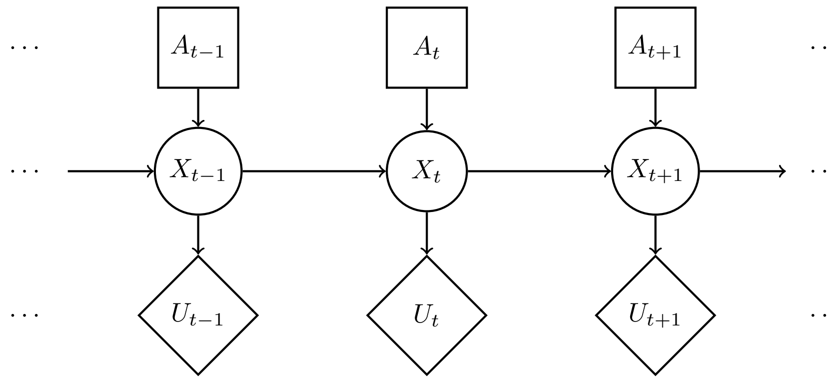
CS 188 Summer 2019 | Introduction to Artificial Intelligence

Written HW 5 Sol.

Self-assessment due: Tuesday 8/6/2019 at 11:59pm (submit via Gradescope)

Q1. Planning ahead with HMMs

Pacman is tired of using HMMs to estimate the location of ghosts. He wants to use HMMs to plan what actions to take in order to maximize his utility. Pacman uses the HMM (drawn to the right) of length T to model the planning problem. In the HMM, $X_{1:T}$ is the sequence of hidden states of Pacman's world, $A_{1:T}$ are actions Pacman can take, and U_t is the utility Pacman receives at the particular hidden state X_t. Notice that there are no evidence variables, and utilities are not discounted.



(a) The belief at time t is defined as $B_t(X_t) = p(X_t|a_{1:t})$. The forward algorithm update has the following form:

$$B_t(X_t) = \underline{\quad\textbf{(i)}\quad}\ \underline{\quad\textbf{(ii)}\quad}\ B_{t-1}(x_{t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(i) ○ $\max_{x_{t-1}}$ ● $\sum_{x_{t-1}}$ ○ $\max_{x_t}$ ○ $\sum_{x_t}$ ○ 1

(ii) ○ $p(X_t|x_{t-1})$ ○ $p(X_t|x_{t-1})p(X_t|a_t)$ ○ $p(X_t)$ ● $p(X_t|x_{t-1}, a_t)$ ○ 1

○ None of the above combinations is correct

$$\begin{aligned} B_t(X_t) &= p(X_t|a_{1:t}) \\ &= \sum_{x_{t-1}} p(X_t|x_{t-1}, a_t)p(x_{t-1}|a_{1:t-1}) \\ &= \sum_{x_{t-1}} p(X_t|x_{t-1}, a_t)B_{t-1}(x_{t-1}) \end{aligned}$$

(b) Pacman would like to take actions $A_{1:T}$ that maximizes the expected sum of utilities, which has the following form:

$$\text{MEU}_{1:T} = \underline{\ \textbf{(i)}\ }\ \underline{\ \textbf{(ii)}\ }\ \underline{\ \textbf{(iii)}\ }\ \underline{\ \textbf{(iv)}\ }\ \underline{\ \textbf{(v)}\ }$$

Complete the expression by choosing the option that fills in each blank.

(i) ● $\max_{a_{1:T}}$ ○ $\max_{a_T}$ ○ $\sum_{a_{1:T}}$ ○ $\sum_{a_T}$ ○ 1

(ii) ○ $\max_t$ ○ $\prod_{t=1}^{T}$ ● $\sum_{t=1}^{T}$ ○ $\min_t$ ○ 1

(iii) ○ $\sum_{x_t, a_t}$ ● $\sum_{x_t}$ ○ $\sum_{a_t}$ ○ $\sum_{x_T}$ ○ 1

(iv) ○ $p(x_t|x_{t-1}, a_t)$ ○ $p(x_t)$ ● $B_t(x_t)$ ○ $B_T(x_T)$ ○ 1

(v) ○ U_T ○ $\frac{1}{U_t}$ ○ $\frac{1}{U_T}$ ● U_t ○ 1

○ None of the above combinations is correct

$$\text{MEU}_{1:T} = \max_{a_{1:T}} \sum_{t=1}^{T} \sum_{x_t} B_t(x_t)U_t(x_t)$$

(c) A greedy ghost now offers to tell Pacman the values of some of the hidden states. Pacman needs your help to figure out if the ghost's information is useful. Assume that the transition function $p(x_t|x_{t-1}, a_t)$ is not deterministic. **With respect to the utility** U_t, mark all that can be True:

■ VPI($X_{t-1}|X_{t-2}$) > 0 ☐ VPI($X_{t-2}|X_{t-1}$) > 0 ■ VPI($X_{t-1}|X_{t-2}$) $= 0$ ■ VPI($X_{t-2}|X_{t-1}$) $= 0$
☐ None of the above

It is always possible that VPI $= 0$. Can guarantee VPI($E|e$) is not greater than 0 if E is independent of parents(U) given e.

