CSC 384
Introduction to Artificial Intelligence

Uncertainty 6
The Viterbi Algorithm

Alice Gao and Randy Hickey

Winter 2023

Learning Goals

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

1. Derive the most likely sequence of hidden states given a sequence of observations by executing the Viterbi algorithm.

Outline

The Umbrella Story

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

You want to know whether it is raining today.

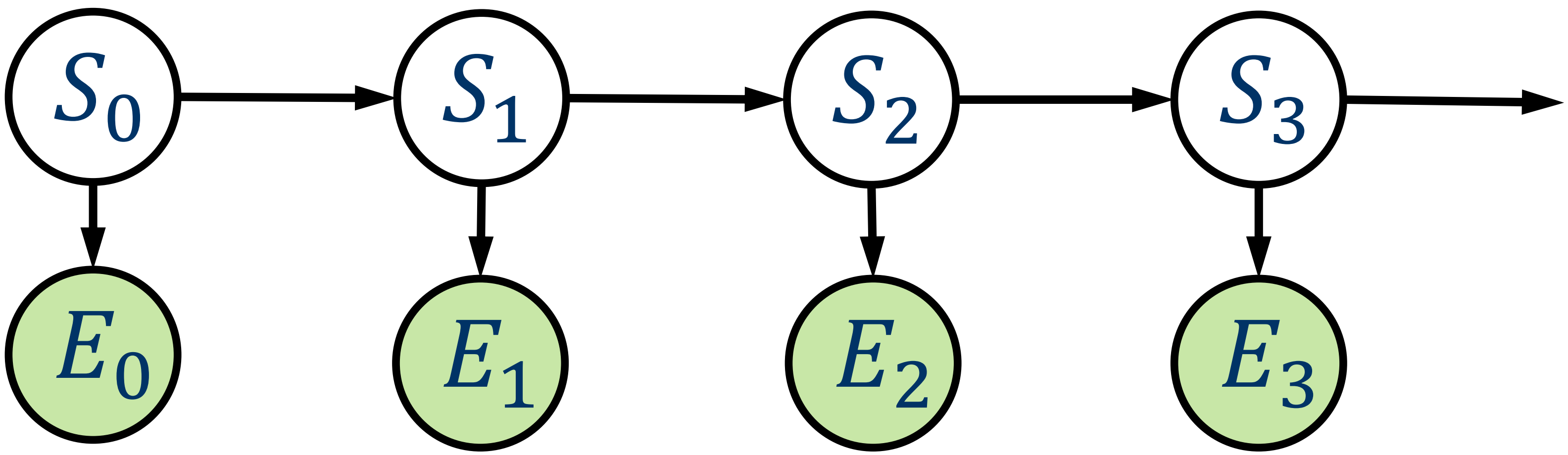
Unfortunately, your only access to the outside world occurs each morning when you see the director coming in with, or without, an umbrella.

HMM with the Umbrella Story

$P(s_0) = 0.5$

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

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



A VITERBI EXAMPLE FOR DAYS 0 AND 1

Viterbi Example

The first four observations are $e_0, \neg e_1, e_2, e_3$ (or umbrella, no, umbrella, umbrella).

What is the most likely sequence of hidden states that generated these four observations?

Viterbi for $t = 0$

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

$$P(s_0|e_0) = 0.818, P(\neg s_0|e_0) = 0.182$$

Given that we saw umbrella on day 0,

the most likely explanation is it rained on day 0.

