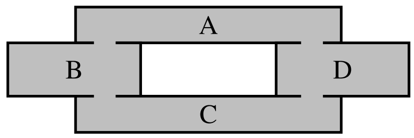
CS 188 Summer 2021 Introduction to Artificial Intelligence **Exam Prep 5B Solutions**

Q1. HMMs: Help Your House Help You

Imagine you have a smart house that wants to track your location within itself so it can turn on the lights in the room you are in and make you food in your kitchen. Your house has 4 rooms (A, B, C, D) in the floorplan below (A is connected to B and D, B is connected to A and C, C is connected to B and D, and D is connected to A and C):



At the beginning of the day ($t = 0$), your probabilities of being in each room are p_A, p_B, p_C, and p_D for rooms A, B, C, and D, respectively, and at each time t your position (following a Markovian process) is given by X_t. At each time, your probability of staying in the same room is q_0, your probability of moving clockwise to the next room is q_1, and your probability of moving counterclockwise to the next room is $q_{-1} = 1 - q_0 - q_1$.

(a) Initially, assume your house has no way of sensing where you are. What is the probability that you will be in room D at time $t = 1$?

○ $q_0 p_D$ ○ $q_0 p_D + q_1 p_A + q_{-1} p_C + 2q_1 p_B$ ● $q_0 p_D + q_1 p_A + q_{-1} p_C$

○ $q_0 p_D + q_{-1} p_A + q_1 p_C$ ○ $q_1 p_A + q_1 p_C + q_0 p_D$ ○ None of these

This probability is given by the sum of three probabilities: 1) $q_0 p_D$: You are in room D to start (p_D) and stay there (q_0), 2) $q_1 p_A$: You are in room A to start (p_A) and move clockwise to room D (q_1), and 3) $q_{-1} p_C$: You are in room C to start (p_C) and move counterclockwise to room D (q_{-1}).

Now assume your house contains a sensor M^A that detects motion ($+m$) or no motion ($-m$) in room A. However, the sensor is a bit noisy and can be tricked by movement in adjacent rooms, resulting in the conditional distributions for the sensor given in the table below. The prior distribution for the sensor's output is also given.

M^A	$P(M^A \mid X = A)$	$P(M^A \mid X = B)$	$P(M^A \mid X = C)$	$P(M^A \mid X = D)$
$+m^A$	$1 - 2\gamma$	γ	0.0	γ
$-m^A$	2γ	$1 - \gamma$	1.0	$1 - \gamma$

M^A	$P(M^A)$
$+m^A$	0.5
$-m^A$	0.5

(b) You decide to help your house to track your movements using a particle filter with three particles. At time $t = T$, the particles are at $X^0 = A, X^1 = B, X^2 = D$. What is the probability that the particles will be resampled as $X^0 = X^1 = X^2 = A$ after time elapse? Select **all terms in the product.**

■ q_0 ☐ q_0^2 ☐ q_0^3 ■ q_1 ☐ q_1^2 ☐ q_1^3 ■ q_{-1} ☐ q_{-1}^2 ☐ q_{-1}^3 ○ None of these

The probability that all particles will be resampled as being in room A is $q_0 q_1 q_{-1}$ since particle X^0 stays in A with probability q_0, particle X^1 moves clockwise to A with probability q_1, and particle X^2 moves counterclockwise with probability q_{-1}.

(c) Assume that the particles are actually resampled after time elapse as $X^0 = D, X^1 = B, X^2 = C$, and the sensor observes

$M^A = -m^A$. What are the particle weights given the observation?

Particle	Weight
$X^0 = D$	○ γ ● $1-\gamma$ ○ $1-2\gamma$ ○ 0.0 ○ 1.0 ○ 2γ ○ None of these
$X^1 = B$	○ γ ● $1-\gamma$ ○ $1-2\gamma$ ○ 0.0 ○ 1.0 ○ 2γ ○ None of these
$X^2 = C$	○ γ ○ $1-\gamma$ ○ $1-2\gamma$ ○ 0.0 ● 1.0 ○ 2γ ○ None of these

We can read these weights off of the tables given above. The weight for X^0 is given by $P(M^A = -m^A|X = D) = 1-\gamma$, the weight for X^1 is given by $P(M^A = -m^A|X = B) = 1-\gamma$, and the weight for X^2 is given by $P(M^A = -m^A|X = C) = 1$.