(d) Pacman notices that calculating the beliefs under this model is very slow using exact inference. He therefore decides to try out various particle filter methods to speed up inference. Order the following methods by how accurate their estimate of $B_T(X_T)$ is? If different methods give an equivalently accurate estimate, mark them as the same number.

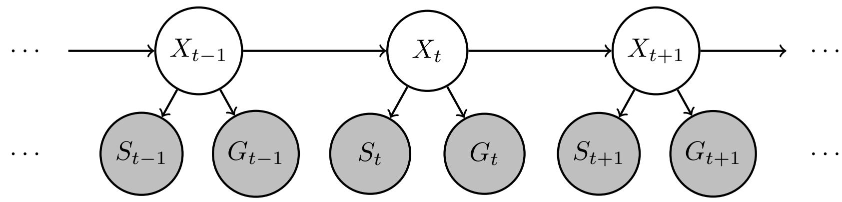
	Most accurate			Least accurate
Exact inference	● 1	○ 2	○ 3	○ 4
Particle filtering with no resampling	○ 1	● 2	○ 3	○ 4
Particle filtering with resampling before every time elapse	○ 1	○ 2	○ 3	● 4
Particle filtering with resampling before every other time elapse	○ 1	○ 2	● 3	○ 4

Exact inference will always be more accurate than using a particle filter. When comparing the particle filter resampling approaches, notice that because there are no observations, each particle will have weight 1. Therefore resampling when particle weights are 1 could lead to particles being lost and hence prove bad.

Q2. HMM: Where is the Car?

Transportation researchers are trying to improve traffic in the city but, in order to do that, they first need to estimate the location of each of the cars in the city. They need our help to model this problem as an inference problem of an HMM. For this question, assume that only *one* car is being modeled.

(a) The structure of this modified HMM is given below, which includes X, the location of the car; S, the noisy location of the car from the signal strength at a nearby cell phone tower; and G, the noisy location of the car from GPS.



We want to perform filtering with this HMM. That is, we want to compute the belief $P(x_t|s_{1:t}, g_{1:t})$, the probability of a state x_t given all past and current observations.

The **dynamics update** expression has the following form:

$$P(x_t|s_{1:t-1}, g_{1:t-1}) = \underline{\quad\textbf{(i)}\quad} \; \underline{\quad\textbf{(ii)}\quad} \; \underline{\quad\textbf{(iii)}\quad} \; P(x_{t-1}|s_{1:t-1}, g_{1:t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(i)	○ $P(s_{1:t-1}, g_{1:t-1})$	○ $P(s_{1:t}, g_{1:t})$	○ $P(s_{1:t-1})P(g_{1:t-1})$	○ $P(s_{1:t})P(g_{1:t})$	● 1
(ii)	● $\sum_{x_{t-1}}$	○ $\sum_{x_t}$	○ $\max_{x_{t-1}}$	○ $\max_{x_t}$	○ 1
(iii)	○ $P(x_{t-2}, x_{t-1})$	○ $P(x_{t-1} \mid x_{t-2})$	○ $P(x_{t-1}, x_t)$	● $P(x_t|x_{t-1})$	○ 1

The derivation of the dynamics update is similar to the one for the canonical HMM, but with two observation variables instead.

$$\begin{aligned} P(x_t|s_{1:t-1}, g_{1:t-1}) &= \sum_{x_{t-1}} P(x_{t-1}, x_t|s_{1:t-1}, g_{1:t-1}) \\ &= \sum_{x_{t-1}} P(x_t|x_{t-1}, s_{1:t-1}, g_{1:t-1}) P(x_{t-1}|s_{1:t-1}, g_{1:t-1}) \\ &= \sum_{x_{t-1}} P(x_t|x_{t-1}) P(x_{t-1}, x_t|s_{1:t-1}, g_{1:t-1}) \end{aligned}$$

In the last step, we use the independence assumption given in the HMM, $X_t S_{1:t-1}, G_{1:t-1}|X_{t-1}$.

