1 Particle Filtering: Warmup

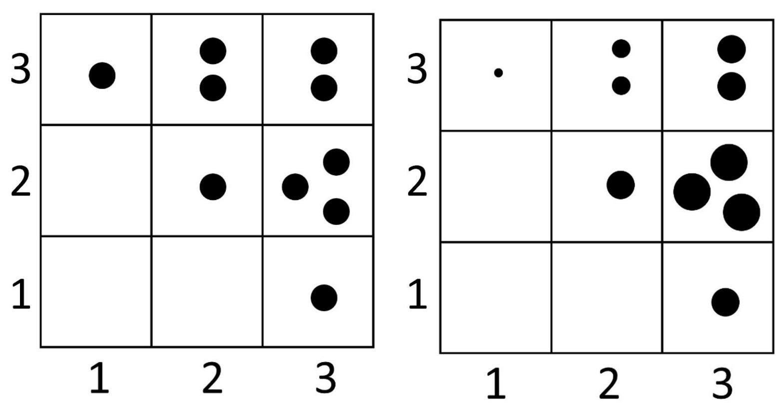
(a) **True / False:** The particle filtering algorithm is consistent since it gives correct probabilities as the number of samples N tends to infinity.

True

(b) **True / False:** The number of samples we use in the particle filtering algorithm increases from one time step to the next.

False. The number of samples stays constant from one time step to the next. The last step for each iteration of the algorithm is resampling, which builds a new population of N samples from the belief distribution updated by observation weights.

(c) The following state space contains 10 particles. The left grid shows the prior belief distribution of the particles at time t, while the grid on the right shows the particles weighted by the observations $P(e_t|S_t)$.



State	Weight
(1,3)	0.1
(2,2)	0.4
(2,3)	0.2
(3,1)	0.4
(3,2)	0.9
(3,3)	0.4

Fill in the following grids to update the belief distribution. Each square in the "Belief" grid should correspond to $\hat{P}(S_t|e_{1:t-1})$, the estimated probability of a particle being in state S at time t. Each square in the "Unnormalized" grid should correspond to the probability $P(S_t, e_t|e_{1:t-1})$. The "Normalized" grid should contain our updated belief distribution $\hat{P}(S_t|e_t, e_{1:t-1})$.

	1	2	3
3	/10	/10	/10
2	/10	/10	/10
1	/10	/10	/10

Belief

	1	2	3
3			
2			
1			

Unnormalized

	1	2	3
3			
2			
1			

Normalized

Solution: Note that states which did not appear in the weight table have a weight of 0.

3	1/10	2/10	2/10
2	0/10	1/10	3/10
1	0/10	0/10	1/10
	1	2	3

Belief

3	1/100	4/100	8/100
2	0	4/100	27/100
1	0	0	4/100
	1	2	3

Unnormalized

3	1/48	4/48	8/48
2	0	4/48	27/48
1	0	0	4/48
	1	2	3

Normalized

2 Tracking the Jabberwock

Lewis' Jabberwock is in the wild: its position is in a two-dimensional discrete grid, but this time the grid is not bounded. In other words, the position of the Jabberwock is a pair of integers $z = (x, y) \in \mathbb{Z}^2 = \{\dots, -2, -1, 0, 1, 2, \dots\} \times \{\cdots, -2, -1, 0, 1, 2, \cdots\}$. At each time step $t = 1, 2, 3, \dots$, the Jabberwock is in some cell $Z_t = z \in \mathbb{Z}^2$, and it moves to cell Z_{t+1} randomly as follows: with probability 1/2, it stays where it is; otherwise, it chooses one of its four neighboring cells uniformly at random (fortunately, no teleportation is allowed this week!).

