



Viterbi Algorithm

CS 181 Fall 2021



Definition: Inputs and Outputs

- Inputs $\langle \Theta, \lambda = (n, m, A, B, \pi) \rangle$
 - Θ = a sequence of observations
 - λ = a Hidden Markov Model
 - n = the number of hidden states, $S = \{s_1, s_2, \dots, s_n\}$
 - m = the number of observation symbols, $V = \{v_1, v_2, \dots, v_m\}$
 - A = the transition probability distribution $A = \{a_{ij} = \mathbb{P}(q_{t+1} = s_j | q_t = s_i)\}$
 - The (i, j) -th entry is the probability that the HMM is in state j at time $t+1$, given that it was in state i at time t
 - B = the emission probability distribution $B = \{b_j(k) = \mathbb{P}(v_k \text{ at time } t | q_t = s_j)\}$
 - The (j, k) -th entry is the probability that the HMM emits symbol k at time t , given that it was in state j at time t
 - π = the initial state distribution $\pi = \{\pi_i = \mathbb{P}(q_1 = s_i)\}$
- Outputs: Q for which $\mathbb{P}(Q|\Theta)$ is maximal
 - The most likely sequence of states (and the probability of observing it)

Definition: Auxiliary Data Structures

- δ = Scoring matrix $\delta = \{\delta_t(i) = \text{MAX}_{q_1 \dots q_{t-1}} \{\mathbb{P}(\theta_1 \dots \theta_t, q_t = s_i)\}$
 - The (i, t) -th entry gives the maximum probability of observing $\theta_1, \dots, \theta_t$ along any sequence of states and ending in state i at time t
- ψ = Backtracking matrix
 - The (i, t) -th entry gives the state at time $t-1$ which produces the maximum probability path ending in state i at time t

The Algorithm

1) Initialize the matrices

2) Apply the recurrence relations to fill each matrix

3) Compute the maximum probability

4) Initialize the backtracking process

5) Complete the backtracking

6) Output p^* and q_1^*, \dots, q_T^*

```
/* Initialization
```

```
for  $i \leftarrow 1$  to  $n$  do
```

```
     $\delta_1(i) \leftarrow \pi_i b_i(\theta_1)$ 
```

```
     $\psi_1(i) \leftarrow 0$ 
```

```
end
```

```
/* Recurrence
```

```
for  $t \leftarrow 2$  to  $T$  do
```

```
    for  $j \leftarrow 1$  to  $n$  do
```

```
         $\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$ 
```

```
         $\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$ 
```

```
    end
```

```
end
```

```
/* Termination
```

```
 $p^* \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_T(i)]$ 
```

```
 $q_T^* \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_T(i)]$ 
```

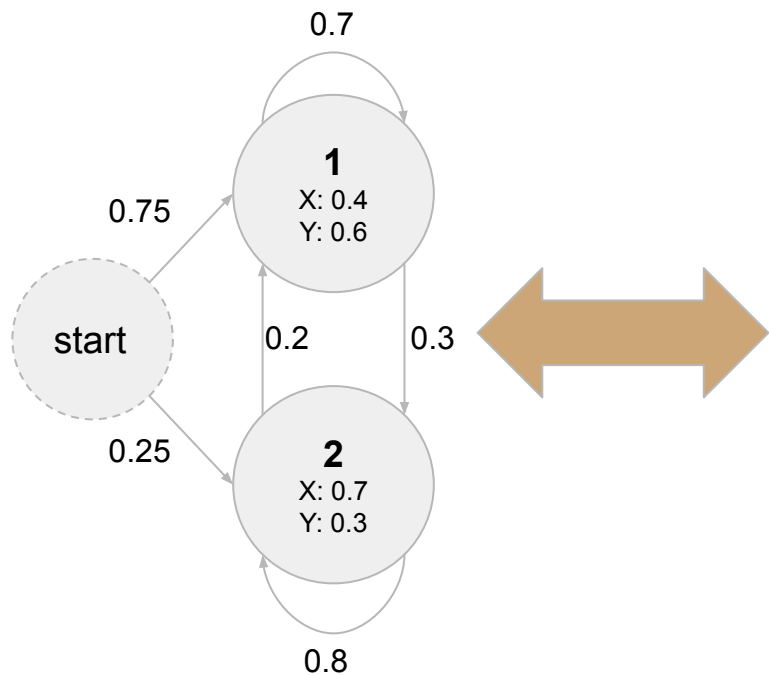
```
/* Backtracking
```

```
for  $t \leftarrow T - 1$  to 1 do
```

```
     $q_t^* \leftarrow \psi_{t+1}(q_{t+1}^*)$ 
```

```
end
```