Now, assume your house also contains sensors M^B and M^D in rooms B and D, respectively, with the conditional distributions of the sensors given below and the prior equivalent to that of sensor M^A.

M^B	$P(M^B \mid X = A)$	$P(M^B \mid X = B)$	$P(M^B \mid X = C)$	$P(M^B \mid X = D)$
$+m^B$	γ	$1-2\gamma$	γ	0.0
$-m^B$	$1-\gamma$	2γ	$1-\gamma$	1.0

M^D	$P(M^D \mid X = A)$	$P(M^D \mid X = B)$	$P(M^D \mid X = C)$	$P(M^D \mid X = D)$
$+m^D$	γ	0.0	γ	$1-2\gamma$
$-m^D$	$1-\gamma$	1.0	$1-\gamma$	2γ

(d) Again, assume that the particles are actually resampled after time elapse as $X^0 = D, X^1 = B, X^2 = C$. The sensor readings are now $M^A = -m^A, M^B = -m^B, M^D = +m^D$. What are the particle weights given the observations?

Particle	Weight
$X^0 = D$	○ $\gamma^2 - 2\gamma^3$ ○ $3-2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ● $1-3\gamma+2\gamma^2$ ○ $2-\gamma$ ○ $1-2\gamma+\gamma^2$ ○ None of these
$X^1 = B$	○ $\gamma^2 - 2\gamma^3$ ○ $3-2\gamma$ ● 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1-3\gamma+2\gamma^2$ ○ $2-\gamma$ ○ $1-2\gamma+\gamma^2$ ○ None of these
$X^2 = C$	○ $\gamma^2 - 2\gamma^3$ ○ $3-2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1-3\gamma+2\gamma^2$ ○ $2-\gamma$ ○ $1-2\gamma+\gamma^2$ ● None of these

The weight for X^0 is given by $P(M^A = -m^A|X = D)P(M^B = -m^B|X = D)P(m^D = +m^D|X = D) = (1-\gamma)(1.0)(1-2\gamma) = 1-3\gamma+2\gamma^2$, the weight for X^1 is given by $P(M^A = -m^A|X = B)P(M^B = -m^B|X = B)P(M^D = +m^D|X = B) = (1-\gamma)(2\gamma)(0.0) = 0.0$, and the weight for X^2 is given by $P(M^A = -m^A|X = C)P(M^B = -m^B|X = C)P(M^A = +m^D|X = C) = (1.0)(1-\gamma)(\gamma) = \gamma - \gamma^2$.

The sequence of observations from each sensor are expressed as the following: $m^A_{0:t}$ are all measurements $m^A_0, m^A_1, \ldots, m^A_t$ from sensor M^A, $m^B_{0:t}$ are all measurements $m^B_0, m^B_1, \ldots, m^B_t$ from sensor M^B, and $m^D_{0:t}$ are all measurements $m^D_0, m^D_1, \ldots, m^D_t$ from sensor M^D. Your house can get an accurate estimate of where you are at a given time t using the forward algorithm. The forward algorithm update step is shown here:

$$P(X_t \mid m^A_{0:t}, m^B_{0:t}, m^D_{0:t}) \propto P(X_t, m^A_{0:t}, m^B_{0:t}, m^D_{0:t}) \tag{1}$$

$$= \sum_{x_{t-1}} P(X_t, x_{t-1}, m^A_t, m^B_t, m^D_t, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) \tag{2}$$

$$= \sum_{x_{t-1}} \_\_\_\_\_\_\_\_ P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) \tag{3}$$

(e) Which of the following expression(s) correctly complete the missing expression in line (3) above (regardless of whether they are available to the algorithm during execution)? Fill in **all** that apply.