(a) Write a function for the transition probability $P(Z_{t+1} = (x', y')|Z_t = (x, y))$.

$$P(Z_{t+1} = (x', y')|Z_t = (x, y)) = \begin{cases} \frac{1}{2} & \text{if } x = x', y = y' \\ \frac{1}{8} & \text{if } |x - x'| + |y - y'| = 1 \\ 0 & \text{otherwise} \end{cases}$$

We will use the particle filtering algorithm to track the Jabberwock. As a source of randomness use values in order from the following sequence $\{a_i\}_{1 \le i \le 14}$. Use these values to sample from any discrete distribution of the form $P(X)$ where X takes values in $\{1, 2, \dots, N\}$. Given $a_i \sim U[0, 1]$, return j such that $\sum_{k=1}^{j-1} P(X = k) \le a_i < \sum_{k=1}^{j} P(X = k)$. You have to fix an ordering of the elements for this procedure to make sense.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
0.142	0.522	0.916	0.792	0.703	0.231	0.036	0.859	0.677	0.221	0.156	0.249

At each time step t you get an observation of the x coordinate R_t in which the Jabberwock sits, but it is a noisy observation. Given the true position $Z_t = (x, y)$, you observe the correct value according to the following probability:

$$P(R_t = r|Z_t = (x, y)) \propto (0.5)^{|x-r|}$$

(b) Suppose that you know that half of the time, the Jabberwock starts at $z_1 = (0, 0)$, and the other half, at $z_1 = (1, 1)$. You get the following observations: $R_1 = 1, R_2 = 0, R_3 = 1$. Fill out the table for each time step using a particle filter with 2 particles to compute an approximation to $P(Z_1, Z_2, Z_3|r_1, r_2, r_3)$. Sample transitions from the table below using the a_i's as our source of randomness. The a_i's you should use for each step habe been indicated in the last row of each table. Note that going "left" decrements the x-coordinate by 1, and going "down" decrements the y-coordinate by 1.

[0; 0.5)	Stay
[0.5; 0.625)	Up
[0.625; 0.75)	Left
[0.75; 0.875)	Right
[0.875; 1)	Down

Initial	Belief $\hat{P}(z_1)$	Weights $P(r_1\|z_1)$	Unnormalized $\hat{P}(z_1, r_1)$	Normalized $\hat{P}(z_1\|r_1)$	Resampling
$p_1 =$ (0,0) $p_2 =$ (1,1) a_1, a_2	1/2 1/2				$p_1 =$ (,) $p_2 =$ (,) a_3, a_4

Transition $P(z_2\|z_1)$	Belief $\hat{P}(z_2\|r_1)$	Weights $P(r_2\|z_2)$	Unnormalized $\hat{P}(z_2, r_2\|r_1)$	Normalized $\hat{P}(z_2\|r_1, r_2)$	Resampling
$p_2 =$ (,) $p_2 =$ (,) a_5, a_6					$p_1 =$ (,) $p_2 =$ (,) a_7, a_8

Transition $P(z_3\|z_2)$	Belief $\hat{P}(z_3\|r_1, r_2)$	Weights $P(r_3\|z_3)$	Unnormalized $\hat{P}(z_3, r_3\|r_1, r_2)$	Normalized $\hat{P}(z_3\|r_1, r_2, r_3)$	Resampling
$p_1 =$ (,) $p_2 =$ (,) a_9, a_{10}					$p_1 =$ (,) $p_2 =$ (,) a_{11}, a_{12}

For each time step, we use our random numbers a_i to sample from the prior or from the transitions. Next, we find the weight of the sample based on the observation at that time step. We update our belief distribution with the weight by taking the product $\hat{P}(z_t|r_{1:t-1})P(r_t|z_t)$ and normalizing to get $\hat{P}(z_t|r_{1:t})$. Note that since the two particles are in different locations at each time step, the belief $\hat{P}(z_t|r_{1:t-1})$ is always 1/2. Finally, we resample the particles from this updated belief distribution.