An Example



Observation
sequence, θ :
XY YXX

π	1	2
$P(q_i)$	0.75	0.25

A	1	2
1	0.7	0.3
2	0.2	0.8

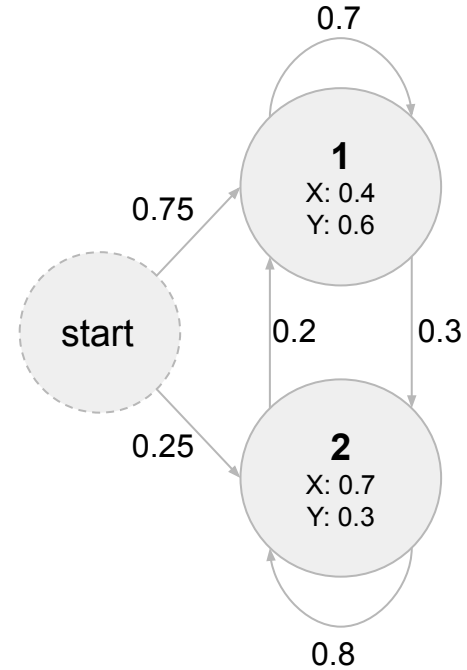
B	X	Y
1	0.4	0.6
2	0.7	0.3

XYYXX

δ	1	2
t = 1		
2		
3		
4		
5		

ψ	1	2
t = 1		
2		
3		
4		
5		

```
/* Initialization
for i ← 1 to n do
  |  $\delta_1(i) \leftarrow \pi_i b_i(\theta_1)$ 
  |  $\psi_1(i) \leftarrow 0$ 
end
```

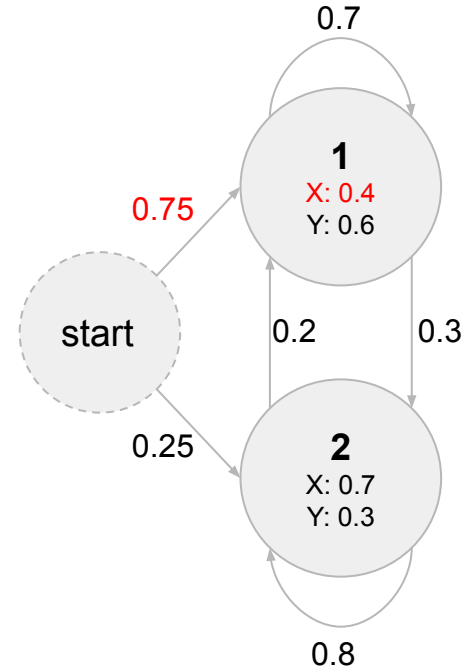


XYYXX

δ	1	2
t = 1	0.3	
2		
3		
4		
5		

ψ	1	2
t = 1		
2		
3		
4		
5		

```
/* Initialization
for i ← 1 to n do
  |  $\delta_1(i) \leftarrow \pi_i b_i(\theta_1)$ 
  |  $\psi_1(i) \leftarrow 0$ 
end
```