■ $P(m^A_t, m^B_t, m^D_t \mid X_t, x_{t-1})$ ☐ $P(m^A_t, m^B_t, m^D_t \mid x_{t-1})$ ☐ $P(m^A_t \mid x_{t-1}) P(m^B_t \mid x_{t-1}) P(m^D_t \mid x_{t-1})$

☐ $P(m^A_t \mid m^A_{0:t-1}) P(m^B_t \mid m^B_{0:t-1}) P(m^D_t \mid m^D_{0:t-1})$ ■ $P(m^A_t, m^B_t, m^D_t \mid X_t, x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1})$

■ $P(m^A_t \mid X_t) P(m^B_t \mid X_t) P(m^D_t \mid X_t)$ ■ $P(m^A_t, m^B_t, m^D_t \mid X_t)$ ○ None of these

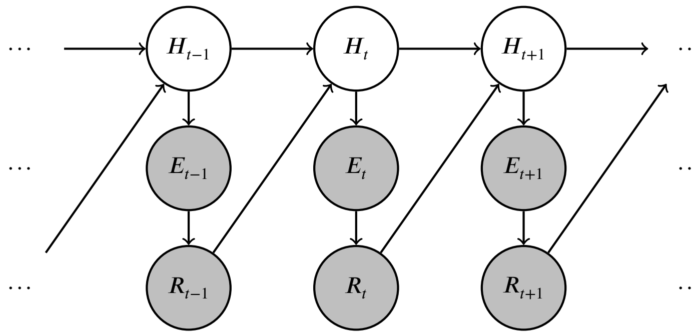
Using the chain rule from the previous step,

$$\begin{aligned} P(X_t, x_{t-1}, m^A_t, m^B_t, m^D_t, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) &= [P(m^A_t, m^B_t, m^D_t \mid X_t, x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) \\ &\quad P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1} m^D_{0:t-1})] \\ &= [P(m^A_t, m^B_t, m^D_t \mid X_t, x_{t-1}) \\ &\quad P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1} m^D_{0:t-1})] \\ &\quad \text{(indep. of measurements from prev. measurements)} \\ &= [P(m^A_t, m^B_t, m^D_t \mid X_t) \\ &\quad P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1} m^D_{0:t-1})] \\ &\quad \text{(indep. of measurements from prev. states)} \\ &= [P(m^A_t \mid X_t) P(m^B_t \mid X_t) P(m^D_t \mid X_t) \\ &\quad P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1} m^D_{0:t-1})] \\ &\quad \text{(indep. of measurements from each other)} \end{aligned}$$

All of the expressions on the right side of the above equations should be selected.

Q2. HMM: Human-Robot Interaction

In the near future, autonomous robots would live among us. Therefore, it is important for the robots to know how to properly act in the presence of humans. In this question, we are exploring a simplified model of this interaction. Here, we are assuming that we can observe the robot's actions at time t, R_t, and an evidence observation, E_t, directly caused by the human action, H_t. Human's actions and Robot's actions from the past time-step affect the Human's and Robot's actions in the next time-step. In this problem, we will remain consistent with the convention that capital letters (H_t) refer to random variables and lowercase letters (h_t) refer to a particular value the random variable can take. The structure is given below:



You are supplied with the following probability tables: $P(R_t \mid E_t)$, $P(H_t \mid H_{t-1}, R_{t-1})$, $P(H_0)$, $P(E_t \mid H_t)$.

Let us derive the forward algorithm for this model. We will split our computation into two components, a **time-elapse update** expression and a **observe update** expression.

(a) We would like to incorporate the evidence that we observe at time t. Using the time-lapse update expression we will derive separately, we would like to find the **observe update** expression:

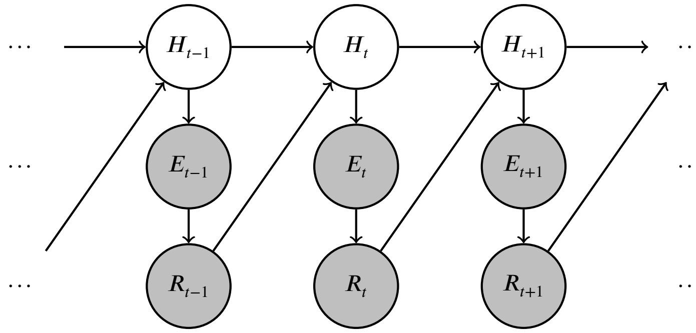
$$O(H_t) = P(H_t | e_{0:t}, r_{0:t})$$

In other words, we would like to compute the distribution of potential human states at time t given all observations up to and including time t. In addition to the conditional probability tables associated with the network's nodes, we are given $T(H_t) = P(H_t \mid e_{0:t-1}, r_{0:t-1})$, which we will assume is correctly computed in the time-elapse update that we will derive in the next part. From the options below, select *all* the options that **both** make valid independence assumptions and would evaluate to the observe update expression.