The **observation update** expression has the following form:

$$P(x_t|s_{1:t}, g_{1:t}) = \underline{\quad\textbf{(iv)}\quad}\ \underline{\quad\textbf{(v)}\quad}\ \underline{\quad\textbf{(vi)}\quad}\ P(x_t|s_{1:t-1}, g_{1:t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(iv)
- ○ $P(s_t, g_t|s_{1:t-1}, g_{1:t-1})$
- ○ $P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)$
- ○ $P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})$
- ○ $P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)$
- ● $\frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})}$
- ○ $\frac{1}{P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)}$
- ○ $\frac{1}{P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})}$
- ○ $\frac{1}{P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)}$
- ○ 1

(v)
- ○ $\sum_{x_{t-1}}$
- ○ $\sum_{x_t}$
- ○ $\max_{x_{t-1}}$
- ○ $\max_{x_t}$
- ● 1

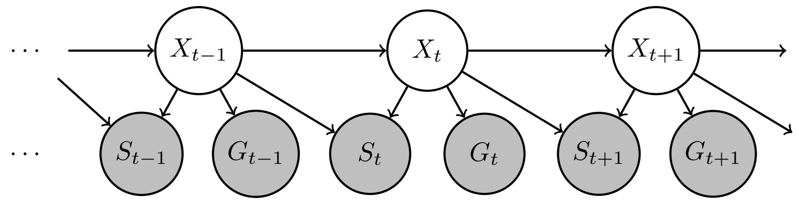
(vi)
- ○ $P(x_{t-1}, s_{t-1})P(x_{t-1}, g_{t-1})$
- ○ $P(x_{t-1}, s_{t-1}, g_{t-1})$
- ○ $P(x_t|s_t)P(x_t|g_t)$
- ○ $P(s_{t-1}|x_{t-1})P(g_{t-1}|x_{t-1})$
- ○ $P(x_t, s_t)P(x_t, g_t)$
- ○ $P(x_t, s_t, g_t)$
- ○ $P(x_{t-1}|s_{t-1})P(x_{t-1}|g_{t-1})$
- ● $P(s_t|x_t)P(g_t|x_t)$
- ○ 1

Again, the derivation of the observation update is similar to the one for the canonical HMM, but with two observation variables instead.

$$\begin{aligned}
P(x_t|s_{1:t}, g_{1:t}) &= P(x_t|s_t, g_t, s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} P(x_t, s_t, g_t|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} P(s_t, g_t|x_t, s_{1:t-1}, g_{1:t-1}) P(x_t|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} P(s_t, g_t|x_t) P(x_t|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} P(s_t|x_t) P(g_t|x_t) P(x_t|s_{1:t-1}, g_{1:t-1})
\end{aligned}$$

In the second to last step, we use the independence assumption $S_t, G_t S_{1:t-1}, G_{1:t-1}|X_t$; and in the last step, we use the independence assumption $S_t G_t|X_t$.

(b) It turns out that if the car moves too fast, the quality of the cell phone signal decreases. Thus, the signal-dependent location S_t not only depends on the current state X_t but it also depends on the previous state X_{t-1}. Thus, we modify our original HMM for a new more accurate one, which is given below.



Again, we want to compute the belief $P(x_t|s_{1:t}, g_{1:t})$. In this part we consider an update that combines the dynamics and observation update in a *single* update.

$P(x_t|s_{1:t}, g_{1:t}) =$ ____**(i)**____ ____**(ii)**____ ____**(iii)**____ ____**(iv)**____ $P(x_{t-1}|s_{1:t-1}, g_{1:t-1})$.

Complete the **forward update** expression by choosing the option that fills in each blank.

(i)
- ○ $P(s_t, g_t|s_{1:t-1}, g_{1:t-1})$
- ○ $P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)$
- ○ $P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})$
- ○ $\frac{1}{P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)}$
- ● $\frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})}$
- ○ $P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)$
- ○ $\frac{1}{P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})}$
- ○ $\frac{1}{P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)}$
- ○ 1

(ii)
- ● $\sum_{x_{t-1}}$
- ○ $\sum_{x_t}$
- ○ $\max_{x_{t-1}}$
- ○ $\max_{x_t}$
- ○ 1