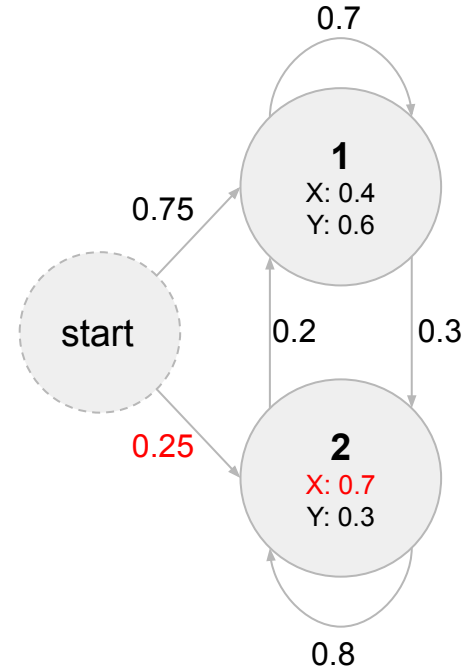


XYYXX

δ	1	2
t = 1	0.3	0.175
2		
3		
4		
5		

ψ	1	2
t = 1		
2		
3		
4		
5		

```
/* Initialization
for i ← 1 to n do
  |  $\delta_1(i) \leftarrow \pi_i b_i(\theta_1)$ 
  |  $\psi_1(i) \leftarrow 0$ 
end
```

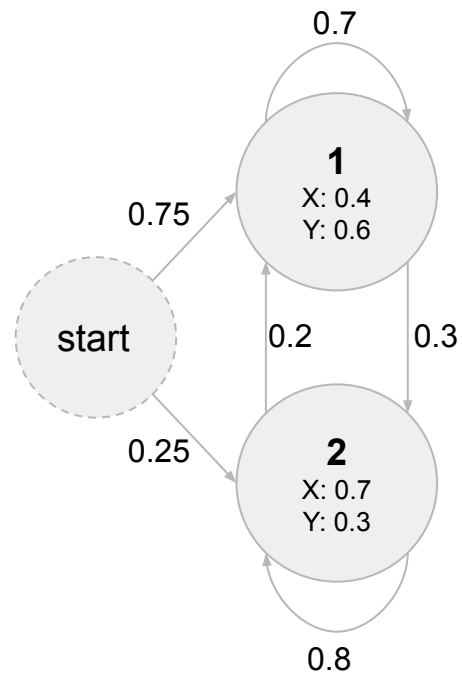


XYYXX

```
/* Initialization
for i ← 1 to n do
  |  $\delta_1(i) \leftarrow \pi_i b_i(\theta_1)$ 
  |  $\psi_1(i) \leftarrow 0$ 
end
```

δ	1	2
t = 1	0.3	0.175
2		
3		
4		
5		

ψ	1	2
t = 1	0	0
2		
3		
4		
5		



XYYXX

j=1: max(0.21, 0.035)

δ	1	2
t = 1	0.3	0.175
2		
3		
4		
5		

ψ	1	2
t = 1	0	0
2		
3		
4		
5		

/* Recurrence

for t ← 2 to T do

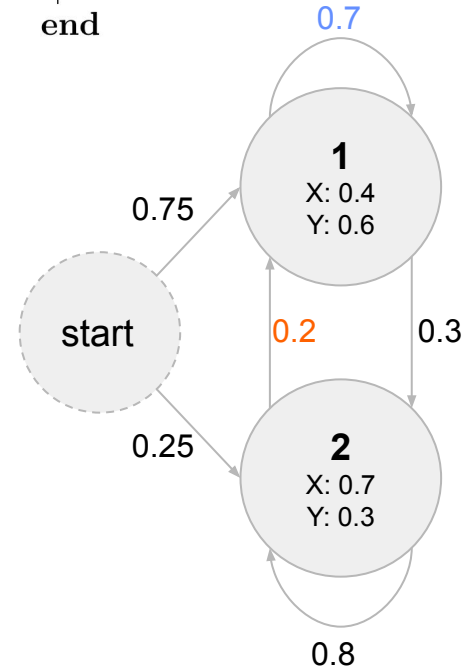
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



XYYXX

j=1: max(0.21, 0.035)