e_0

$\boxed{s_0}$

0.818

$\boxed{\neg s_0}$

0.182

Sequence	Probability
s_0	0.818
$\neg s_0$	0.182

Viterbi for $t = 1$

$P(S_0 \wedge S_1 | e_0 \wedge \neg e_1)$

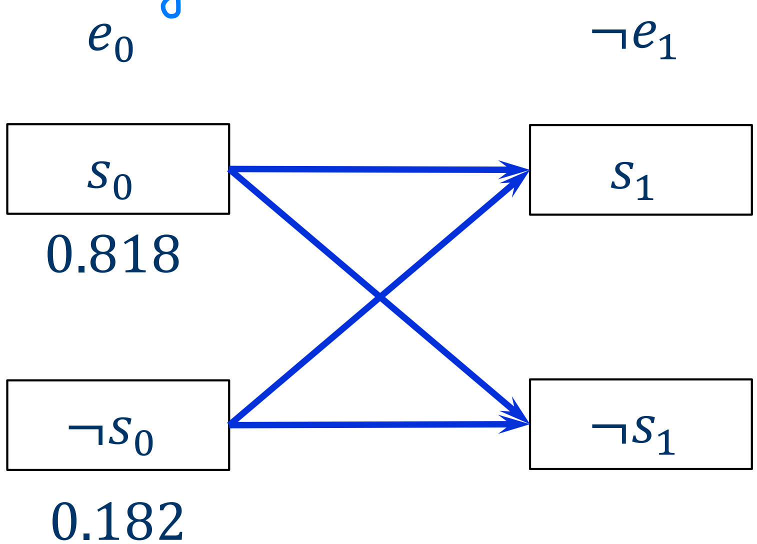
$= \alpha\, P(S_0) P(e_0|S_0)\, P(S_1|S_0) P(\neg e_1|S_1)$

$= \alpha\, P(S_0|e_0)\, P(S_1|S_0) P(\neg e_1|S_1)$

probs from day 0

transition probs

observation probs for day 1



Sequence	Probability
$s_0 \to s_1$	
$\neg s_0 \to s_1$	
$s_0 \to \neg s_1$	
$\neg s_0 \to \neg s_1$	

Viterbi for $t = 1$ (part 1)

$$P(S_0 \land S_1 | e_0 \land \neg e_1) = \alpha P(S_0 | e_0) P(S_1 | S_0) P(\neg e_1 | \mathrm{S}_1)$$

Suppose that it rained on day 1 (s_1):

- If it rained on day 0 (s_0) (i.e., sequence $s_0 \rightarrow s_1$),

$P(s_0 \land s_1 | e_0 \land \neg e_1) =$

- If it didn't rain on day 0 ($\neg s_0$) (i.e., sequence $\neg s_0 \rightarrow s_1$),

$P(\neg s_0 \land s_1 | e_0 \land \neg e_1) =$

Viterbi for $t = 1$ (part 1)

$$P(S_0 \wedge S_1 | e_0 \wedge \neg e_1) = \alpha P(S_0|e_0) P(S_1|S_0) P(\neg e_1|S_1)$$

Suppose that it rained on day 1 (s_1):

- If it rained on day 0 (s_0) (i.e., sequence $s_0 \rightarrow s_1$),

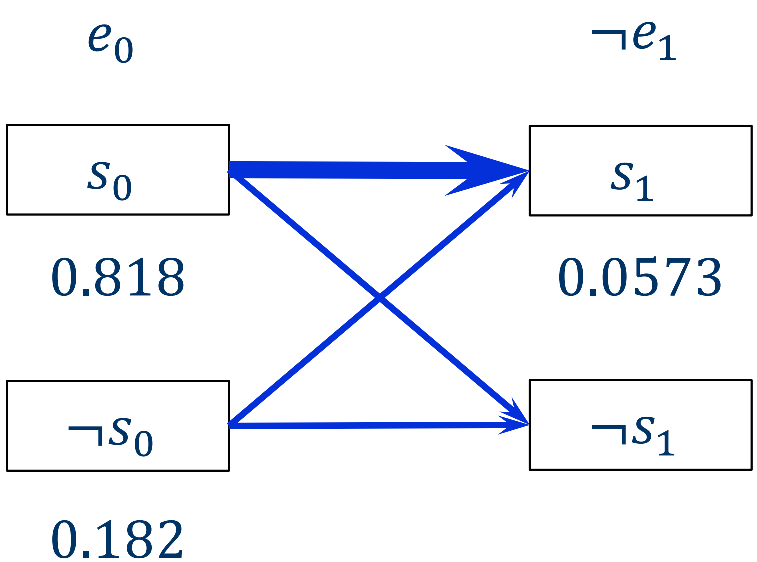
$$P(s_0 \wedge s_1 | e_0 \wedge \neg e_1) = P(s_0|e_0) P(s_1|s_0) P(\neg e_1|s_1)$$
$$= 0.818 * 0.7 * 0.1 = 0.0573$$