- [x] $\frac{P(H_t|e_{0:t-1}, r_{0:t-1})P(e_t|H_t)P(r_t|e_t)}{\sum_{h_t} P(h_t|e_{0:t-1}, r_{0:t-1})P(e_t|h_t)P(r_t|e_t)}$
- [x] $\frac{P(H_t|e_{0:t-1}, r_{0:t-1})P(e_t|H_t)}{\sum_{h_t} P(h_t|e_{0:t-1}, r_{0:t-1})P(e_t|h_t)}$
- [ ] $\frac{\sum_{e_t} P(H_t|e_{0:t-1}, r_{0:t-1})P(e_t|H_t)}{\sum_{h_t} P(h_t|e_{0:t-1}, r_{0:t-1})P(e_t|r_{t-1}, H_{t-1})}$
- [ ] $\sum_{r_{t-1}} P(H_t \mid e_{0:t-1}, r_{0:t-1})P(r_{t-1} \mid e_{t-1})$
- [ ] $\sum_{r_t} P(H_t \mid e_{0:t-1}, r_{0:t-1})P(r_t \mid r_{t-1}, e_t)$
- [ ] $\sum_{h_{t+1}} P(H_t \mid e_{0:t-1}, r_{0:t-1})P(h_{t+1} \mid r_t)$

$$P(H_t|e_{0:t}, r_{0:t}) = \frac{P(H_t, e_{0:t}, e_{0:t})}{\sum_{h_t} P(h_t, e_{0:t}, e_{0:t})} = \frac{P(H_t \mid e_{0:t-1}, r_{0:t-1})P(e_t \mid H_t)P(r_t \mid e_t)}{P(r_t \mid e_t) \sum_{h_t} P(h_t \mid e_{0:t-1}, r_{0:t-1})P(e_t \mid h_t)}$$

The structure below is identical to the one in the beginning of the question and is repeated for your convenience.



(b) We are interested in predicting what the state of human is at time t (H_t), given all the observations through $t-1$. Therefore, the **time-elapse update** expression has the following form:

$$T(H_t) = P(H_t|e_{0:t-1}, r_{0:t-1})$$

Derive an expression for the given time-elapse update above using the probability tables provided in the question and the observe update expression, $O(H_{t-1}) = P(H_{t-1}|e_{0:t-1}, r_{0:t-1})$. Write your final expression in the space provided at below. You may use the function O in your solution if you prefer.

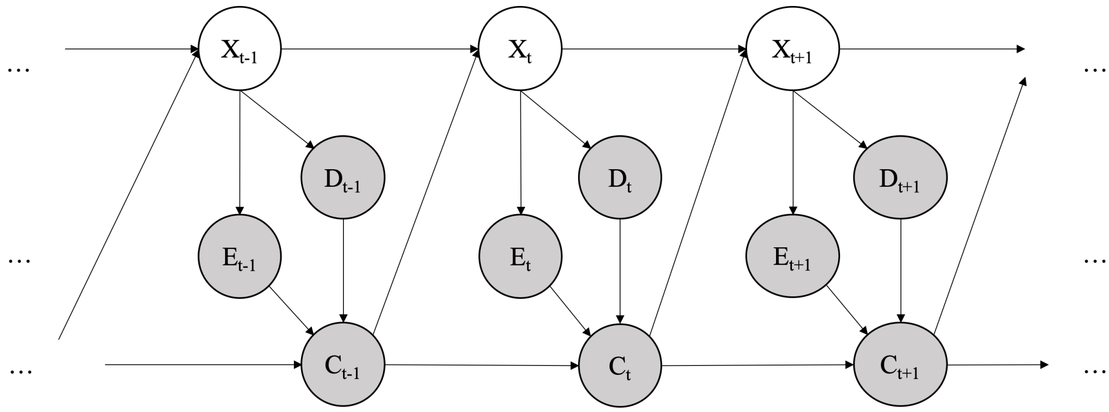
The derivation of the time-elapse update for this setup is similar to the one we have seen in lecture; however, here, we have additional observations and dependencies.