Initial	Belief $\hat{P}(z_1)$	Weights $P(r_1\|z_1)$	Unnormalized $\hat{P}(z_1, r_1)$	Normalized $\hat{P}(z_1\|r_1)$	Resampling
$p_1 =$ (0 , 0) $p_2 =$ (1 , 1) a_1, a_2	1/2 1/2	1/2 1	1/4 1/2	$4/3 * 1/4 = 1/3$ $4/3 * 1/2 = 2/3$	$p_1 =$ (1 , 1) $p_2 =$ (1 , 1) a_3, a_4

Transition $P(z_2 z_1)$	Belief $\hat{P}(z_2\|r_1)$	Weights $P(r_2\|z_2)$	Unnormalized $\hat{P}(z_2, r_2\|r_1)$	Normalized $\hat{P}(z_2\|r_1, r_2)$	Resampling
$p_1 =$ (0 , 1) $p_2 =$ (1 , 1) a_5(left), a_6(stay)	1/2 1/2	1 1/2	1/2 1/4	$4/3 * 1/2 = 2/3$ $4/3 * 1/4 = 1/3$	$p_1 =$ (0 , 1) $p_2 =$ (1 , 1) a_7, a_8

Transition $P(z_3\|z_2)$	Belief $\hat{P}(z_3\|r_1, r_2)$	Weights $P(r_3\|z_3)$	Unnormalized $\hat{P}(z_3, r_3\|r_1, r_2)$	Normalized $\hat{P}(z_3\|r_1, r_2, r_3)$	Resampling
$p_1 =$ (-1 , 1) $p_2 =$ (1 , 1) a_9(left), a_{10}(stay)	1/2 1/2	1/4 1	1/8 1/2	$8/5 * 1/8 = 1/5$ $8/5 * 1/2 = 4/5$	$p_1 =$ (-1, 1) $p_2 =$ (1 , 1) a_{11}, a_{12}

(d) Use your samples (the unweighted particles in the last step) to evaluate the posterior probability that the x-coordinate of Z_3 is different than the column of Z_3, i.e. $X_3 \neq Y_3$.

Out of the two unweighted particles in the last step, exactly one satisfies $X_3 = Y_3$, so the estimate is 1/2.

(e) What is the problem of using the elimination algorithm instead of a particle filter for tracking Jabberwock?

The state space is infinite, so factors of infinite size (distributions over all points on the plane) would need to be computed and stored when using the variable elimination algorithm.

3 Game Theory: Equilibrium

(a) What is a Nash Equilibrium?

A Nash Equilibrium is a set of strategies where no player benefits from changing their strategy alone.

(b) Does a Nash Equilibrium always exist?

If there are a finite number of players and a finite number of actions, then there always exists a Nash Equilibrium (Nash, 1950). This strategy can be either pure or mixed (consider Rock-Paper-Scissors as an example of a game that doesn't have a pure strategy but does have a mixed one).

(c) What is another example of a solution concept that might be useful? What is the difference between this concept and Nash Equilibrium?

In a Stackelberg equilibrium (from lecture), instead of assuming all players play at once, players play one at a time and have some knowledge of what previous players did. This equilibrium involves different assumptions about game dynamics. If you're interested in other solution concepts, maybe try looking up subgame perfect Nash equilibrium or correlated equilibrium!

(d) What are some drawbacks to following the strategy of Nash equilibrium as a solution concept?

It makes assumptions about the other players (e.g., other players also play according to that same Nash equilibrium, no players and information about other players' strategies). Also imagine a game like Prisoner's Dilemma where following the Nash Equilibrium results in an action outcome strictly worse off for all players than another outcome possible in the game.

(e) What is a strategy? What is the difference between a pure and a mixed strategy?

A strategy is a probability distribution over actions. A pure strategy is deterministic (e.g. chosen with probability 1), while a mixed strategy incorporates a randomized action selection based on a distribution.

(f) What is the difference between a weakly dominant and strictly dominant strategy?