- If it didn't rain on day 0 ($\neg s_0$) (i.e., sequence $\neg s_0 \rightarrow s_1$),

$$P(\neg s_0 \wedge s_1 | e_0 \wedge \neg e_1) = P(\neg s_0|e_0) P(s_1|\neg s_0) P(\neg e_1|s_1)$$
$$= 0.182 * 0.4 * 0.1 = 0.0073$$

Viterbi for $t = 1$ Summary

$$P(S_0 \wedge S_1 | e_0 \wedge \neg e_1)$$
$$= \alpha P(S_0) P(e_0|S_0) P(S_1|S_0) P(\neg e_1|S_1)$$
$$= \alpha P(S_0|e_0) P(S_1|S_0) P(\neg e_1|S_1)$$



Sequence	Probability
$s_0 \rightarrow s_1$	0.0573
$\neg s_0 \rightarrow s_1$	0.0073
$s_0 \rightarrow \neg s_1$	
$\neg s_0 \rightarrow \neg s_1$	

Viterbi for $t = 1$ (part 2)

$P(S_0 \wedge S_1 | e_0 \wedge \neg e_1) = \alpha P(S_0 | e_0) P(S_1 | S_0) P(\neg e_1 | \mathrm{S}_1)$

Suppose that it didn't rain on day 1 ($\neg s_1$):

- If it rained on day 0 (s_0) (i.e., sequence $s_0 \rightarrow \neg s_1$),

$P(s_0 \wedge \neg s_1 | e_0 \wedge \neg e_1) =$

- If it did not rain on day 0 ($\neg s_0$) (i.e., sequence $\neg s_0 \rightarrow \neg s_1$),

$P(\neg s_0 \wedge \neg s_1 | e_0 \wedge \neg e_1) =$

Viterbi for $t = 1$ (part 2)

$$P(S_0 \wedge S_1 | e_0 \wedge \neg e_1) = \alpha P(S_0|e_0) P(S_1|S_0) P(\neg e_1|S_1)$$

Suppose that it didn't rain on day 1 ($\neg s_1$):

- If it rained on day 0 (s_0) (i.e., sequence $s_0 \rightarrow \neg s_1$),

$$P(s_0 \wedge \neg s_1 | e_0 \wedge \neg e_1) = P(s_0|e_0) P(\neg s_1|s_0) P(\neg e_1|\neg s_1)$$
$$= 0.818 * 0.3 * 0.8 = 0.1963$$

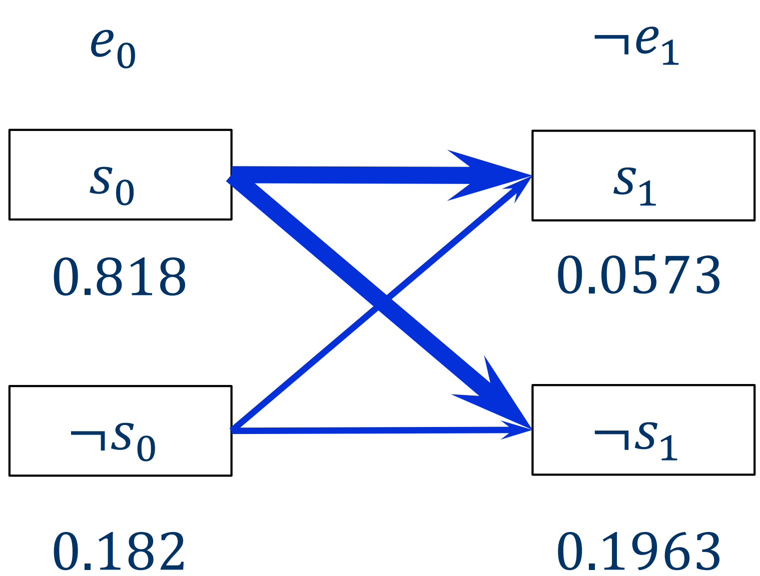
- If it did not rain on day 0 ($\neg s_0$) (i.e., sequence $\neg s_0 \rightarrow \neg s_1$),

$$P(\neg s_0 \wedge \neg s_1 | e_0 \wedge \neg e_1) = P(\neg s_0|e_0) P(\neg s_1|\neg s_0) P(\neg e_1|\neg s_1)$$
$$= 0.182 * 0.6 * 0.8 = 0.0874$$