(iii)
- ○ $P(x_{t-2}, x_{t-1}, s_{t-1})P(x_{t-1}, g_{t-1})$
- ○ $P(x_{t-1}, x_t, s_t)P(x_t, g_t)$
- ○ $P(s_{t-1}, g_{t-1}|x_{t-1})$
- ○ $P(s_{t-1}|x_{t-2}, x_{t-1})P(g_{t-1}|x_{t-1})$
- ● $P(s_t|x_{t-1}, x_t)P(g_t|x_t)$
- ○ $P(s_t, g_t|x_t)$
- ○ $P(x_{t-2}, x_{t-1}|s_{t-1})P(x_{t-1}|g_{t-1})$
- ○ $P(x_{t-1}, x_t|s_t)P(x_t|g_t)$
- ○ 1
- ○ $P(x_{t-2}, x_{t-1}, s_{t-1}, g_{t-1})$
- ○ $P(x_{t-1}, x_t, s_t, g_t)$

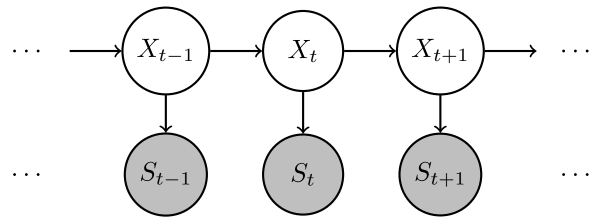
(iv)
- ○ $P(x_{t-1}, x_t)$
- ● $P(x_t|x_{t-1})$
- ○ $P(x_{t-2}, x_{t-1})$
- ○ $P(x_{t-1}|x_{t-2})$
- ○ 1

For this modified HMM, we have the dynamics and observation update in a single update because one of the previous independence assumptions does not longer holds.

$$
\begin{aligned}
P(x_t|s_{1:t}, g_{1:t}) &= \sum_{x_{t-1}} P(x_{t-1}, x_t|s_t, g_t, s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} \sum_{x_{t-1}} P(x_{t-1}, x_t, s_t, g_t|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} \sum_{x_{t-1}} P(s_t, g_t|x_{t-1}, x_t, s_{1:t-1}, g_{1:t-1})P(x_{t-1}, x_t|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} \sum_{x_{t-1}} P(s_t, g_t|x_{t-1}, x_t)P(x_t|x_{t-1}, s_{1:t-1}, g_{1:t-1})P(x_{t-1}|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} \sum_{x_{t-1}} P(s_t|x_{t-1}, x_t)P(g_t|x_{t-1}, x_t)P(x_t|x_{t-1})P(x_{t-1}|s_{1:t-1}, g_{1:t-1}) \\
&= \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} \sum_{x_{t-1}} P(s_t|x_{t-1}, x_t)P(g_t|x_t)P(x_t|x_{t-1})P(x_{t-1}|s_{1:t-1}, g_{1:t-1})
\end{aligned}
$$

In the third to last step, we use the independence assumption $S_t, G_t S_{1:t-1}, G_{1:t-1}|X_{t-1}, X_t$; in the second to last step, we use the independence assumption $S_t G_t|X_{t-1}, X_t$ and $X_t S_{1:t-1}, G_{1:t-1}|X_{t-1}$; and in the last step, we use the independence assumption $G_t X_{t-1}|X_t$.

(c) The Viterbi algorithm finds the most probable sequence of hidden states $X_{1:T}$, given a sequence of observations $s_{1:T}$, for some time $t = T$. Recall the canonical HMM structure, which is shown below.



For this canonical HMM, the Viterbi algorithm performs the following dynamic programming computations:

$$m_t[x_t] = P(s_t|x_t) \max_{x_{t-1}} P(x_t|x_{t-1}) m_{t-1}[x_{t-1}].$$

We consider extending the Viterbi algorithm for the modified HMM from part **(b)**. We want to find the most likely sequence of states $X_{1:T}$ given the sequence of observations $s_{1:T}$ and $g_{1:T}$. The dynamic programming update for $t > 1$ for the modified HMM has the following form:

$$m_t[x_t] = \underline{\quad\textbf{(i)}\quad} \; \underline{\quad\quad\textbf{(ii)}\quad\quad} \; \underline{\quad\quad\textbf{(iii)}\quad\quad} \; m_{t-1}[x_{t-1}].$$

Complete the expression by choosing the option that fills in each blank.