δ	1	2
t = 1	0.3	0.175
2	0.126	
3		
4		
5		

ψ	1	2
t = 1	0	0
2	1	
3		
4		
5		

/* Recurrence

for t ← 2 to T do

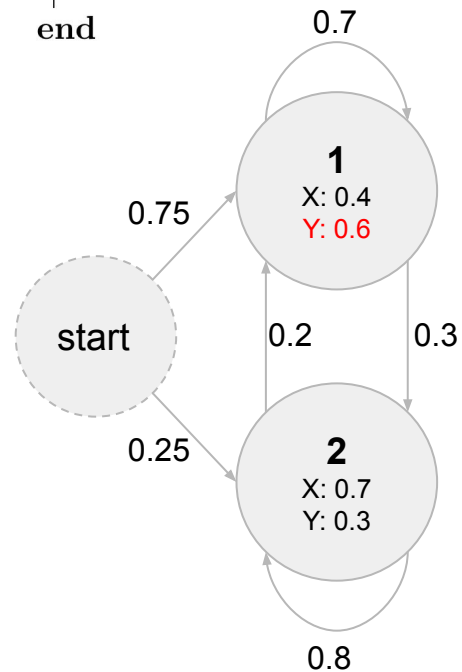
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



XYYXX

j=1: max(0.09, 0.14)

δ	1	2
t = 1	0.3	0.175
2	0.126	
3		
4		
5		

ψ	1	2
t = 1	0	0
2	1	
3		
4		
5		

/* Recurrence

for t ← 2 to T do

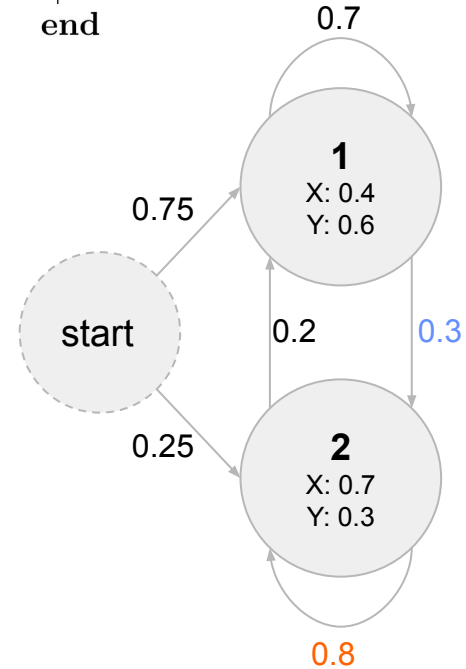
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



X**Y**YXX

j=1: max(0.09, **0.14**)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3		
4		
5		

ψ	1	2
t = 1	0	0
2	1	2
3		
4		
5		

/* Recurrence

for t ← 2 to T do

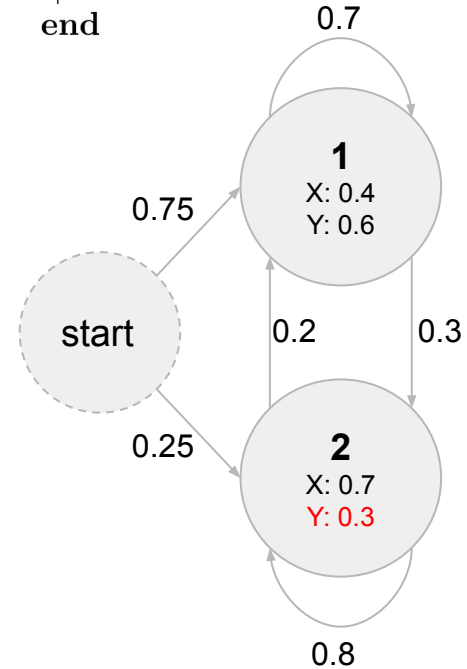
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



XYYXX

j=1: max(0.0882, 0.0084)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3		
4		
5		

ψ	1	2
t = 1	0	0
2	1	2
3		
4		
5		

```
/* Recurrence
```

```
for t ← 2 to T do
```

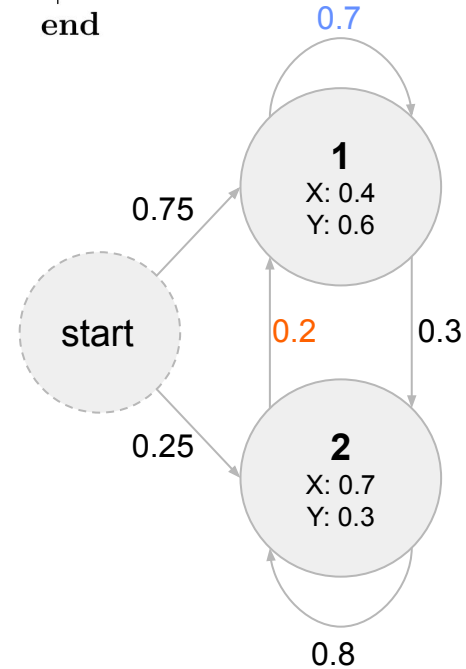
```
  for j ← 1 to n do
```

```
     $\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$ 
```

```
     $\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$ 
```

```
  end
```

```
end
```



