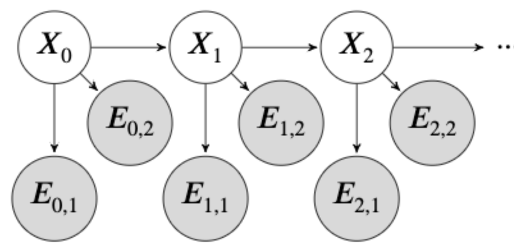
CS 188 Spring 2021 — Introduction to Artificial Intelligence — Exam Prep 8 Solutions

Q1. Particle Filtering

You've chased your arch-nemesis Leland to the Stanford quad. You enlist two robo-watchmen to help find him! The grid below shows the campus, with ID numbers to label each region. Leland will be moving around the campus. His location at time step t will be represented by random variable X_t. Your robo-watchmen will also be on campus, but their locations will be fixed. Robot 1 is always in region 1 and robot 2 is always in region 9. (See the * locations on the map.) At each time step, each robot gives you a sensor reading to help you determine where Leland is. The sensor reading of robot 1 at time step t is represented by the random variable $E_{t,1}$. Similary, robot 2's sensor reading at time step t is $E_{t,2}$. The Bayes' Net to the right shows your model of Leland's location and your robots' sensor readings.

1*	2	3	4	5
6	7	8	9*	10
11	12	13	14	15



In each time step, Leland will either stay in the same region or move to an adjacent region. For example, the available actions from region 4 are (WEST, EAST, SOUTH, STAY). He chooses between all available actions with equal probability, regardless of where your robots are. Note: moving off the grid is not considered an available action.

Each robot will detect if Leland is in an adjacent region. For example, the regions adjacent to region 1 are 1, 2, and 6. If Leland is in an adjacent region, then the robot will report $NEAR$ with probability 0.8. If Leland is not in an adjacent region, then the robot will still report $NEAR$, but with probability 0.3.

For example, if Leland is in region 1 at time step t the probability tables are:

E	$P(E_{t,1} \mid X_t = 1)$	$P(E_{t,2} \mid X_t = 1)$
$NEAR$	0.8	0.3
FAR	0.2	0.7

(a) Suppose we are running particle filtering to track Leland's location, and we start at $t = 0$ with particles [$X = 6$, $X = 14$, $X = 9$, $X = 6$]. Apply a forward simulation update to each of the particles using the random numbers in the table below.

Assign region IDs to sample spaces in numerical order. For example, if, for a particular particle, there were three possible sucessor regions 10, 14 and 15, with associated probabilities, $P(X = 10)$, $P(X = 14)$ and $P(X = 15)$, and the random number was 0.6, then 10 should be selected if $0.6 \leq P(X = 10)$, 14 should be selected if $P(X = 10) < 0.6 < P(X = 10) + P(X = 14)$, and 15 should be selected otherwise.

Particle at $t = 0$	Random number for update	Particle after forward simulation update
$X = 6$	0.864	11
$X = 14$	0.178	9
$X = 9$	0.956	14
$X = 6$	0.790	11

(b) Some time passes and you now have particles $[X = 6, X = 1, X = 7, X = 8]$ at the particular time step, but you have not yet incorporated your sensor readings at that time step. Your robots are still in regions 1 and 9, and both report $NEAR$. What weight do we assign to each particle in order to incorporate this evidence?

Particle	Weight
$X = 6$	0.8 * 0.3
$X = 1$	0.8 * 0.3
$X = 7$	0.3 * 0.3
$X = 8$	0.3 * 0.8

(c) To decouple this question from the previous question, let's say you just incorporated the sensor readings and found the following weights for each particle (these are not the correct answers to the previous problem!):

Particle	Weight
$X = 6$	0.1
$X = 1$	0.4
$X = 7$	0.1
$X = 8$	0.2

Normalizing gives us the distribution

$$X = 1 : 0.4/0.8 = 0.5$$
$$X = 6 : 0.1/0.8 = 0.125$$
$$X = 7 : 0.1/0.8 = 0.125$$
$$X = 8 : 0.2/0.8 = 0.25$$

Use the following random numbers to resample you particles. As on the previous page, **assign region IDs to sample spaces in numerical order.**

Random number:	0.596	0.289	0.058	0.765
Particle:	6	1	1	8

Q2. Lotteries in Ghost Kingdom

(a) Diverse Utilities. Ghost-King (GK) was once great friends with Pacman (P) because he observed that Pacman and he shared the same preference order among all possible event outcomes. Ghost-King, therefore, assumed that he and Pacman shared the same utility function. However, he soon started realizing that he and Pacman had a different preference order when it came to lotteries and, alas, this was the end of their friendship.

Let Ghost-King and Pacman's utility functions be denoted by U_{GK} and U_P respectively. Assume both U_{GK} and U_P are guaranteed to output non-negative values.

(i) Which of the following relations between U_{GK} and U_P are consistent with Ghost King's observation that U_{GK} and U_P agree, with respect to all event outcomes but not all lotteries?