Viterbi for $t = 1$ Summary

$$P(S_0 \land S_1 | e_0 \land \neg e_1) = \alpha P(S_0|e_0) P(S_1|S_0) P(\neg e_1|S_1)$$

Most likely sequence so far: $s_0 \rightarrow \neg s_1$



Sequence	Probability
$s_0 \rightarrow s_1$	0.0573
$\neg s_0 \rightarrow s_1$	0.0073
$s_0 \rightarrow \neg s_1$	0.1963
$\neg s_0 \rightarrow \neg s_1$	0.0874

THE VITERBI ALGORITHM

The Viterbi Algorithm

Which sequence of hidden state is
most likely to have generated the observations?

- Goal is to find the most likely sequence through the graph.

The likelihood of a sequence
= the initial probability of state on day 0 *
the transition probabilities *
the probabilities of the observations

The most likely sequence to a state on day k is

- the most likely sequence to some state on day k-1 followed by
- a transition to the state on day k.

① $P(S_0)$ ② $P(E_k|S_k)$ ③ $P(S_k|S_{k-1})$ HMM = ① + ② + ③

Viterbi Pseudocode

```
# I: initial probabilities. T: transition matrix.
# M: observation matrix.
Viterbi(E, S, I, T, M):
  prob = matrix(length(E),length(S))
  prev = matrix(length(E),length(S))

  # Determine values for time step 0
  for i in [0,…,length(S)-1]:
    prob[0,i] = I[i] * M[i,E[0]]
    prev[0,i] = None

  # For time steps 1 to length(E)-1,
  # find each current state's most likely prior state x.
  for t in [1,…,length(E)-1]:
    for i in [0,…,length(S)-1]:
      x = argmax_j in (prob[t-1,j] * T[j,i] * M[i,E[t]])
      prob[t,i] = prob[t-1,x] * T[x,i] * M[i,E[t]]
      prev[t,i] = x

  return prob, prev
```

$E = \{E_0, E_1, \dots, E_t\}$ set of observations over time steps.

S is set of hidden state values.

length(E) = # of time steps

Base case: $= P(S_0) * P(E_0|S_0)$

t: current time step

i: hidden state value for curr time t.

j: hidden state value for prev time t-1.

T[j,i]: $P(S_t|S_{t-1})$ transition probs from time t-1 to t.

Recursive case:

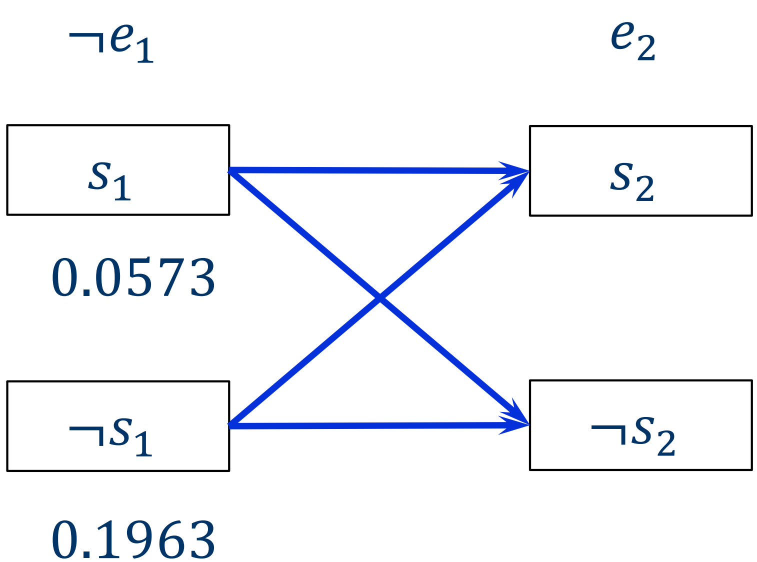
prob[t-1,x]: $P(S_0 \wedge \cdots \wedge S_{t-1} | E_0 \wedge \cdots \wedge E_{t-1})$ prob from prev time t-1

M[i,E[t]]: $P(E_t|S_t)$ observation prob for curr time t.