XY^YXX

j=1: max(0.0882, 0.0084)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	
4		
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	
4		
5		

/* Recurrence

for t ← 2 to T do

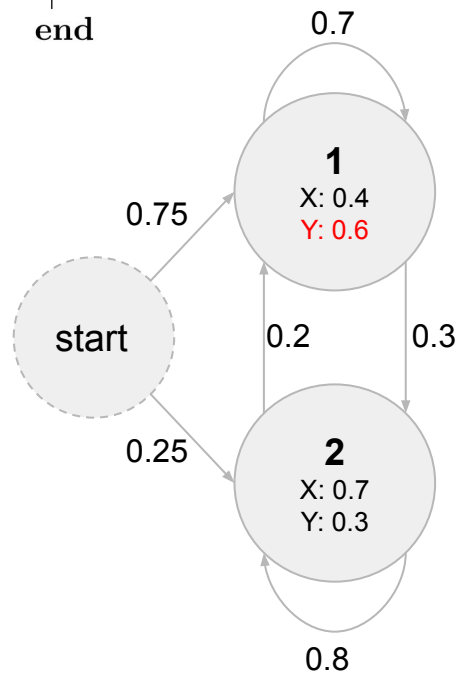
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



XY YXX

j=1: max(0.0378, 0.0336)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	
4		
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	
4		
5		

/* Recurrence

for t ← 2 to T do

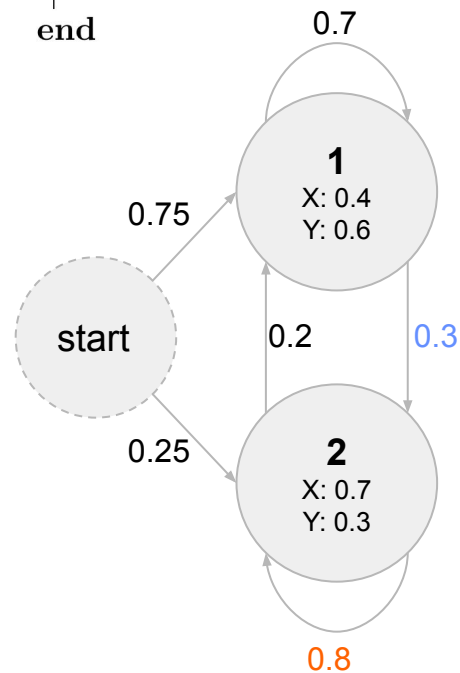
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



XY^YXX

j=1: max(0.0378, 0.0336)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4		
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4		
5		

/* Recurrence

for t ← 2 to T do

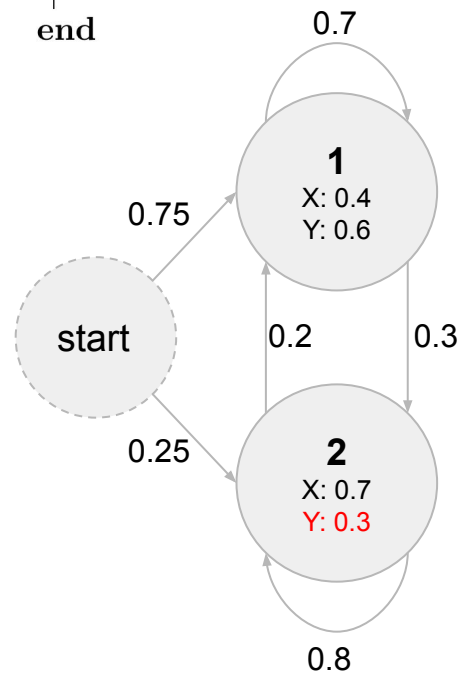
 for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

 end

end



XYYXX

j=1: max(0.0370, 0.0027)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4		
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4		
5		