A strategy is weakly dominant if the player's payoff is at least as good as the opponent's (no matter the opponent's strategy). A strategy is strictly dominant if the player's payoff is always strictly better than the opponent's.

(g) Consider rock paper scissors where Player 1's strategy is to always play rock, and Player 2's strategy is to play scissors or paper with equal probability. Is this a Nash Equilibrium? What strategy would be best for Player 1 given Player 2's current strategy? What strategy would be best for Player 2 given Player 1's current strategy?

This is not a Nash Equilibrium. Player 1 would benefit from changing their strategy to always playing scissors. Player 2 would benefit from changing their strategy to always playing paper.

(h) Recursively remove dominated strategies to find the Nash Equilibrium of the following game. The order of utilities in each cell is the roman numeral player then the alphabet player.

	A	B	C
i	3,0	0,-5	0,-4
ii	1,-1	3,3	-2,4
iii	2,4	4,1	-1,8

Firstly, strategy ii is dominated by strategy iii. Also, strategy B is dominated by strategy C. This leaves the four corners of the table to be considered. Considering only the corners, strategy i dominates strategy iii. This leaves only the top two corners, in which case strategy A dominates strategy C and we are left with (3,0) as the Nash Equilibrium.

4 Game Theory: Pacman Hunt

Let us define the Pacman Hunt game for Simrit and Olivia (derivative of the well-known Stag Hunt game!). Suppose Simrit and Olivia are both hunters, they can choose to hunt Pacman or ghosts. If they hunt a ghost, they will always be successful and gain the modest payoff of 1. If they hunt Pacman, they are successful only if they both chose to hunt Pacman because then they can cooperate. In this case, they gain payoff of 2 each. However, if one hunts Pacman and one hunts ghosts, the one that hunted Pacman gets a payoff of 0. We can formulate the game in the below utility table:

Simrit, Olivia	Pacman	Ghost
Pacman	2,2	0,1
Ghost	1,0	1,1

(a) What are the pure Nash Equilibria of this problem?

The two pure Nash Equilibria are (Pacman, Pacman) and (Ghost, Ghost). Clearly, if both choose Pacman, they have gotten the best possible outcome, so switching does not help. If both do Ghost, unilateral switching would change the payoff to 0. The other two pure outcomes are not Nash Equilibria because the player that chose Pacman would want to switch to Ghost knowing the other player fixed Ghost, and the player that chose Ghost would want to switch to Pacman, knowing the other player fixed Pacman. Note that we only need one player to be dissatisfied for an outcome to not be a Nash; it just happened to be that both are dissatisfied in these non-Nash outcomes.

(b) We will now investigate the possibility of a mixed Nash equilibrium. Recall that in a mixed Nash Equilibrium, the utilities of the weighted actions are equal. Let p be the probability that Olivia picks Pacman.

(i) What is the expected value of action Pacman for Simrit?

$2p + 0(1 - p) = 2p$. In this case, Simrit's payoff is contingent on Olivia also hunting Pacman.

(ii) What is the expected value of action Ghost for Simrit?

$1p - 1(1 - p) = 1$. In this case, Olivia's action does not affect Simrit's hunt for Ghost.

(iii) What value of p makes these two expected values the same?

$2p = 1$
$p = 1/2$

(iv) Since the table is symmetric, the probability that equalizes the value of action Pacman and Ghost for both players is the same. What is the expected utility for both Olivia and Simrit if they play according to the mixed Nash Equilibria? How does this utility compare to the equilibria from (a)?

All action outcomes have probability 1/4 in this game. Therefore, we can take the utility to be the average of the payoff of (Pacman, Pacman), (Pacman, Ghost), (Ghost, Pacman), (Ghost, Ghost), which is $\frac{2+0+1+1}{4} = 1$. We may also use the utility equations above. This equilibrium is as good as always playing Ghost, and is worse than always playing Pacman.