A VITERBI EXAMPLE FOR DAY 2

Viterbi for $t = 2$ Summary

$$P(S_0 \wedge S_1 \wedge S_2 | e_0 \wedge \neg e_1 \wedge e_2)$$
$$= \alpha \, P(S_0 \wedge S_1 | e_0 \wedge \neg e_1) \, P(S_2|S_1) P(e_2|S_2)$$



Sequence	Probability
$s_1 \rightarrow s_2$	
$\neg s_1 \rightarrow s_2$	
$s_1 \rightarrow \neg s_2$	
$\neg s_1 \rightarrow \neg s_2$	

Viterbi for $t = 2$ (part 1)

$P(S_0 \wedge S_1 \wedge S_2 | e_0 \wedge \neg e_1 \wedge e_2)$
$= \alpha P(S_0 \wedge S_1 | e_0 \wedge \neg e_1) P(S_2 | S_1) P(e_2 | S_2)$

Suppose that it rained on day 2 (s_2):

- If it rained on day 1 (s_1) (i.e., sequence $s_0 \to s_1 \to s_2$),

$P(s_0 \wedge s_1 \wedge s_2 | e_0 \wedge \neg e_1 \wedge e_2)$
=

- If it didn't rain on day 1 ($\neg s_1$) (i.e., sequence $s_0 \to \neg s_1 \to s_2$),

$P(s_0 \wedge \neg s_1 \wedge s_2 | e_0 \wedge \neg e_1 \wedge e_2)$
=

Viterbi for $t = 2$ (part 1)

$P(S_0 \land S_1 \land S_2 | e_0 \land \neg e_1 \land e_2)$
$= \alpha P(S_0 \land S_1 | e_0 \land \neg e_1) P(S_2 | S_1) P(e_2 | S_2)$

Suppose that it rained on day 2 (s_2):

- If it rained on day 1 (s_1) (i.e., sequence $s_0 \rightarrow s_1 \rightarrow s_2$),

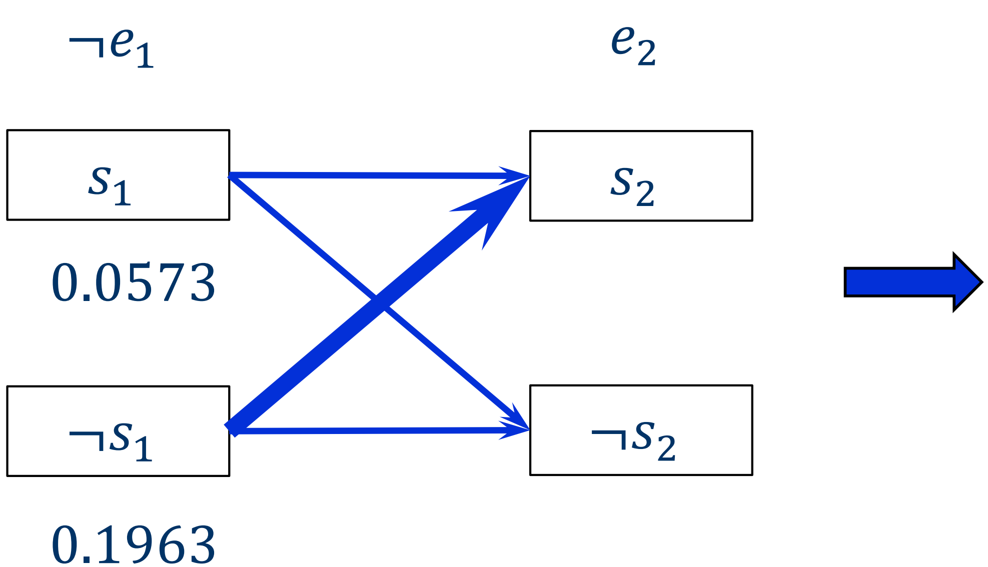
$P(s_0 \land s_1 \land s_2 | e_0 \land \neg e_1 \land e_2)$
$= P(s_0 \land s_1 | e_0 \land \neg e_1) P(s_2 | s_1) P(e_2 | s_2)$
$= 0.0573 * 0.7 * 0.9 = 0.0361$

- If it didn't rain on day 1 ($\neg s_1$) (i.e., sequence $s_0 \rightarrow \neg s_1 \rightarrow s_2$),