/* Recurrence

for t ← 2 to T do

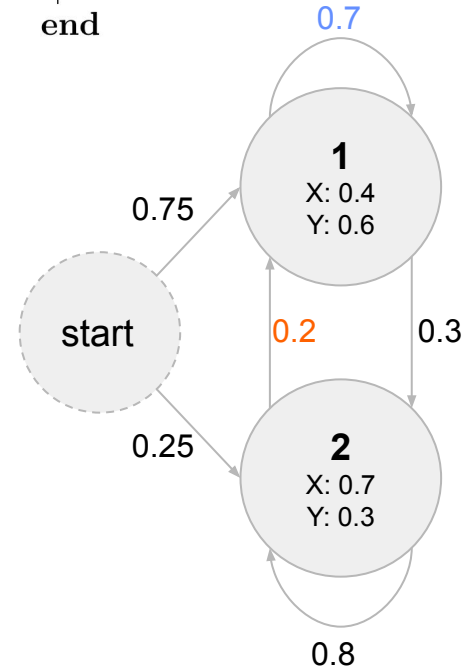
for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

end

end



XYYXX

j=1: max(0.0370, 0.0027)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	
5		

/* Recurrence

for t ← 2 to T do

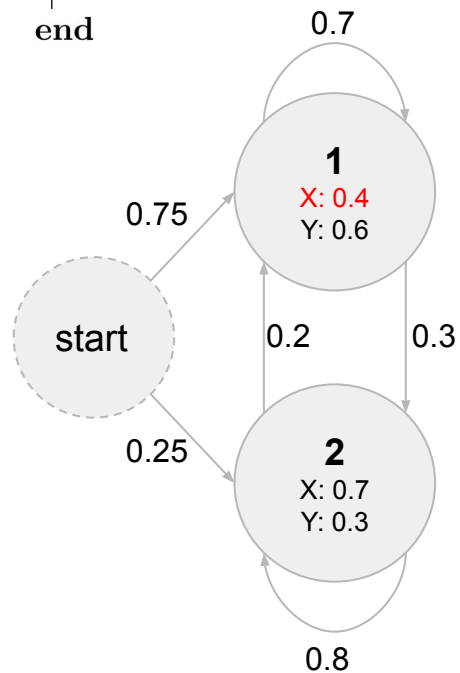
 for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