(i) ○ $\sum_{x_{t-1}}$ ○ $\sum_{x_t}$ ● $\max_{x_{t-1}}$ ○ $\max_{x_t}$ ○ 1

(ii) ○ $P(x_{t-2}, x_{t-1}, s_{t-1})P(x_{t-1}, g_{t-1})$ ○ $P(x_{t-1}, x_t, s_t)P(x_t, g_t)$ ○ $P(s_{t-1}, g_{t-1}|x_{t-1})$
○ $P(s_{t-1}|x_{t-2}, x_{t-1})P(g_{t-1}|x_{t-1})$ ● $P(s_t|x_{t-1}, x_t)P(g_t|x_t)$ ○ $P(s_t, g_t|x_t)$
○ $P(x_{t-2}, x_{t-1}|s_{t-1})P(x_{t-1}|g_{t-1})$ ○ $P(x_{t-1}, x_t|s_t)P(x_t|g_t)$ ○ 1
○ $P(x_{t-2}, x_{t-1}, s_{t-1}, g_{t-1})$ ○ $P(x_{t-1}, x_t, s_t, g_t)$

(iii) ○ $P(x_{t-1}, x_t)$ ● $P(x_t|x_{t-1})$ ○ $P(x_{t-2}, x_{t-1})$ ○ $P(x_{t-1}|x_{t-2})$ ○ 1

If we remove the summation from the forward update equation of part **(b)**, we get a joint probability of the states,

$$P(x_{1:t}|s_{1:t}, g_{1:t}) = \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} P(s_t|x_{t-1}, x_t)P(g_t|x_t)P(x_t|x_{t-1})P(x_{1:t-1}|s_{1:t-1}, g_{1:t-1}).$$

We can define $m_t[x_t]$ to be the maximum joint probability of the states (for a particular x_t) given all past and current observations, times some constant, and then we can find a recursive relationship for $m_t[x_t]$,

$$\begin{aligned}
m_t[x_t] &= P(s_{1:t}, g_{1:t}) \max_{x_{1:t-1}} P(x_{1:t}|s_{1:t}, g_{1:t}) \\
&= P(s_{1:t}, g_{1:t}) \max_{x_{1:t-1}} \frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} P(s_t|x_{t-1}, x_t)P(g_t|x_t)P(x_t|x_{t-1})P(x_{1:t-1}|s_{1:t-1}, g_{1:t-1}) \\
&= \max_{x_{t-1}} P(s_t|x_{t-1}, x_t)P(g_t|x_t)P(x_t|x_{t-1}) \frac{P(s_{1:t}, g_{1:t})}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})} \max_{x_{1:t-2}} P(x_{1:t-1}|s_{1:t-1}, g_{1:t-1}) \\
&= \max_{x_{t-1}} P(s_t|x_{t-1}, x_t)P(g_t|x_t)P(x_t|x_{t-1})P(s_{1:t-1}, g_{1:t-1}) \max_{x_{1:t-2}} P(x_{1:t-1}|s_{1:t-1}, g_{1:t-1}) \\
&= \max_{x_{t-1}} P(s_t|x_{t-1}, x_t)P(g_t|x_t)P(x_t|x_{t-1}) m_{t-1}[x_{t-1}].
\end{aligned}$$

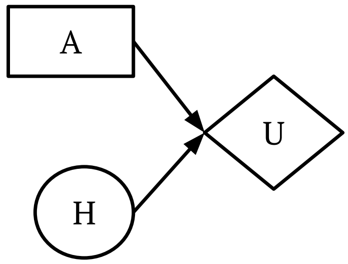
Notice that the maximum joint probability of states up to time $t = T$ given all past and current observations is given by

$$\max_{x_{1:T}} P(x_{1:T}|s_{1:T}, g_{1:T}) = \frac{\max_{x_t} m_T[x_t]}{P(s_{1:T}, g_{1:T})}.$$

We can recover the actual most likely sequence of states by bookkepping back pointers of the states the maximized the Viterbi update equations.

Q3. Decision Networks

After years of battles between the ghosts and Pacman, the ghosts challenge Pacman to a winner-take-all showdown, and the game is a coin flip. Pacman has a decision to make: whether to accept the challenge (*accept*) or decline (*decline*). If the coin comes out heads ($+h$) Pacman wins. If the coin comes out tails ($-h$), the ghosts win. No matter what decision Pacman makes, the outcome of the coin is revealed.