$$\begin{aligned} P(H_t|e_{0:t-1}, r_{0:t-1}) &= \sum_{h_{t-1}} P(H_t, h_{t-1} \mid e_{0:t-1}, r_{0:t-1}) \\ &= \sum_{h_{t-1}} P(H_t \mid h_{t-1}, r_{t-1}) P(h_{t-1} \mid e_{0:t-1}, r_{0:t-1}) \end{aligned}$$

$P(H_t|e_{0:t-1}, r_{0:t-1}) =$ $\sum_{h_{t-1}} P(H_t \mid h_{t-1}, r_{t-1}) P(h_{t-1} \mid e_{0:t-1}, r_{0:t-1})$

Q3. [Timed: 15mins] We Are Getting Close...

The CS 188 TAs have built an autonomous vehicle, and it's finally on the street! Approaching a crossroad, our vehicle must avoid bumping into pedestrians. However, how close are we?
X is the signal received from sensors on our vehicle. We have a estimation model E, which estimates the current distance of any object in our view. Our vehicle also needs a model to detect objects and label their classes as one of {pedestrian, stop sign, road, other}. The TAs trained a detection model D that does the above and with a simple classification, outputs one of {no pedestrian, pedestrian on the road, pedestrian beside the stop sign}. Our vehicle has a control operator C, which determines the velocity by changing the acceleration.



(a) For the above Dynamic Bayes Net, complete the equations for performing updates. (Hint: think about the prediction update and observation update equations in the forward algorithm for HMMs.)

Time elapse: **(i)** = **(ii)** **(iii)** **(iv)** $P\left(x_{t-1} | e_{0:t-1}, d_{0:t-1}, c_{0:t-1}\right)$

(i) ○ $P(x_t)$ ● $P\left(x_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1}\right)$ ○ $P\left(e_t, d_t, c_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1}\right)$

(ii) ○ $P(c_{0:t-1})$ ○ $P(x_{0:t-1}, c_{0:t-1})$ ○ $P(e_{0:t-1}, d_{0:t-1}, c_{0:t-1})$
○ $P(e_{0:t}, d_{0:t}, c_{0:t})$ ● 1

(iii) ● $\Sigma_{x_{t-1}}$ ○ Σ_{x_t} ○ $\max_{x_{t-1}}$ ○ $\max_{x_t}$ ○ 1

(iv) ○ $P(x_{t-1} | x_{t-2})$ ○ $P(x_{t-1}, x_{t-2})$ ○ $P(x_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1})$
○ $P(x_t | x_{t-1})$ ○ $P(x_t, x_{t-1})$ ○ $P(x_t, e_{0:t-1}, d_{0:t-1}, c_{0:t-1})$
● $P(x_t | x_{t-1}, c_{t-1})$ ○ $P(x_t, x_{t-1}, c_{t-1})$ ○ 1

Recall the prediction update of forward algorithm: $P(x_t | o_{0:t-1}) = \Sigma_{x_{t-1}} P(x_t | x_{t-1}) P\left(x_{t-1} | o_{0:t-1}\right)$, where o is the observation. Here it is similar, despite that there are several observations at each time, which means o_t corresponds to e_t, d_t, c_t for each t, and that X is dependent on the C value of the previous time, so we need $P(x_t | x_{t-1}, c_{t-1})$ instead of $P(x_t | x_{t-1})$. Also note that X is independent of D_{t-1}, E_{t-1} given C_{t-1}, X_{t-1}.

Update to incorporate new evidence at time t:
$P\left(x_t | e_{0:t}, d_{0:t}, c_{0:t}\right)$ = **(v)** **(vi)** **(vii)** **Your choice for (i)**

(v) ○ $\left(P\left(c_t | c_{0:t-1}\right)\right)^{-1}$ ○ $\left(P\left(e_t | e_{0:t-1}\right) P\left(d_t | d_{0:t-1}\right) P\left(c_t | c_{0:t-1}\right)\right)^{-1}$
● $\left(P\left(e_t, d_t, c_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1}\right)\right)^{-1}$ ○ $\left(P\left(e_{0:t-1} | e_t\right) P\left(d_{0:t-1} | d_t\right) P\left(c_{0:t-1} | c_t\right)\right)^{-1}$
○ $\left(P\left(e_{0:t-1}, d_{0:t-1}, c_{0:t-1} | e_t, d_t, c_t\right)\right)^{-1}$ ○ 1