- ☐ $U_P = aU_{GK} + b \quad (0 < a < 1, b > 0)$
- ☐ $U_P = aU_{GK} + b \quad (a > 1, b > 0)$
- ■ $U_P = U_{GK}^2$
- ■ $U_P = \sqrt{(U_{GK})}$

For all the above options, U_P and U_{GK} result in the same preference order between two non-lottery events ($U_P(e_1) > U_P(e_2) \Leftrightarrow U_{GK}(e_1) > U_{GK}(e_2)$). However, options 1 and 2 also share the same preference order among all lotteries as well.

(ii) In addition to the above, Ghost-King also realized that Pacman was more risk-taking than him . Which of the relations between U_{GK} and U_P are possible?

- ☐ $U_P = aU_{GK} + b \quad (0 < a < 1, b > 0)$
- ☐ $U_P = aU_{GK} + b \quad (a > 1, b > 0)$
- ■ $U_P = U_{GK}^2$
- ☐ $U_P = \sqrt{(U_{GK})}$

As an example, say Ghost-King prefers winning \$2 as much as a lottery : winning \$0 or \$4 with equal probability. For option c), Pacman prefers the lottery much more (more risk-taking) and for option d), Pacman prefers the guaranteed reward.

(b) Guaranteed Return. Pacman often enters lotteries in the Ghost Kingdom. A particular Ghost vendor offers a lottery (for free) with three possible outcomes that are each equally likely: winning \$1, \$4, or \$5.

Let $U_P(m)$ denote Pacman's utility function for \$$m$. Assume that Pacman always acts rationally.

(i) The vendor offers Pacman a special deal - if Pacman pays \$1, the vendor will manipulate the lottery such that Pacman **always gets the highest reward possible**. For which of these utility functions would Pacman choose to pay the \$1 to the vendor for the manipulated lottery over the original lottery? (Note that if Pacman pays \$1 and wins \$$m$ in the lottery, his actual winnings are \$$m$-1.)

- ■ $U_P(m) = m$
- ■ $U_P(m) = m^2$

a) $U_P(m) = m$:
If pacman does not pay, expected utility $= \frac{1}{3}(5) + \frac{1}{3}(4) + \frac{1}{3}(1) = \frac{10}{3}$
If pacman pays up, expected utility $= 1(5-1) + 0(4-1) + 0(1-1) = 4$.

b) $U_P(m) = m^2$:
If pacman does not pay, expected utility $= \frac{1}{3}(5)^2 + \frac{1}{3}(4)^2 + \frac{1}{3}(1)^2 = 14$
If pacman pays up, expected utility $= 1(5-1)^2 + 0(4-1)^2 + 0(1-1)^2 = 16$.

(ii) Now assume that the ghost vendor can only manipulate the lottery such that Pacman **never gets the lowest reward** and the remaining two outcomes become equally likely. For which of these utility functions would Pacman choose to pay the \$1 to the vendor for the manipulated lottery over the original lottery?

■ $U_P(m) = m$

□ $U_P(m) = m^2$

a) $U_P(m) = m$:
If pacman does not pay, expected utility $= \frac{1}{3}(5) + \frac{1}{3}(4) + \frac{1}{3}(1) = \frac{10}{3}$
If pacman pays up, expected utility $= \frac{1}{2}(5-1) + \frac{1}{2}(4-1) + 0(1-1) = 3.5$.

b) $U_P(m) = m^2$:
If pacman does not pay, expected utility $= \frac{1}{3}(5)^2 + \frac{1}{3}(4)^2 + \frac{1}{3}(1)^2 = 14$
If pacman pays up, expected utility $= \frac{1}{2}(5-1)^2 + \frac{1}{2}(4-1)^2 + 0(1-1)^2 = 12.5$.

(c) Minimizing Other Utility.

The Ghost-King, angered by Pacman's continued winnings, decided to revolutionize the lotteries in his Kingdom. There are now 4 lotteries (A1, A2, B1, B2), each with two equally likely outcomes. Pacman, who wants to maximize his expected utility, can pick one of two lottery types (A, B). The ghost vendor thinks that Pacman's utility function is $U'_P(m) = m$ and minimizes accordingly. However, Pacman's real utility function $U_P(m)$ may be different.
For each of the following utility functions for Pacman, select the lottery corresponding to the outcome of the game. Note that Pacman knows how the ghost vendor is going to behave.

Pacman's expected utility for the 4 lotteries, under various utility functions, are as follows :
$U_P(m) = m$: [A1 : 10.5; A2 : 10; B1 : 9; B2 : 8]
$U_P(m) = m^2$: [A1 : 200.5; A2 : 100; B1 : 90; B2 : 113]
$U_P(m) = \sqrt{m}$: [A1 : 2.74; A2 : 3.16; B1 : 2.96; B2 : 2.44]

(i) $U_P(m) = m$:

○ A1 ● A2 ○ B1 ○ B2

(ii) $U_P(m) = m^2$:

○ A1 ○ A2 ○ B1 ● B2

(iii) $U_P(m) = \sqrt{m}$:

○ A1 ● A2 ○ B1 ○ B2

Since vendor minimizes $U'_P(m) = m$, If Pacman chooses A, vendor would pick A2 and if Pacman chooses B, vendor would pick B2. For $U_P(m) = m$ and $U_P(m) = \sqrt{m}$, Pacman prefers A2 over B2 and for $U_P(m) = m^2$, Pacman prefers B2 over A2 and acts accordingly.