CS 188
Summer 2022

Regular Discussion 5B Solutions

Q1. VPI

You are the latest contestant on Monty Hall's game show. In the game, there are n closed doors: behind one door is a car ($U(car) = 1000$), while the other $n-1$ doors each have a goat behind them ($U(goat) = 10$). You are permitted to open exactly one door and claim the prize behind it. You begin by choosing a door uniformly at random.

(a) What is your expected utility? ($1000 * \frac{1}{n} + 10 * \frac{n-1}{n} = 10 + 990 * \frac{1}{n}$)
We can calculate the expected utility via the usual formula of expectation, or we can note that there is a guaranteed utility of 10, with a small probability of a bonus utility. The latter is a bit simpler, so the answers to the following parts use this form.

(b) After you choose a door but before you open it, Monty offers to open k other doors, each of which are guaranteed to have a goat behind it. If you accept this offer, should you keep your original choice of a door, or switch to a new door?

$EU(keep)$: $10 + 990 * \frac{1}{n}$

$EU(switch)$: $10 + 990 * \frac{(n-1)}{n*(n-k-1)}$

Action that achieves MEU: switch

The expected utility if we keep must be the same as the answer from the previous part: the probability that we have a winning door has not changed at all, since we have gotten no meaningful information.
In order to win a car by switching, we must have chosen a goat door previously (probability $\frac{n-1}{n}$) and then switch to the car door (probability $\frac{1}{n-k-1}$).
Since $n-1 > n-k-1$ for positive k, switching gets a larger expected utility.

(c) What is the value of the information that Monty is offering you? $990 * \frac{1}{n} * \frac{k}{n-k-1}$
The formula for VPI is $MEU(e) - MEU(\emptyset)$. Thus, we want the difference between $EU(switch)$ (the optimal action if Monty opens the doors) and our expected utility from part (a).
(It is true that $EU(keep)$ happens to have the same numeric expression as in part (a), but this fact is not meaningful in answering this part.)

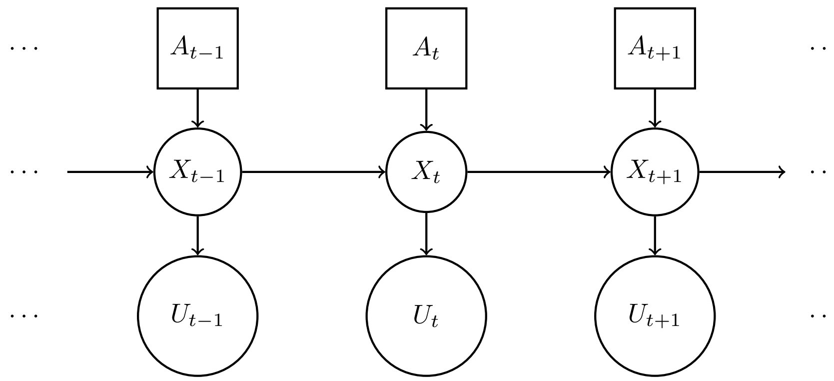
(d) Monty is changing his offer! After you choose your initial door, you are given the offer to choose any other door and open this second door. If you do, after you see what is inside the other door, you may switch your initial choice (to the newly opened door) or keep your initial choice. What is the value of this new offer? $\frac{990}{n}$
Intuitively, if we take this offer, it is as if we just chose two doors in the beginning, and we win if either door has the car behind it. Unlike in the previous parts, if the new door has a goat behind it, it is not more optimal to switch doors.
Mathematically, letting D_i be the event that door i has the car, we can calculate this as $P(D_2 \cup D_1) = P(D_1) + P(D_2) = 1/n + 1/n = 2/n$, to see that $MEU(\text{offer}) = 10 + 990 * \frac{2}{n}$. Subtracting the expected utility without taking the offer, we are left with $990 * \frac{1}{n}$.

(e) Monty is generalizing his offer: you can pay $\$d^3$ to open d doors as in the previous part. (Assume that $U(\$x) = x$.) You may now switch your choice to any of the open doors (or keep your initial choice). What is the largest value of d for which it would be rational to accept the offer? $d = \sqrt{\frac{990}{n}}$
It is a key insight (whether intuitive or determined mathematically) that the answer to the previous part

is constant for each successive door we open. Thus, the value of opening d doors is just $d * 990 * \frac{1}{n}$. Setting this equal to d^3, we can solve for d.

Q2. 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:

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

It is always possible that VPI $= 0$. Can guarantee $\text{VPI}(E|e)$ 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.