(vi) ○ $\Sigma_{x_{t-1}}$ ○ Σ_{x_t} ○ Σ_{x_{t-1}, x_t} ○ $\max_{x_{t-1}}$ ○ $\max_{x_t}$ ● 1

(vii) ☐ $P(x_t | e_t, d_t, c_t)$ ☐ $P(x_t, e_t, d_t, c_t)$
☐ $P(x_t | e_t, d_t, c_t, c_{t-1})$ ☐ $P(x_t, e_t, d_t, c_t, c_{t-1})$
■ $P(e_t, d_t | x_t) P(c_t | e_t, d_t, c_{t-1})$ ☐ $P(e_t, d_t, c_t | x_t)$ ○ 1

Recall the observation update of forward algorithm: $P(x_t | o_{0:t}) \propto P(x_t, o_t | o_{0:t-1}) = P(o_t | x_t) P(x_t | o_{0:t-1})$.
Here the observations o_t corresponds to e_t, d_t, c_t for each t. Apply the Chain Rule, we are having
$P\left(x_t | e_{0:t}, d_{0:t}, c_{0:t}\right) \propto P\left(x_t, e_t, d_t, c_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1}\right) = P(e_t, d_t, c_t | x_t, c_{t-1}) P(x_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1})$

$= P(e_t, d_t | x_t) P(c_t | e_t, d_t, c_{t-1}) P(x_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1})$.
Note that in $P(e_t, d_t, c_t | x_t, c_{t-1})$, we cannot omit c_{t-1} due to the arrow between c_t and c_{t-1}.
To calculate the normalizing constant, use Bayes Rule: $P\left(x_t | e_{0:t}, d_{0:t}, c_{0:t}\right) = \frac{P(x_t, e_t, d_t, c_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1})}{P(e_t, d_t, c_t | e_{0:t-1}, d_{0:t-1}, c_{0:t-1})}$.

(viii) Suppose we want to do the above updates in one step and use normalization to reduce computation. Select all the terms that are not explicitly calculated in this implementation.
DO **NOT** include the choices if their values are 1.

☐ **(ii)** ☐ **(iii)** ☐ **(iv)** ■ **(v)** ☐ **(vi)** ☐ **(vii)** ○ None of the above

(v) is a constant, so we don't calculate it during implementation and simply do a normalization instead. Everything else is necessary.

(b) Suppose X outputs 1024×1024 greyscale images and our vehicle stays stationary. As before, E includes precise estimation of the distance between our vehicle and the pedestrian evaluated from outputs of X. Unfortunately, a power outage happened, and before the power is restored, E will not be available for our vehicle. But we still have the detection model D, which outputs one of {no pedestrian, pedestrian on the road, pedestrian beside the stop sign} for each state.

(i) During the power outage, it is best to
- ○ do particle filtering because the particles are easier to track for D than for both D and E
- ● do particle filtering because of memory constraints
- ○ do particle filtering, but not for the reasons above
- ○ do exact inference because it saves computation
- ○ do exact inference, but not for the reason above

E is unavailable and C does not change its value since our vehicle stays stationary, so we only considers X and D. Although D has only 3 possible values, X is huge and it is impossible to store the belief distribution.

(ii) The power outage was longer than expected. As the sensor outputs of X have degraded to 2×2 binary images, it is best to
- ○ do particle filtering because the particles are easier to track for D than for both D and E
- ○ do particle filtering because of memory constraints
- ○ do particle filtering, but not for the reasons above
- ○ do exact inference because it saves computation
- ● do exact inference, but not for the reason above

In this case we do not have the "X is huge" problem in **(i)**, and we can do exact inference, which is always more accurate than particle filtering and thus more favorable in this setting.

(iii) After power is restored and we have E, it is reasonable to
- ● do particle filtering because of memory constraints
- ○ do particle filtering, but not for the reason above
- ○ do exact inference because E gives more valuable information than D
- ○ do exact inference because it's impractical to do particle filtering for E
- ○ do exact inference, but not for the reasons above

The belief distribution is too big to store in memory.