 end

end



XY YXX

j=1: max(0.0159, 0.0091)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	
5		

/* Recurrence

for t ← 2 to T do

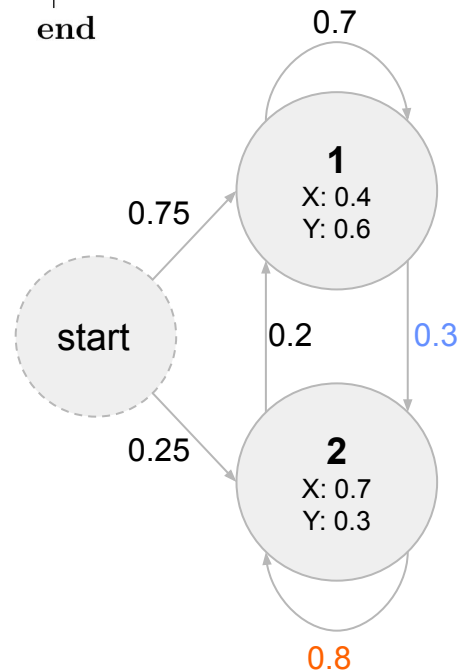
 for j ← 1 to n do

$$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$$

$$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$$

 end

end



XYYXX

j=1: max(0.0159, 0.0091)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5		

/* Recurrence

for t ← 2 to T do

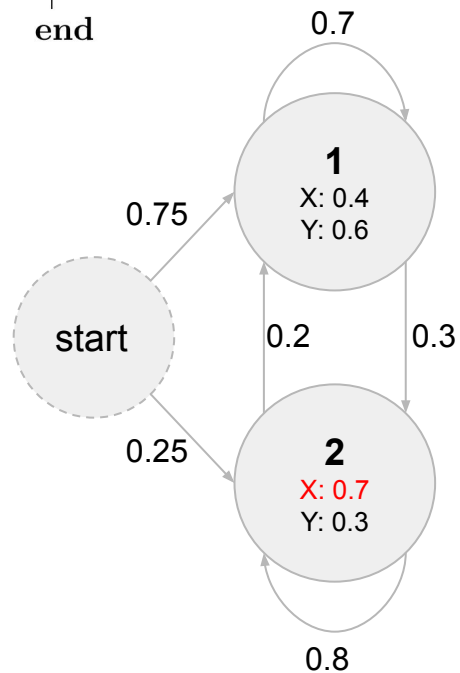
 for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

 end

end



XYYXX

j=1: max(0.0104, 0.0022)

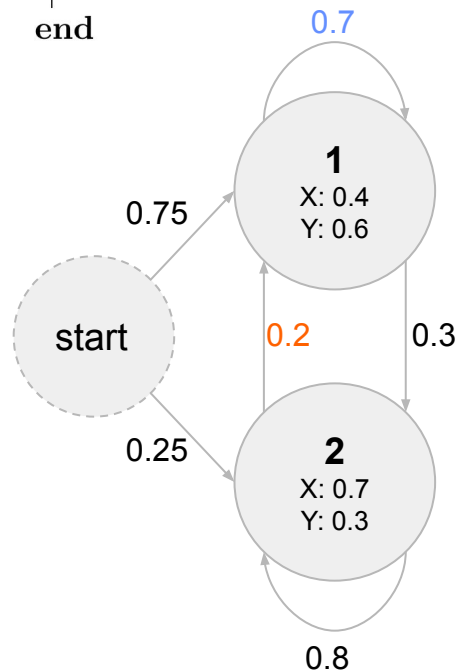
δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5		

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5		

```

/* Recurrence
for t ← 2 to T do
  for j ← 1 to n do
     $\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$ 
     $\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$ 
  end
end
end

```



XYYXX

j=1: max(0.0104, 0.0022)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	

/* Recurrence

for t ← 2 to T do

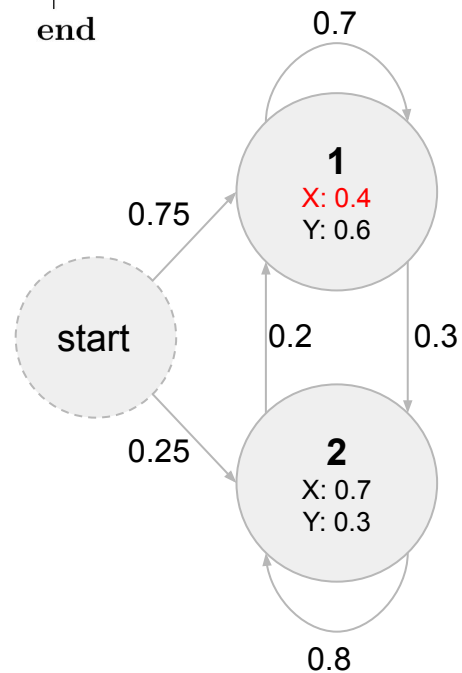
 for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

 end

end



XY YXX

j=1: max(0.0044, 0.0089)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	

/* Recurrence

for t ← 2 to T do

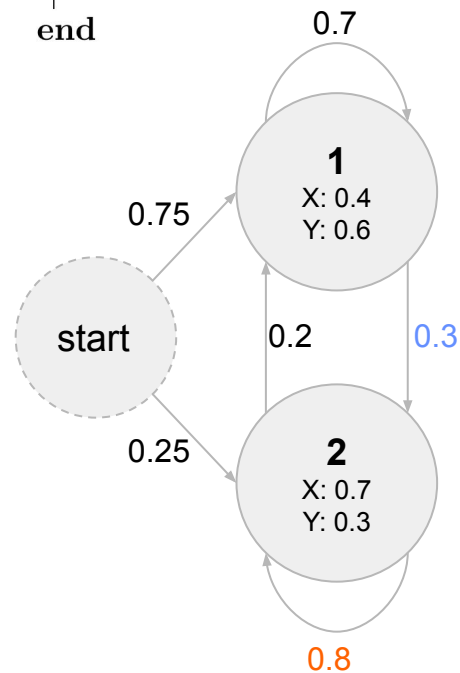
 for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

 end

end



XYYXX

j=1: max(0.0044, 0.0089)

