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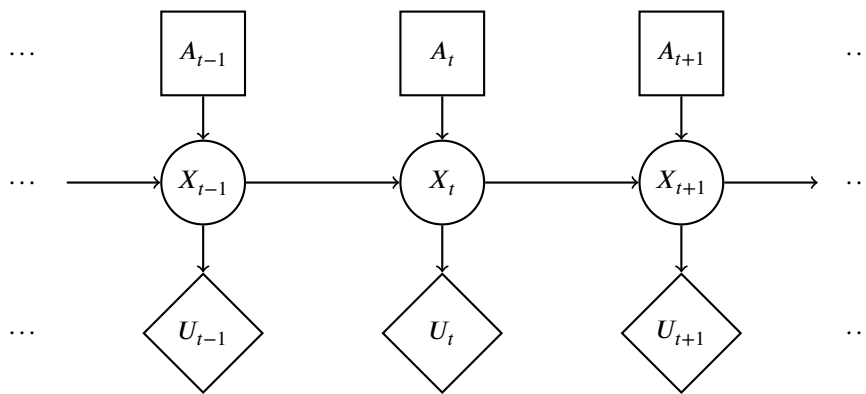
(a) The belief at time t is defined as $B_t(X_t) = p(X_t|a_{1:t})$. The forward algorithm update has the following form:

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

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

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

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

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

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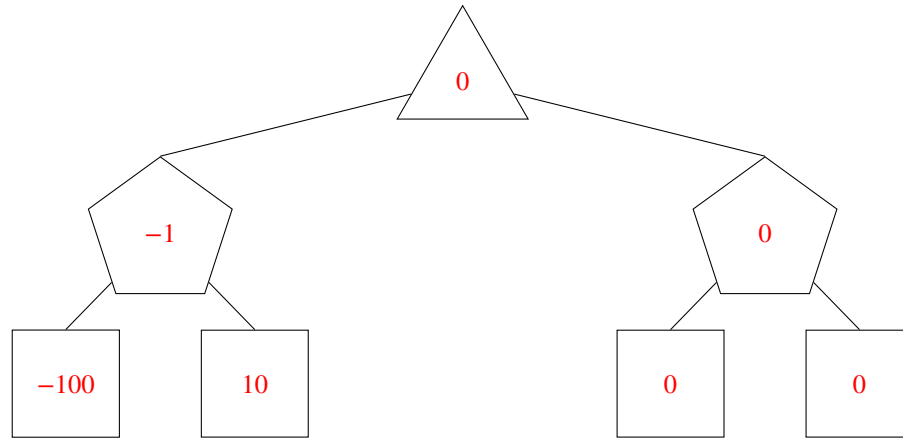
	Most accurate			Least accurate
Exact inference	● 1	○ 2	○ 3	○ 4
Particle filtering with no resampling	○ 1	● 2	○ 3	○ 4
Particle filtering with resampling before every time elapse	○ 1	○ 2	○ 3	● 4
Particle filtering with resampling before every other time elapse	○ 1	○ 2	● 3	○ 4

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.
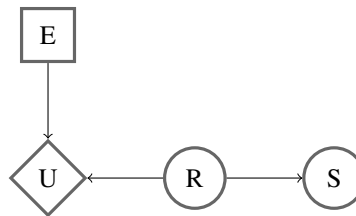
Q2. Decision Networks and VPI

Valerie has just found a cookie on the ground. She is concerned that the cookie contains raisins, which she really dislikes but she still wants to eat the cookie. If she eats the cookie and it contains raisins she will receive a utility of -100 and if the cookie doesn't contain raisins she will receive a utility of 10. If she doesn't eat the cookie she will get 0 utility. The cookie contains raisins with probability 0.1.

