MDL Assignment 2

Value Iteration

Part 2

- Dayitva Goel: 2019101005
- Prajneya Kumar: 2019114011

1 Overview of the Value Iteration Algorithm

Value Iteration is an algorithm that tries to calculate the optimal value of any given state, given a sample world consisting of different states, by iteratively improvising utility values based on the Bellman Equation until convergence.

The Bellman Update Equation is given by:

$$U_{t+1}(I) = \max_A \left[R(I, A) + \sum_J P(J \mid I, A) \cdot U_t(J) \right]$$

$$R(I, A) = \sum_J P(J \mid I, A) \cdot R(J, A, I)$$

We implemented this in the following way:

```
final_util = 0

for state in states:
    add_factor = state[0] * utilities[state[1].show()]
    final_util += add_factor

new_util = max(new_util, cost + gamma_inp * final_util)
```

where *states* consists of each state that IJ can go to from current state.

The cost is calculated as:

```
cost = 0
for choice in choices:
    cost += choice[0] * (costs[ACTION] + REWARD[choice[1].show()])
```

2 Task 1

Hyperparameter	Value
DELTA	0.001
GAMMA	0.999
ALL STEP COSTS	-20
PRIZE	+50
PUNISHMENT	-40
ITERATIONS	142

Analysis

The agent IJ tries to maximise the utility for each action-state pair. With a high gamma of 0.999, he prefers long-term rewards and therefore wants to kill MM as soon as possible to gain the final rewards of +50.

When IJ is in North Square:

If MM is in Dormant State, IJ moves down to Center Square to take an attacking position. If IJ has material, he chooses to craft arrows, else prefers to stay in North Square.

When IJ is in East Square:

IJ choses to HIT or SHOOT depending upon the number of arrows he has and state of MM.

When IJ is in South Square:

If MM's health is above 25, IJ decides to stay or gather, depending upon the amount of materials he has.

If MM's health is below 25, IJ decides to move up to Center Square to take an attacking position.

When IJ is in West Square:

If IJ has arrows, he shoots. If he doesn't and MM is in Dormant State, then he moves to attacking position by moving right to Center Square. Otherwise, he stays.

When IJ is in Center Square:

If IJ has arrows, he shoots. Else he chooses to craft more arrows by moving up if he has materials.

If MM is in ready state, then he moves left to West Square to protect himself from MM's attack. Else, he moves to the right to East Square to have a better chance of killing MM.

IJ never goes down to South Square to gather materials because the total step cost of gathering, crafting, and moving is too high for the final reward.

Simulation

Starting Point: (W, 0, 0, D, 100)

State	Policy
('W', 0, 0, 'D', 100)	RIGHT
('C', 0, 0, 'R', 100)	RIGHT
('C', 0, 0, 'D', 100)	RIGHT
('E', 0, 0, 'D', 100)	HIT
('E', 0, 0, 'D', 100)	HIT
('E', 0, 0, 'D', 100)	HIT
('E', 0, 0, 'R', 100)	HIT
('E', 0, 0, 'R', 100)	HIT
('E', 0, 0, 'D', 100)	HIT
('E', 0, 0, 'R', 100)	HIT
('E', 0, 0, 'R', 100)	HIT
('E', 0, 0, 'D', 100)	HIT
('E', 0, 0, 'D', 100)	HIT
('E', 0, 0, 'R', 100)	HIT
('E', 0, 0, 'R', 50)	HIT
('E', 0, 0, 'R', 50)	HIT
('E', 0, 0, 'D', 75)	HIT
('E', 0, 0, 'D', 75)	HIT
('E', 0, 0, 'R', 25)	HIT
('E', 0, 0, 'R', 25)	HIT
('E', 0, 0, 'D', 50)	HIT
('E', 0, 0, 'D', 50)	HIT
('E', 0, 0, 'D', 50)	HIT
('E', 0, 0, 'D', 50)	HIT

Simulation Analysis

Since IJ does not have any arrows, he moves to the East Square to have the best chance to HIT MM. He keeps hitting MM until he dies and he receives the final reward of +50.

Starting Point: (C, 2, 0, R, 100)

State	Policy
(('C', 2, 0, 'R', 100)	UP
('N', 2, 0, 'R', 100)	CRAFT
('N', 1, 1, 'R', 100)	CRAFT
('N', 0, 3, 'R', 100)	DOWN
('C', 0, 3, 'R', 100)	LEFT
('E', 0, 3, 'R', 100)	SHOOT
('E', 0, 2, 'R', 75)	SHOOT
('E', 0, 1, 'R', 50)	SHOOT
('E', 0, 0, 'D', 75)	HIT
('E', 0, 0, 'R', 25)	HIT
('E', 0, 0, 'D', 50)	HIT

Simulation Analysis

In this starting state, IJ has 2 materials, so he moves to the North Square to craft arrows. He then moves down to Center Square to attack MM and then to East Square to have to best chance of shooting MM. He shoots until he has no arrows left after which he keeps hitting MM until he dies and he receives the final reward of +50.

3 Task 2 Case 1

Hyperparameter	Value
DELTA	0.001
GAMMA	0.999
ALL STEP COSTS	-20
PRIZE	+50
PUNISHMENT	-40
ITERATIONS	145

Analysis

In this case, IJ on the LEFT action at East Square goes to the West Square. Most of IJ's actions are similar to Task 1 with the only difference being that when IJ is in the East Square,

IJ chooses to go Left and land in West Square when MM is in Ready State.

This happens because IJ tries to avoid punishment of -40 and now the total step cost of going from East Square to West Square has reduced from -40 to -20.

4 Task 2 Case 2

Hyperparameter	Value
DELTA	0.001
GAMMA	0.999
ALL STEP COSTS EXCEPT STAY	-20
STAY COST	0
PRIZE	+50
PUNISHMENT	-40
ITERATIONS	59

Analysis

Since the cost of staying has now become zero, IJ prefers to stay in places where MM cannot harm him.

When MM's health is above 25 and IJ is in Center or East Square(the places when MM can attack)

IJ chooses to run away to a safe square (North, West or South) and chooses to stay there for infinite amount of time.

When IJ is already in a safe square, he chooses to stay there. This means that the end state of killing MM cannot be reached in these conditions.

When MM's health is below 25

IJ chooses to shoot or hit depending on the amount of arrows he has.

5 Task 2 Case 3

Hyperparameter	Value
DELTA	0.001
GAMMA	0.25
ALL STEP COSTS	-20
PRIZE	+50
PUNISHMENT	-40
ITERATIONS	9

Analysis

Since the value of gamma is now decreased to 0.25, convergence is much faster. Furthermore, IJ now does not look for long-term rewards and is biased by short term gains and punishments.

In Task1, IJ used to keep hitting MM regardless of his health since he saw the long term reward of +50 of killing MM. But in this case, when gamma is near 0, IJ cannot see the long term reward of +50 and prefers to save himself when MM is in Ready state with health above 50.

When MM's health is above 50, he chooses to either stay or gather in South Square and stay or craft in North Square instead of moving to the Center Square and taking an attacking position.