
$P(e_t | \neg s_t) = 0.2$

Filtering for $t = 1$ calculations

Calculate $P(S_1|e_0 \wedge e_1)$ (brought umbrella on days 0 & 1)
Given that $P(S_0|e_0) = (0.818, 0.182)$

2. Calculate $P(S_1|e_0 \wedge e_1) = \alpha P(e_1|S_1)P(S_1|e_0)$
Given that $P(S_1|e_0) = (0.6454, 0.3546)$

$$P(e_1|s_1)P(s_1|e_0) = 0.9 * 0.6454 = 0.5809$$
$$P(e_1|\neg s_1)P(\neg s_1|e_0) = 0.2 * 0.3546 = 0.0709$$

$$P(s_1|e_0 \wedge e_1) = \alpha P(e_1|s_1)P(s_1|e_0) = \frac{0.5809}{0.5809+0.0709} = 0.8912$$

$$P(\neg s_1|e_0 \wedge e_1) = 1 - P(s_1|e_0 \wedge e_1) = 0.1088$$

$P(s_0) = 0.5$

$P(s_t|s_{t-1}) = 0.7$
$P(s_t|\neg s_{t-1}) = 0.4$

$P(e_t|s_t) = 0.9$
$P(e_t|\neg s_t) = 0.2$

Revisiting the Learning Goals

By the end of this lecture, you should be able to

1. Explain the independence assumptions in a Hidden Markov Model.
2. Calculate the filtering probabilities for a time step in a Hidden Markov Model.