$P(s_0 \land \neg s_1 \land s_2 | e_0 \land \neg e_1 \land e_2)$
$= P(s_0 \land \neg s_1 | e_0 \land \neg e_1) P(s_2 | \neg s_1) P(e_2 | s_2)$
$= 0.1963 * 0.4 * 0.9 = 0.0707$

Viterbi for $t = 2$ Summary

$$P(S_0 \wedge S_1 \wedge S_2 | e_0 \wedge \neg e_1 \wedge e_2)$$
$$= \alpha\, P(S_0 \wedge S_1 | e_0 \wedge \neg e_1)\, P(S_2 | S_1) P(e_2 | S_2)$$



Sequence	Probability
$s_1 \to s_2$	0.0361
$\neg s_1 \to s_2$	0.0707
$s_1 \to \neg s_2$	
$\neg s_1 \to \neg s_2$	

Viterbi for $t = 2$ (part 2)

$P(S_0 \wedge S_1 \wedge S_2 | e_0 \wedge \neg e_1 \wedge e_2)$
$= \alpha P(S_0 \wedge S_1 | e_0 \wedge \neg e_1) P(S_2 | S_1) P(e_2 | S_2)$

Suppose that it didn't rain on day 2 (s_2):

- If it rained on day 1 (s_1) (i.e., sequence $s_0 \to s_1 \to \neg s_2$),

$P(s_0 \wedge s_1 \wedge \neg s_2 | e_0 \wedge \neg e_1 \wedge e_2)$
=

- If it didn't rain on day 1 ($\neg s_1$) (i.e., sequence $s_0 \to \neg s_1 \to \neg s_2$),

$P(s_0 \wedge \neg s_1 \wedge \neg s_2 | e_0 \wedge \neg e_1 \wedge e_2)$
=

Viterbi for $t = 2$ (part 2)

$P(S_0 \land S_1 \land S_2 | e_0 \land \neg e_1 \land e_2)$
$= \alpha P(S_0 \land S_1 | e_0 \land \neg e_1) P(S_2 | S_1) P(e_2 | S_2)$

Suppose that it didn't rain on day 2 (s_2):

- If it rained on day 1 (s_1) (i.e., sequence $s_0 \to s_1 \to \neg s_2$),

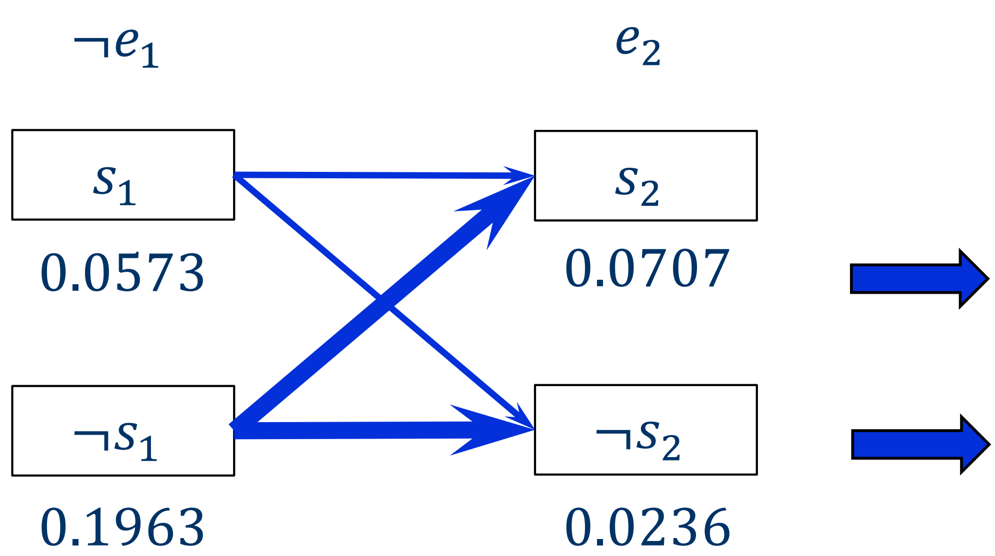
$P(s_0 \land s_1 \land \neg s_2 | e_0 \land \neg e_1 \land e_2)$
$= P(s_0 \land s_1 | e_0 \land \neg e_1) P(\neg s_2 | s_1) P(e_2 | \neg s_2)$
$= 0.0573 * 0.3 * 0.2 = 0.0034$

