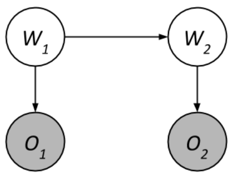
1 HMMs

Consider the following Hidden Markov Model. O_1 and O_2 are supposed to be shaded.



W_1	$P(W_1)$
0	0.3
1	0.7

W_t	W_{t+1}	$P(W_{t+1}\|W_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

W_t	O_t	$P(O_t\|W_t)$
0	a	0.9
0	b	0.1
1	a	0.5
1	b	0.5

Suppose that we observe $O_1 = a$ and $O_2 = b$.
Using the forward algorithm, compute the probability distribution $P(W_2|O_1 = a, O_2 = b)$ one step at a time.

1. Compute $P(W_1, O_1 = a)$.

   $P(W_1, O_1 = a) = P(W_1)P(O_1 = a|W_1)$
   $P(W_1 = 0, O_1 = a) = (0.3)(0.9) = 0.27$
   $P(W_1 = 1, O_1 = a) = (0.7)(0.5) = 0.35$

2. Using the previous calculation, compute $P(W_2, O_1 = a)$.

   $P(W_2, O_1 = a) = \sum_{w_1} P(w_1, O_1 = a)P(W_2|w_1)$
   $P(W_2 = 0, O_1 = a) = (0.27)(0.4) + (0.35)(0.8) = 0.388$
   $P(W_2 = 1, O_1 = a) = (0.27)(0.6) + (0.35)(0.2) = 0.232$

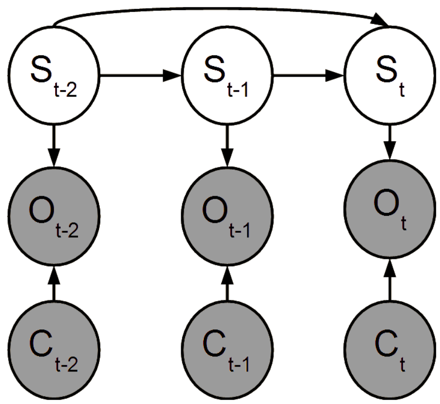
3. Using the previous calculation, compute $P(W_2, O_1 = a, O_2 = b)$.

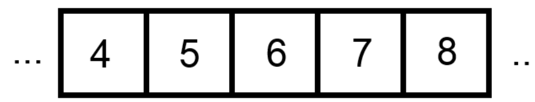
   $P(W_2, O_1 = a, O_2 = b) = P(W_2, O_1 = a)P(O_2 = b|W_2)$
   $P(W_2 = 0, O_1 = a, O_2 = b) = (0.388)(0.1) = 0.0388$
   $P(W_2 = 1, O_1 = a, O_2 = b) = (0.232)(0.5) = 0.116$

4. Finally, compute $P(W_2|O_1 = a, O_2 = b)$.

   Renormalizing the distribution above, we have
   $P(W_2 = 0|O_1 = a, O_2 = b) = 0.0388/(0.0388 + 0.116) \approx 0.25$
   $P(W_2 = 1|O_1 = a, O_2 = b) = 0.116/(0.0388 + 0.116) \approx 0.75$

Q2. Particle Filtering





Pacman is trying to hunt a ghost in an infinite hallway with positions labeled as in the picture above. He's become more technologically savvy, and decided to locate find the ghosts actual position, $\mathbf{S_t}$, using some sensors he set up. From the sensors, Pacman can find, at each time step, a noisy reading of the ghost's location, $\mathbf{O_t}$. However, just as Pacman has gained technology, so has the ghost. It is able to cloak itself at each time step, given by $\mathbf{C_t}$, adding extra noise to Pacman's sensor readings.

Pacman has generated an error model, given in the table below, for the sensor depending on whether the ghost is cloaked or not.

Pacman has also generated a dynamics model, given in the table below, that takes into account the position of the ghost at the two previous timesteps.

Dynamics model:

$P(S_t|S_{t-1}, S_{t-2}) = F(D_1, D_2)$
$D_1 = |S_t - S_{t-1}|$
$D_2 = |S_t - S_{t-2}|$

D_1	D_2	F(D_1, D_2)
0	0	0.7
0	1	0.2
0	2	0
1	0	0.3
1	1	0.3
1	2	0.5

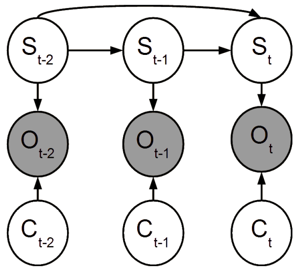
Observation model:

$P(O_t|S_t, C_t) = E(C_t, D)$
$D = |O_t - S_t|$

C	D	E(C, D)
+	0	0.4
+	1	0.2
+	2	0.1
-	0	0.6
-	1	0.2
-	2	0

(a) Assume that you currently have the following two particles: $(S_6 = 7, S_7 = 8)$ and $(S_6 = 6, S_7 = 6)$. Compute the weights for each particle given the observations $C_6 = +, C_7 = -, O_6 = 5, O_7 = 8$:

$(S_6 = 7, S_7 = 8)$	$Pr(O_6 = 5 \mid C_6 = +, S_6 = 7) * Pr(O_7 = 8 \mid C_7 = -, S_7 = 8) = 0.1 * 0.6 = 0.06$
$(S_6 = 6, S_7 = 6)$	$Pr(O_6 = 5 \mid C_6 = +, S_6 = 6) * Pr(O_7 = 8 \mid C_7 = -, S_7 = 6) = 0.2 * 0 = 0$



C	P(C)
+	0.5
-	0.5

(b) Assume that Pacman can no longer see whether the ghost is cloaked or not, but assumes that it will be cloaked at each timestep with probability 0.5. Compute the weights for each particle given the observations $O_6 = 5, O_7 = 8$:

For each of the particle's states: we want to find $Pr(o_t|s_t)$, as this is the contribution to the weight of the sample. However, we have C unobserved, so: $Pr(o_t|s_t) = \sum_{c_t} Pr(o_t, c_t|s_t) = \sum_{c_t} Pr(c_t|s_t)Pr(o_t|s_t, c_t) = \sum_{c_t} Pr(c_t)Pr(o_t|s_t, c_t)$. Last equality is due to independence between C_t and S_t.

$(S_6 = 7, S_7 = 8)$	$\sum_{c_6} Pr(c_6)Pr(O_6 = 5\|S_6 = 7, c_6) * \sum_{c_7} Pr(c_7)Pr(O_7 = 8\|S_7 = 8, c_7) = 0.05 * 0.5 = 0.025$
$(S_6 = 6, S_7 = 6)$	$\sum_{c_6} Pr(c_6)Pr(O_6 = 5\|S_6 = 6, c_6) * \sum_{c_7} Pr(c_7)Pr(O_7 = 8\|S_7 = 6, c_7) = 0.2 * 0.05 = 0.01$

(c) To prevent error propagation, assume that after weighting the particles and resampling, one of the particles you end up with is $(S_6 = 6, S_7 = 7)$.

(i) What is the probability that after passing this particle through the dynamics model it becomes $(S_7 = 6, S_8 = 6)$?
0. It's invalid for a particle to start at $S_7 = 7$ and after one transition, become $S_7 = 6$.

(ii) What is the probability the particle becomes $(S_7 = 7, S_8 = 8)$?
0.5. This is just $Pr(S_8 = 8|S_6 = 6, S_7 = 7) = F(D_1 = 1, D_2 = 2) = 0.5$.

(d) To again decouple this part from previous parts, assume that you have the following three particles with the specified weights.

Particle	weight
$(S_7 = 5, S_8 = 6)$	.1
$(S_7 = 7, S_8 = 6)$	.25
$(S_7 = 7, S_8 = 7)$	.3

What is Pacman's belief for the ghost's position at time $t = 8$?

Position	P(S_8)
$S_8 = 5$	$\frac{0}{.1+.25+.3} = 0$
$S_8 = 6$	$\frac{.1+.25}{.1+.25+.3} = \frac{.35}{.65} = \frac{7}{13}$
$S_8 = 7$	$\frac{.3}{.1+.25+.3} = \frac{.3}{.65} = \frac{6}{13}$
$S_8 = 8$	$\frac{0}{.1+.25+.3} = 0$

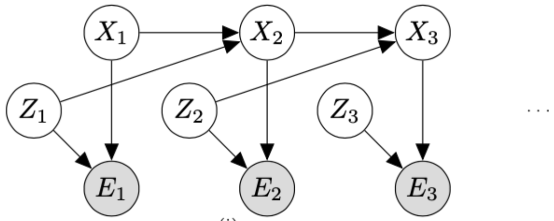
Q3. Modified HMM Updates

(a) Recall that for a standard HMM the Elapse Time update and the Observation update are of the respective forms:

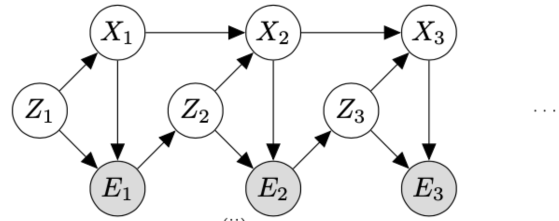
$P(X_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1:t-1})$

$P(X_t \mid e_{1:t}) \propto P(X_t \mid e_{1:t-1}) P(e_t \mid x_t)$

We now consider the following two HMM-like models:



(i)



(ii)

Mark the modified Elapse Time update and the modified Observation update that correctly compute the beliefs from the quantities that are available in the Bayes' Net. (Mark one of the first set of six options, and mark one of the second set of six options for (i), and same for (ii).)