(a) We want to represent this decision network as an expectimax game tree. Fill in the nodes of the tree below, with the top node representing her maximizing choice.

```
                          ___
                         /   \
                        /  0  \
                       /_____\
                      /         \
                     /           \
               _____              _____
              /     \            /     \
             /  -1   \          /   0   \
            /_____\        /_____\
            /         \        /         \
       _____   _____  _____   _____
      |          | |          ||          | |          |
      |  -100    | |    10    ||    0     | |    0     |
      |_____| |_____||_____| |_____|
```

(b) Should Valerie eat the cookie? ◯ Yes   ● No

(c) Valerie can now smell the cookie to judge whether it has raisins before she eats it. However, since she dislikes raisins she does not have much experience with them and cannot recognize their smell well. As a result she will incorrectly identify raisins when there are no raisins with probability 0.2 and will incorrectly identify no raisins when there are raisins with probability 0.3. This decision network can be represented by the diagram below where E is her choice to eat, U is her utility earned, R is whether the cookie contains raisins, and S is her attempt at smelling.

```
        ┌───┐
        │ E │
        └───┘
          │
          ▼
        ◇───◇        ◯        ◯
        │ U │◄───────│ R │───►│ S │
        ◇───◇        ◯        ◯
```

Valerie has just smelled the cookie and she thinks it doesn't have raisins. Write the probability, X, that the cookie has raisins given that she smelled no raisins as a simplest form fraction or decimal.

$X = \boxed{0.04}$

$P(+r \mid -s) = \frac{P(-s \mid +r)P(+r)}{P(-s)} = \frac{P(-s \mid +r)P(+r)}{P(-s \mid +r)P(+r) + P(-s \mid -r)P(-r)} = \frac{.3 * .1}{.3 * .1 + .8 * .9} = \frac{.03}{.75} = .04$

(d) What is her maximum expected utility, Y given that she smelled no raisins? You can answer in terms of X or as a simplest form fraction or decimal.

$Y = \boxed{-100X + 10(1 - X), 5.6}$

$MEU(-s) = max(MEU(eating \mid -s), MEU(noteating \mid -s)) =$
$max(P(+r \mid -s) * EU(eating, +r) + P(-r \mid -s) * EU(eating, -r), MEU(noteating)) =$
$max(X * (-100) + (1 - X) * 10, 0) =$
$X * 100 + (1 - X) * 10$

(e) What is the Value of Perfect Information (VPI) of smelling the cookie? You can answer in terms of X and Y or as a

simplest form fraction or decimal.

$VPI =$ | $\quad 0.75 * Y, 4.2$ |

$VPI(S) = MEU(S) - MEU(\emptyset)$
$MEU(S) = P(-s)MEU(-s) + P(+s)MEU(+s)$
$P(-s) = .75$ from part (c), $MEU(-s) = Y$
$MEU(+s) = 0$ because it was better for her to not eat the raisin without knowing anything, smelling raisins will only make it more likely for the cookie to have raisins and it will still be best for her to not eat and earn a utility of 0. Note this means we do not have to calculate P(+s).
$MEU(\emptyset) = 0$
$VPI(S) = .75 * Y + 0 - 0 = .75 * Y$

(f) Valerie is unsatisfied with the previous model and wants to incorporate more variables into her decision network. First, she realizes that the air quality (A) can affect her smelling accuracy. Second, she realizes that she can question (Q) the people around to see if they know where the cookie came from. These additions are reflected in the decision network below.
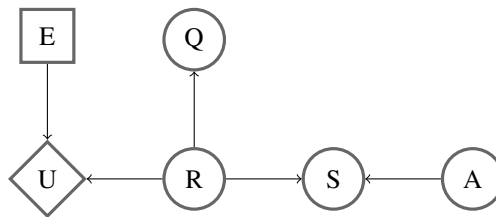
Choose one for each equation:

	Could Be True	Must Be True	Must Be False
$VPI(A, S) > VPI(A) + VPI(S)$	●	○	○
$VPI(A) = 0$	○	●	○
$VPI(Q, R) \leq VPI(Q) + VPI(R)$	○	●	○
$VPI(S, R) > VPI(R)$	○	○	●
$VPI(Q) \geq 0$	○	●	○
$VPI(Q, A) > VPI(Q)$	○	○	●
$VPI(S|A) < VPI(S)$	○	○	●
$VPI(A|S) > VPI(A)$	●	○	○