H	$P(H)$
+h	0.5
-h	0.5

H	A	U(H,A)
+h	*accept*	100
-h	*accept*	-100
+h	*decline*	-30
-h	*decline*	50

(a) Maximum Expected Utility

Compute the following quantities:

$EU(accept) = P(+h)U(+h, accept) + P(-h)U(-h, accept) = 0.5 * 100 + 0.5 * -100 = 0$

$EU(decline) = P(+h)U(+h, decline) + P(-h)U(-h, decline) = 0.5 * -30 + 0.5 * 50 = 10$

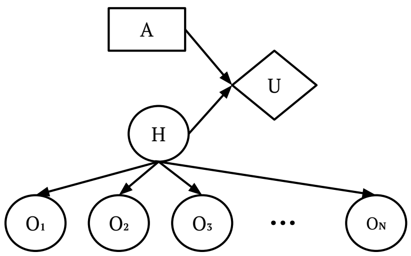
$MEU(\{\}) = max(0, 10) = 10$

Action that achieves $MEU(\{\})$ = *decline*

(b) VPI relationships When deciding whether to accept the winner-take-all coin flip, Pacman can consult a few fortune tellers that he knows. There are N fortune tellers, and each one provides a prediction O_n for H.

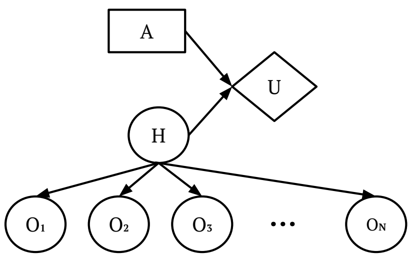
For each of the questions below, select **all** of the VPI relations that are guaranteed to be true, or select *None of the above.*

(i) In this situation, the fortune tellers give perfect predictions.
Specifically, $P(O_n = +h \mid H = +h) = 1$, $P(O_n = -h \mid H = -h) = 1$, for all n from 1 to N.



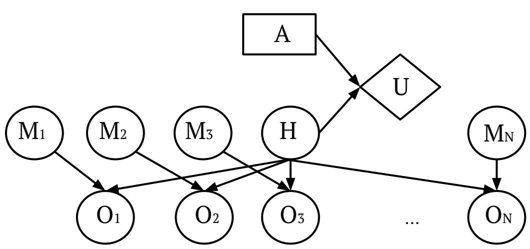
- ☐ $\text{VPI}(O_1,\ O_2) \geq \text{VPI}(O_1) + \text{VPI}(O_2)$
- ■ $\text{VPI}(O_i) = \text{VPI}(O_j)$ where $i \neq j$
- ☐ $\text{VPI}(O_3 \mid O_2,\ O_1) > \text{VPI}(O_2 \mid O_1)$.
- ☐ $\text{VPI}(H) > \text{VPI}(O_1, O_2, \ldots O_N)$
- ☐ None of the above.

(ii) In another situation, the fortune tellers are pretty good, but not perfect.
Specifically, $P(O_n = +h \mid H = +h) = 0.8$, $P(O_n = -h \mid H = -h) = 0.5$, for all n from 1 to N.



- ☐ $\text{VPI}(O_1,\ O_2) \geq \text{VPI}(O_1) + \text{VPI}(O_2)$
- ■ $\text{VPI}(O_i) = \text{VPI}(O_j)$ where $i \neq j$
- ☐ $\text{VPI}(O_3 \mid O_2,\ O_1) > \text{VPI}(O_2 \mid O_1)$.
- ■ $\text{VPI}(H) > \text{VPI}(O_1, O_2, \ldots O_N)$
- ☐ None of the above.

(iii) In a third situation, each fortune teller's prediction is affected by their mood. If the fortune teller is in a good mood ($+m$), then that fortune teller's prediction is guaranteed to be correct. If the fortune teller is in a bad mood ($-m$), then that teller's prediction is guaranteed to be incorrect. Each fortune teller is happy with probability $P(M_n = +m) = 0.8$.

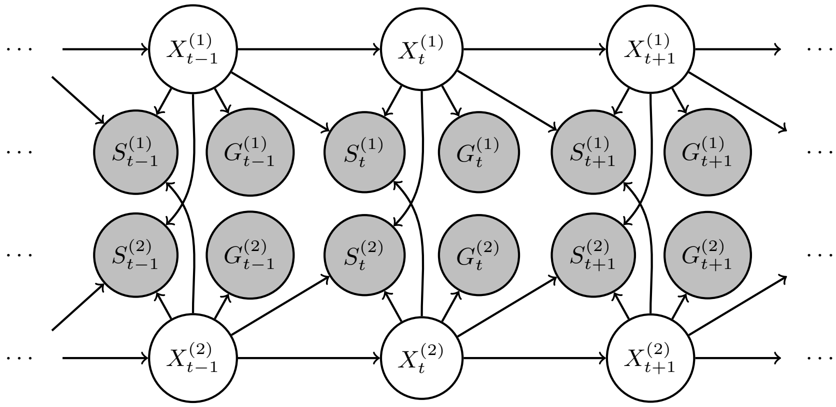


- ☐ $\text{VPI}(M_1) > 0$
- ■ $\forall i\ \text{VPI}(M_i|O_i) > 0$
- ☐ $\text{VPI}(M_1,\ M_2,\ \ldots,\ M_N) > \text{VPI}(M_1)$
- ■ $\forall i\ \text{VPI}(H) = \text{VPI}(M_i, O_i)$
- ☐ None of the above.