(c) Now we formulate the Dynamic Bayes Net for this question into a non-deterministic two-player game (analogous to MDP in single-player setting). Each state $S = (X, E, D)$.
There are 2 agents in the game: our vehicle (with action set A), and a pedestrian (with action set B). **The vehicle and the pedestrian take turns to perform their actions**.

The TAs implemented 3 modes for the autonomous vehicle to act in the same space with the kind pedestrian, the confused pedestrian, and the naughty pedestrian. In each round of testing, a TA will be the pedestrian, and one of the modes will be tested. The vehicle and the pedestrian are both in the corresponding mode.
Below, Q_v^* is the Q-function for the autonomous vehicle. For each subquestion, given the standard notation for an MDP, select the expression f_n that would complete the blank part of the Q-Value Iteration under the specified formulation.

$Q_v^*(s,a) = \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma \_\_\_\_\_\_\_\_\_\_]$

$$f_1 = \sum_{b \in B} \sum_{s''} \left(T\left(s', b, s''\right)\left[R\left(s', b, s''\right) + \gamma \max_{a' \in A} Q_v^*\left(s'', a'\right)\right]\right)$$

$$f_2 = \max_{b \in B} \sum_{s''} \left(T\left(s', b, s''\right)\left[R\left(s', b, s''\right) + \gamma \max_{a' \in A} Q_v^*\left(s'', a'\right)\right]\right)$$

$$f_3 = \min_{b \in B} \sum_{s''} \left(T\left(s', b, s''\right)\left[R\left(s', b, s''\right) + \gamma \max_{a' \in A} Q_v^*\left(s'', a'\right)\right]\right)$$

$$f_4 = \sum_{b \in B} \sum_{s''} \left(T\left(s', b, s''\right)\left[R\left(s', b, s''\right) + \gamma \frac{1}{|B|} \max_{a' \in A} Q_v^*\left(s'', a'\right)\right]\right)$$

$$f_5 = \max_{a' \in A} Q_v^*\left(s', a'\right)$$

$$f_6 = \min_{a' \in A} Q_v^*\left(s', a'\right)$$

$$f_7 = \frac{1}{|A|} \sum_{a' \in A} Q_v^*\left(s', a'\right)$$

(i) The kind pedestrian that acts friendly, maximizing the vehicle's utility.
☐ f_1 ■ f_2 ☐ f_3 ☐ f_4 ☐ f_5 ☐ f_6 ☐ f_7 ○ None of the above

(ii) The confused pedestrian that acts randomly.
☐ f_1 ☐ f_2 ☐ f_3 ☐ f_4 ■ f_5 ☐ f_6 ☐ f_7 ○ None of the above

(iii) The naughty pedestrian that performs adversarial actions, minimizing the vehicle's utility.
☐ f_1 ☐ f_2 ■ f_3 ☐ f_4 ☐ f_5 ☐ f_6 ☐ f_7 ○ None of the above

Recall the q-value iteration formula: $Q_{k+1}(s,a) \leftarrow \sum_{s'} T\left(s,a,s'\right)\left[R\left(s,a,s'\right) + \gamma \max_{a'} Q_k\left(s',a'\right)\right]$.
Here it is similar, but in addition to the vehicle, there's also a pedestrian taking actions from a different set, so we need to do something similar for the pedestrian, as in f_1, f_2, f_3, f_4. That is, instead of using the maximum Q_v right away, we substitute that with the q-value iteration formula for the pedestrian with respect to Q_v.
The value of this formula is maximized (as in f_2) in the case of the kind pedestrian,
minimized (as in f_3) in the case of the naughty pedestrian,
and averaged in the case of the naughty pedestrian, which would be
$\frac{1}{|B|} \sum_{b \in B} \sum_{s''} \left(T\left(s', b, s''\right)\left[R\left(s', b, s''\right) + \gamma \max_{a' \in A} Q_v^*\left(s'', a'\right)\right]\right)$.
Hence f_1 and f_4 are incorrect. Since the pedestrian is acting completely randomly, we can include the pedestrian in the transition dynamics and just use regular q-value iteration, which is f_5.