**A logo for a university

Description automatically generated**

**MACHINE LEARNING COURSEWORK II**

BY

**STUDENT: MAIMUNATU TUNAU**

**STUDENT ID: 240008998**

FOR

**COURSE: MACHINE LEARNING**

**COURSE ID:**

TASK 2

For this task, the optimal policy at each green state should be solved taking into consideration the probabilistic movement, the inaccessible states and the boundaries

To calculate this, we look to the general Bellman Equation which is:

Then to calculate the expected utility for each action:

Implemented based on the criteria given to us would be:

The value I have given is 0.9

|  |  |  |  |
| --- | --- | --- | --- |
| Actions | (2,1) | (4,3) | (5,2) |
| Up | 0.8(7.41\*0.9)+0.1(7.31\*0.9)+0.1(7.31\*0.9)+(-0.1)= 6.55 | 0.8(6.44\*0.9)+0.1(6.77\*0.9)+0.1(5.87\*0.9)+ (-0.1)= 5.67 | 0.8(6.77\*0.9)+0.1(6.90\*0.9)+0.1(6.59\*0.9) +(-0.1)=5.99 |
| Down | 0.8(7.15\*0.9)+0.1(7.31\*0.9)+0.1(7.31\*0.9) +(-0.1)= 6.36 | 0.8(6.59\*0.9)+0.1(5.87\*0.9)+0.1(6.77\*0.9) +(-0.1)= 5.78 | 0.8(6.80\*0.9)+0.1(6.59\*0.9)+0.1(6.90\*0.9) +(-0.1)=6.01 |
| Left | 0.8(7.31\*0.9)+0.1(7.15\*0.9)+0.1(7.41\*0.9)+(-0.1)= 6.47 | 0.8(6.77\*0.9)+0.1(6.59\*0.9)+0.1(6.44\*0.9) +(-0.1)= 5.95 | 0.8(6.90\*0.9)+0.1(6.80\*0.9)+0.1(6.77\*0.9) +(-0.1)=6.09 |
| Right | 0.8(7.31\*0.9)+0.1(7.41\*0.9)+0.1(7.15\*0.9)+(-0.1)= 6.47 | 0.8(5.87\*0.9)+0.1(6.44\*0.9)+0.1(6.59\*0.9) +(-0.1)= 5.29 | 0.8(6.59\*0.9)+0.1(6.77\*0.9)+0.1(6.80\*0.9) +(-0.1)=5.87 |

Based on these calculations, the optimal policy can easily be identified ad the one with the highest expected utility

* (2,1) : Highest utility is 6.55, action is **UP**
* (4,3) :Highest utility is 5.95, action is **LEFT**
* (5,2) : Highest utility is 6.09, action is **LEFT**