Q4. Particle Filtering: Where are the Two Cars?

As before, we are trying to estimate the location of cars in a city, but now, we model two cars jointly, i.e. car i for $i \in \{1, 2\}$. The modified HMM model is as follows:

- $X^{(i)}$ – the location of car i
- $S^{(i)}$ – the noisy location of the car i from the signal strength at a nearby cell phone tower
- $G^{(i)}$ – the noisy location of car i from GPS



d	$D(d)$	$E_L(d)$	$E_N(d)$	$E_G(d)$
-4	0.05	0	0.02	0
-3	0.10	0	0.04	0.03
-2	0.25	0.05	0.09	0.07
-1	0.10	0.10	0.20	0.15
0	0	0.70	0.30	0.50
1	0.10	0.10	0.20	0.15
2	0.25	0.05	0.09	0.07
3	0.10	0	0.04	0.03
4	0.05	0	0.02	0

The signal strength from one car gets noisier if the other car is at the same location. Thus, the observation $S_t^{(i)}$ also depends on the current state of the other car $X_t^{(j)}$, $j \neq i$.

The transition is modeled using a drift model D, the GPS observation $G_t^{(i)}$ using the error model E_G, and the observation $S_t^{(i)}$ using one of the error models E_L or E_N, depending on the car's speed and the relative location of both cars. These drift and error models are in the table above. **The transition and observation models are:**

$$P(X_t^{(i)}|X_{t-1}^{(i)}) = D(X_t^{(i)} - X_{t-1}^{(i)})$$

$$P(S_t^{(i)}|X_{t-1}^{(i)}, X_t^{(i)}, X_t^{(j)}) = \begin{cases} E_N(X_t^{(i)} - S_t^{(i)}), & \text{if } |X_t^{(i)} - X_{t-1}^{(i)}| \geq 2 \text{ or } X_t^{(i)} = X_t^{(j)} \\ E_L(X_t^{(i)} - S_t^{(i)}), & \text{otherwise} \end{cases}$$

$$P(G_t^{(i)}|X_t^{(i)}) = E_G(X_t^{(i)} - G_t^{(i)}).$$

Throughout this problem you may give answers either as unevaluated numeric expressions (e.g. $0.1 \cdot 0.5$) or as numeric values (e.g. 0.05). The questions are decoupled.

(a) Assume that at $t = 3$, we have the single particle $(X_3^{(1)} = -1, X_3^{(2)} = 2)$.

(i) What is the probability that this particle becomes $(X_4^{(1)} = -3, X_4^{(2)} = 3)$ after passing it through the dynamics model?

$$\begin{aligned} P(X_4^{(1)} = -3, X_4^{(2)} = 3|X_3^{(1)} = -1, X_3^{(2)} = 2) &= P(X_4^{(1)} = -3|X_3^{(1)} = -1) \cdot P(X_4^{(2)} = 3|X_3^{(2)} = 2) \\ &= D(-3 - (-1)) \cdot D(3 - 2) \\ &= 0.25 \cdot 0.10 \\ &= 0.025 \end{aligned}$$

Answer: 0.025

(ii) Assume that there are no sensor readings at $t = 4$. What is the joint probability that the *original* single particle (from $t = 3$) becomes $(X_4^{(1)} = -3, X_4^{(2)} = 3)$ and then becomes $(X_5^{(1)} = -4, X_5^{(2)} = 4)$?

$$\begin{aligned}
&P(X_4^{(1)} = -3, X_5^{(1)} = -4, X_4^{(2)} = 3, X_5^{(2)} = 4 | X_3^{(1)} = -1, X_3^{(2)} = 2) \\
&\quad = P(X_4^{(1)} = -3, X_5^{(1)} = -4 | X_3^{(1)} = -1) \cdot P(X_4^{(2)} = 3, X_5^{(2)} = 4 | X_3^{(2)} = 2) \\
&\quad = P(X_5^{(1)} = -4 | X_4^{(1)} = -3) \cdot P(X_4^{(1)} = -3 | X_3^{(1)} = -1) \cdot P(X_5^{(2)} = 4 | X_4^{(2)} = 3) \cdot P(X_4^{(2)} = 3 | X_3^{(2)} = 2) \\
&\quad = D(-4 - (-3)) \cdot D(-3 - (-1)) \cdot D(4 - 3) \cdot D(3 - 2) \\
&\quad = 0.10 \cdot 0.25 \cdot 0.10 \cdot 0.10 \\
&\quad = 0.00025
\end{aligned}$$

Answer: 0.00025

For the remaining of this problem, we will be using 2 particles at each time step.