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

/* Recurrence

for t ← 2 to T do

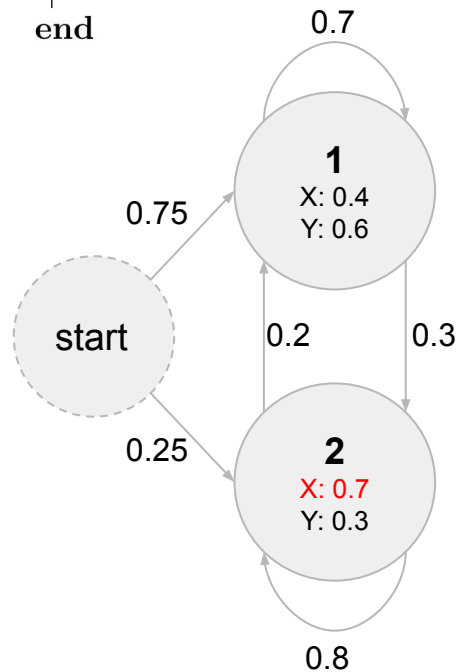
 for j ← 1 to n do

$\delta_t(j) \leftarrow \text{MAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(\theta_t)$

$\psi_t(j) \leftarrow \text{ARGMAX}_{1 \leq i \leq n} [\delta_{t-1}(i) \cdot a_{ij}]$

 end

end



XY YXX

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

XY YXX

$p^* = 0.006234$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

/* Termination

$$p^* \leftarrow \underset{1 \leq i \leq n}{\text{MAX}}[\delta_T(i)]$$

$$q_T^* \leftarrow \underset{1 \leq i \leq n}{\text{ARGMAX}}[\delta_T(i)]$$

	q^*
t=1	
2	
3	
4	
5	2

XY YXX

$p^* = 0.006234$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

```
/* Backtracking
for t ← T - 1 to 1 do
|  $q_t^* \leftarrow \psi_{t+1}(q_{t+1}^*)$ 
end
```

	q^*
t=1	
2	
3	
4	
5	2

XY YXX

$p^* = 0.006234$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

```
/* Backtracking
for t ← T - 1 to 1 do
|  $q_t^* \leftarrow \psi_{t+1}(q_{t+1}^*)$ 
end
```

	q^*
t=1	
2	
3	
4	2
5	2

XY Y XX

$p^* = 0.006234$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

```
/* Backtracking
for t ← T - 1 to 1 do
|  $q_t^* \leftarrow \psi_{t+1}(q_{t+1}^*)$ 
end
```

	q^*
t=1	
2	
3	1
4	2
5	2

XY Y XX

$p^* = 0.006234$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

```
/* Backtracking  
for t ← T - 1 to 1 do  
|  $q_t^* \leftarrow \psi_{t+1}(q_{t+1}^*)$   
end
```

	q^*
t=1	
2	1
3	1
4	2
5	2

XY YXX

$p^* = 0.006234$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

```
/* Backtracking
for t ← T - 1 to 1 do
|  $q_t^* \leftarrow \psi_{t+1}(q_{t+1}^*)$ 
end
```

	q^*
t=1	1
2	1
3	1
4	2
5	2

XY YXX

$$p^* = 0.006234$$

δ	1	2
t = 1	0.3	0.175
2	0.126	0.042
3	0.0529	0.0113
4	0.0148	0.0111
5	0.0041	0.0062

ψ	1	2
t = 1	0	0
2	1	2
3	1	1
4	1	1
5	1	2

	q^*
t=1	1
2	1
3	1
4	2
5	2

θ : XY Y XX

State sequence: 1 1 1 2 2

Probability of observing θ , given the state sequence: 0.0062234

$$p^* = 0.006234$$

	q^*
t=1	1
2	1
3	1
4	2
5	2