Assignment 4 – MDP

Explanation of the method employed:

The representation of state –

We decided to represent a state with two fields,

1. The **current location** of the agent – a vertex object
2. A dictionary with **blockages status**

As the broken probability is given on each **vertex** – but also it blocks the passage through all edges leading to it we decided that the dictionary will hold the probability of an edge be blocked (leading to a broken vertex)

We based our implementation based on the policy iteration algorithm presented in chapter 17 page 13 in the lectures,

Where we hope to minimize the policy value for each state,

We first initialize the policy value to arbitrary values – in our case for each state where the current location is the goal the value is 0, and for all the remaining state the value is the predefined minimum – -10,000

Then we iterate in 4 loops

1. while there are significant changes in the policy values

2. for each of the state excluding the goal state

3. for all legal moves to be taken from a state

– represented by legal edge to traverse based on the blockage status dictionary

4. for each neighboring state – if the edge between the states is a legal move we continue to compute the expected utility of transition

In each iteration we look for the best legal action given the current state, and update the expected value of the path to the destination from the current location accordingly.

After the expected value from each location is calculated the simulation starts, the user is asked to decide weather each one of the paths exists or was affected by the brittle vertex, then the agent start to go down the best path (with the best utility value) – each time he reaches a vertex if his neighbors were with an unknown state they get revealed then with the updated status he choses the best next move and repeats again.