(b) At $t = 6$, we have particles $[(X_6^{(1)} = 3, X_6^{(2)} = 0), (X_6^{(1)} = 3, X_6^{(2)} = 5)]$. Suppose that after weighting, resampling, and transitioning from $t = 6$ to $t = 7$, the particles become $[(X_7^{(1)} = 2, X_7^{(2)} = 2), (X_7^{(1)} = 4, X_7^{(2)} = 1)]$.

(i) At $t = 7$, you get the observations $S_7^{(1)} = 2, G_7^{(1)} = 2, S_7^{(2)} = 2, G_7^{(2)} = 2$. What is the weight of each particle?

Particle	Weight
$(X_7^{(1)} = 2, X_7^{(2)} = 2)$	$P(S_7^{(1)} = 2 \| X_6^{(1)} = 3, X_7^{(1)} = 2, X_7^{(2)} = 2) \cdot P(G_7^{(1)} = 2 \| X_7^{(1)} = 2) \cdot$ $P(S_7^{(2)} = 2 \| X_6^{(2)} = 0, X_7^{(2)} = 2, X_7^{(1)} = 2) \cdot P(G_7^{(2)} = 2 \| X_7^{(2)} = 2)$ $= E_N(2-2) \cdot E_G(2-2) \cdot E_N(2-2) \cdot E_G(2-2)$ $= 0.30 \cdot 0.50 \cdot 0.30 \cdot 0.50$ $= 0.0225$
$(X_7^{(1)} = 4, X_7^{(2)} = 1)$	$P(S_7^{(1)} = 2 \| X_6^{(1)} = 3, X_7^{(1)} = 4, X_7^{(2)} = 1) \cdot P(G_7^{(1)} = 2 \| X_7^{(1)} = 4) \cdot$ $P(S_7^{(2)} = 2 \| X_6^{(2)} = 5, X_7^{(2)} = 1, X_7^{(1)} = 4) \cdot P(G_7^{(2)} = 2 \| X_7^{(2)} = 1)$ $= E_L(4-2) \cdot E_G(4-2) \cdot E_N(1-2) \cdot E_G(1-2)$ $= 0.05 \cdot 0.07 \cdot 0.20 \cdot 0.15$ $= 0.000105$

(ii) Suppose both cars' cell phones died so you only get the observations $G_7^{(1)} = 2, G_7^{(2)} = 2$. What is the weight of each particle?

Particle	Weight
$(X_7^{(1)} = 2, X_7^{(2)} = 2)$	$P(G_7^{(1)} = 2 \| X_7^{(1)} = 2) \cdot P(G_7^{(2)} = 2 \| X_7^{(2)} = 2)$ $= E_G(2-2) \cdot E_G(2-2)$ $= 0.50 \cdot 0.50$ $= 0.25$
$(X_7^{(1)} = 4, X_7^{(2)} = 1)$	$P(G_7^{(1)} = 2 \| X_7^{(1)} = 4) \cdot P(G_7^{(2)} = 2 \| X_7^{(2)} = 1)$ $= E_G(4-2) \cdot E_G(1-2)$ $= 0.07 \cdot 0.15$ $= 0.0105$

(c) To decouple this question, assume that you got the following weights for the two particles.

Particle	Weight
$(X_7^{(1)} = 2, X_7^{(2)} = 2)$	0.09
$(X_7^{(1)} = 4, X_7^{(2)} = 1)$	0.01

What is the belief for the location of car 1 and car 2 at $t = 7$?

Location	$P(X_7^{(1)})$	$P(X_7^{(2)})$
$X_7^{(i)} = 1$	$\frac{0}{0.09+0.01} = 0$	$\frac{0.01}{0.09+0.01} = 0.1$
$X_7^{(i)} = 2$	$\frac{0.09}{0.09+0.01} = 0.9$	$\frac{0.09}{0.09+0.01} = 0.9$
$X_7^{(i)} = 4$	$\frac{0.01}{0.09+0.01} = 0.1$	$\frac{0}{0.09+0.01} = 0$