(c) What if the game changed and somehow the Pacmen in the wild got larger or more profitable so the utility table then became

Simrit, Olivia	Pacman	Ghost
Pacman	3,3	0,1
Ghost	1,0	1,1

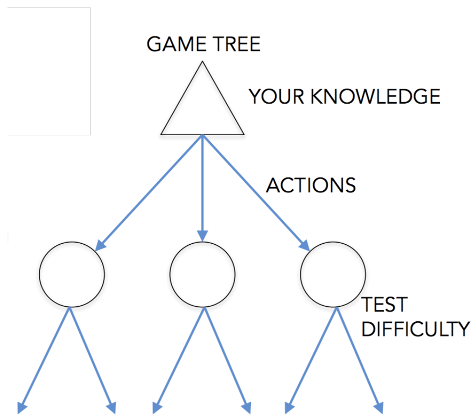
Calculate the mixed Nash Equilibrium for this game. Are the results surprising?

The Nash equilibrium has each player playing Pacman with probability 1/3 and Ghost with probability 2/3. This might feel weird because it seems as though with higher payoffs for Pacman, a solution concept would favor Pacman more. Remember that solution concepts are about stability, and not necessarily making some "correct" strategy!

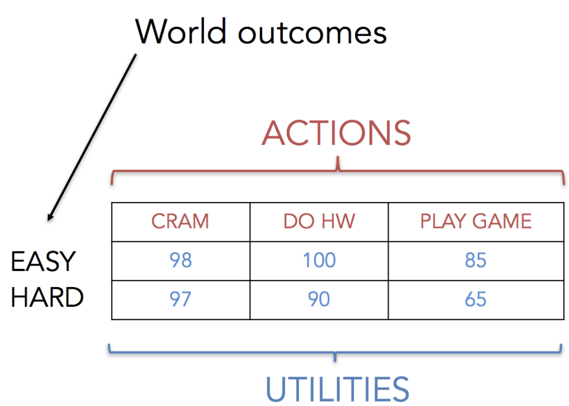
5 Game Theory Search

(a) What are the differences between extensive form and normal form games?

An extensive form game is one where every player is represented, as well as every opportunity that each of those players has to make a move, what moves they can make at that time, what information they have at the time, and the possible payoffs to be received.



On the other hand, a normal form game is one where games are approximated as a single shot. It represents only actions and utilities. This is the table format that we're used to using at this point.



(b) We can practice computing utilities using this example that was given in lecture:

CRAM	DO HW	PLAY GAME
98	100	85
97	90	65

P(EASY) = .2
P(HARD) = .8

- What is the utility of the pure strategy: cram?

  98 * 0.2 + 97 * 0.8 = 97.2

- What is the utility of the pure strategy: do HW?

$100 * 0.2 + 90 * 0.8 = 92$

- What is the utility of the mixed strategy: $\frac{1}{2}$ cram, $\frac{1}{2}$ do HW?

  $\frac{1}{2} * 97.2 + \frac{1}{2} * 92 = 94.6$

(c) Consider a two player game, where each player must simultaneously choose a number from $\{2, 3, ..., 99, 100\}$. Let x_1 represent the value chosen by player 1, and x_2 represent the value chosen by player 2. The rules of the game are such that the utility for a player 1 can be given as:

$$u(p_1) = \begin{cases} x_1 & x_1 = x_2 \\ x_2 - 2 & x_1 > x_2 \\ x_2 + 2 & x_1 < x_2 \end{cases}$$

Because the rules of the game for everyone are the same, the utility function for player 2 is symmetric to $u(p1)$. Does there exist a pure Nash Equilibrium for this game? It may help to try to play a few rounds of this game with someone next to you.

Yes. If $x_1 = x_2 = 2$, then neither player has any incentive to unilaterally change their strategy. There is no option to choose a lower number, and any player who chooses to change their number to anything greater than 2 will then receive a utility of 0, which is strictly worse than their current utility of 2.