- If it didn't rain on day 1 ($\neg s_1$) (i.e., sequence $s_0 \to \neg s_1 \to \neg s_2$),

$P(s_0 \land \neg s_1 \land \neg s_2 | e_0 \land \neg e_1 \land e_2)$
$= P(s_0 \land \neg s_1 | e_0 \land \neg e_1) P(\neg s_2 | \neg s_1) P(e_2 | \neg s_2)$
$= 0.1963 * 0.6 * 0.2 = 0.0236$

Viterbi for $t = 2$ Summary

$$P(S_0 \wedge S_1 \wedge S_2 | e_0 \wedge \neg e_1 \wedge e_2)$$
$$= \alpha\, P(S_0 \wedge S_1 | e_0 \wedge \neg e_1)\, P(S_2 | S_1) P(e_2 | S_2)$$

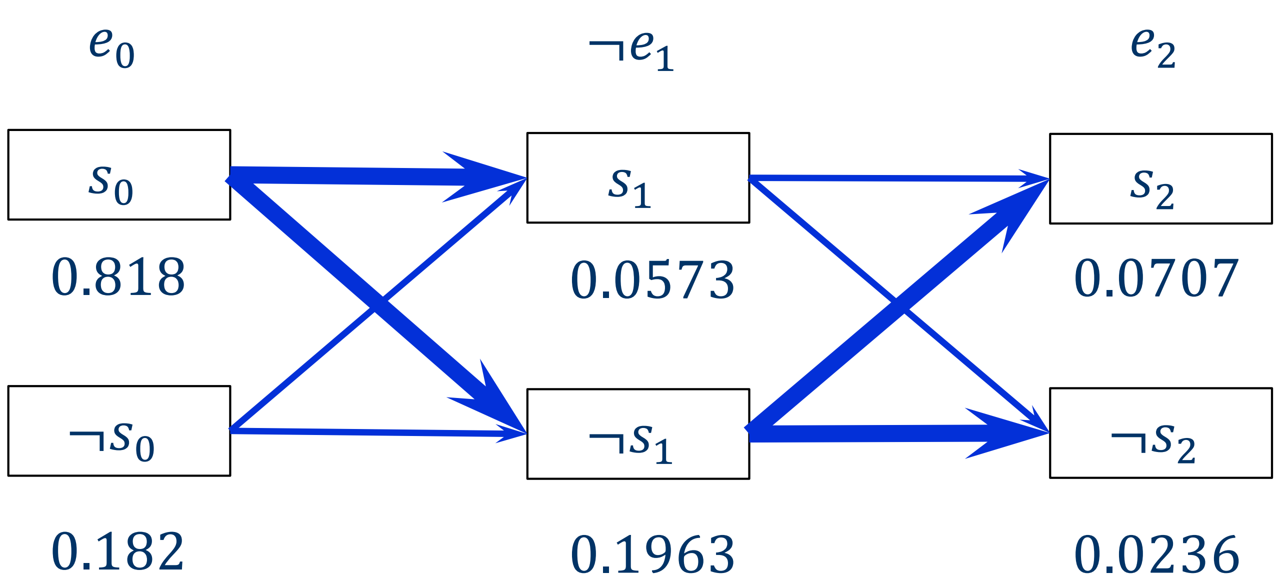


Sequence	Probability
$s_1 \to s_2$	0.0361
$\neg s_1 \to s_2$	0.0707
$s_1 \to \neg s_2$	0.0034
$\neg s_1 \to \neg s_2$	0.0236

Viterbi for $t = 2$ Summary

$P(S_0 \wedge S_1 \wedge S_2 | e_0 \wedge \neg e_1 \wedge e_2)$
$= \alpha\, P(S_0 \wedge S_1 | e_0 \wedge \neg e_1)\, P(S_2 | S_1) P(e_2 | S_2)$

Most likely sequence so far: $s_0 \rightarrow \neg s_1 \rightarrow s_2$



Revisiting the Learning Goals

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

1. Derive the most likely sequence of hidden states given a sequence of observations by executing the Viterbi algorithm.