(i)

● $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}, z_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1}) P(X_t \mid x_{t-1}, z_{t-1}) P(Z_t)$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}, z_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1}) P(X_t \mid x_{t-1}, z_{t-1})$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}, z_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1}) P(X_t, Z_t \mid x_{t-1}, z_{t-1})$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1}) P(X_t \mid x_{t-1}, z_{t-1}) P(Z_t)$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1}) P(X_t \mid x_{t-1}, z_{t-1})$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1}) P(X_t, Z_t \mid x_{t-1}, z_{t-1})$

In the elapse time update, we want to get from $P(X_{t-1}, Z_{t-1} | e_{1:t-1})$ to $P(X_t, Z_t | e_{1:t-1})$.

$$\begin{aligned} P(X_t, Z_t | e_{1:t-1}) &= \sum_{x_{t-1}, z_{t-1}} P(X_t, Z_t, x_{t-1}, z_{t-1} | e_{1:t-1}) \\ &= \sum_{x_{t-1}, z_{t-1}} P(x_{t-1}, z_{t-1} | e_{1:t-1}) P(X_t | x_{t-1}, z_{t-1}, e_{1:t-1}) P(Z_t | X_t, x_{t-1}, z_{t-1}, e_{1:t-1}) \\ &= \sum_{x_{t-1}, z_{t-1}} P(x_{t-1}, z_{t-1} | e_{1:t-1}) P(X_t | x_{t-1}, z_{t-1}) P(Z_t) \end{aligned}$$

First line: marginalization, second line: chain rule, third line: conditional independence assumptions.

● $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1}) P(e_t \mid X_t, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto \sum_{X_t} P(X_t, Z_t \mid e_{1:t-1}) P(e_t \mid X_t, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto \sum_{Z_t} P(X_t, Z_t \mid e_{1:t-1}) P(e_t \mid X_t, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1}) P(e_t \mid X_t) P(e_t \mid Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1}) P(e_t \mid X_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1}) \sum_{X_t} P(e_t \mid X_t)$

In the observation update, we want to get from $P(X_t, Z_t | e_{1:t-1})$ to $P(X_t, Z_t | e_{1:t})$.

$$\begin{aligned} P(X_t, Z_t | e_{1:t}) &\propto P(X_t, Z_t, e_t | e_{1:t-1}) \\ &\propto P(X_t, Z_t | e_{1:t-1}) P(e_t | X_t, Z_t, e_{1:t-1}) \\ &\propto P(X_t, Z_t | e_{1:t-1}) P(e_t | X_t, Z_t) \end{aligned}$$

First line: normalization, second line: chain rule, third line: conditional independence assumptions.

(ii) ○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1},z_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1})P(X_t \mid x_{t-1}, z_{t-1})P(Z_t \mid e_{t-1})$

● $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1},z_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1})P(Z_t \mid e_{t-1})P(X_t \mid x_{t-1}, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1},z_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1})P(X_t, Z_t \mid x_{t-1}, e_{t-1})$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1})P(X_t \mid x_{t-1}, z_{t-1})P(Z_t \mid e_{t-1})$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1})P(Z_t \mid e_{t-1})P(X_t \mid x_{t-1}, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(x_{t-1}, z_{t-1} \mid e_{1:t-1})P(X_t, Z_t \mid x_{t-1}, e_{t-1})$

In the elapse time update, we want to get from $P(X_{t-1}, Z_{t-1}|e_{1:t-1})$ to $P(X_t, Z_t|e_{1:t-1})$.

$$\begin{aligned}
P(X_t, Z_t|e_{1:t-1}) &= \sum_{x_{t-1},z_{t-1}} P(X_t, Z_t, x_{t-1}, z_{t-1}|e_{1:t-1}) \\
&= \sum_{x_{t-1},z_{t-1}} P(x_{t-1}, z_{t-1}|e_{1:t-1})P(Z_t|x_{t-1}, z_{t-1}, e_{1:t-1})P(X_t|Z_t, x_{t-1}, z_{t-1}, e_{1:t-1}) \\
&= \sum_{x_{t-1},z_{t-1}} P(x_{t-1}, z_{t-1}|e_{1:t-1})P(Z_t|e_{t-1})P(X_t|x_{t-1}, Z_t)
\end{aligned}$$

First line: marginalization, second line: chain rule, third line: conditional independence assumptions.

● $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1})P(e_t \mid X_t, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto \sum_{X_t} P(X_t, Z_t \mid e_{1:t-1})P(e_t \mid X_t, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto \sum_{Z_t} P(X_t, Z_t \mid e_{1:t-1})P(e_t \mid X_t, Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1})P(e_t \mid X_t)P(e_t \mid Z_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1})P(e_t \mid X_t)$

○ $P(X_t, Z_t \mid e_{1:t}) \propto P(X_t, Z_t \mid e_{1:t-1}) \sum_{X_t} P(e_t \mid X_t)$

In the observation update, we want to get from $P(X_t, Z_t|e_{1:t-1})$ to $P(X_t, Z_t|e_{1:t})$.

$$\begin{aligned}
P(X_t, Z_t|e_{1:t}) &\propto P(X_t, Z_t, e_t|e_{1:t-1}) \\
&\propto P(X_t, Z_t|e_{1:t-1})P(e_t|X_t, Z_t, e_{1:t-1}) \\
&\propto P(X_t, Z_t|e_{1:t-1})P(e_t|X_t, Z_t)
\end{aligned}$$

First line: normalization, second line: chain rule, third line: conditional independence assumptions.