

MSc in Computer Science and Engineering

**Learning and Decision Making 2016-2017**

Homework 2. Markov Decision Problems

**(a)**

**Note:** we considered the fist square the square 0 instead of 1. This is more coherent with indexes in python.

Table 1 - Model of the board.

|  |  |
| --- | --- |
| 0 | 1 |
| 2 | 3 |

X = {(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (1,3), (2,0), (2,1), (2,2), (2,3), (3,0), (3,1), (3,2), (3,3)},

where the fist number in the tuples (X[i][0]) is the position of the wolf and the second number (X[i][1]) is the position of the hare.

Ex: X [0] = (0,0) means that both wolf and hare are at square 0 and X [6] = (1,2) means that the wolf is at the square 1 and hare ate square 2.

A = {Left, Right, Up, Down, Stay}

**(b)**

For this MDP since the board is Toroidal world the probabilities associated with Up and Down or Left and Right are the same.

Table 2 - Transition probabilities for actions Up and Down

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 00 | 01 | 02 | 03 | 10 | 11 | 12 | 13 | 20 | 21 | 22 | 23 | 30 | 31 | 32 | 33 |
| 00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.48 | 0.16 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 01 | 0.04 | 0.12 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.48 | 0.00 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 |
| 02 | 0.04 | 0.00 | 0.12 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.00 | 0.48 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 |
| 03 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.16 | 0.48 | 0.00 | 0.00 | 0.00 | 0.00 |
| 10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.48 | 0.16 | 0.16 | 0.00 |
| 11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.12 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.48 | 0.00 | 0.16 |
| 12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.12 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.00 | 0.48 | 0.16 |
| 13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.16 | 0.48 |
| 20 | 0.48 | 0.16 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 21 | 0.16 | 0.48 | 0.00 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.12 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 |
| 22 | 0.16 | 0.00 | 0.48 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.12 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 |
| 23 | 0.00 | 0.16 | 0.16 | 0.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 |
| 30 | 0.00 | 0.00 | 0.00 | 0.00 | 0.48 | 0.16 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 |
| 31 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.48 | 0.00 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.12 | 0.00 | 0.04 |
| 32 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.00 | 0.48 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.12 | 0.04 |
| 33 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.16 | 0.16 | 0.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.12 |

Table 3 - Transition probabilities for actions Left and Right

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 00 | 01 | 02 | 03 | 10 | 11 | 12 | 13 | 20 | 21 | 22 | 23 | 30 | 31 | 32 | 33 |
| 00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.48 | 0.16 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 01 | 0.04 | 0.12 | 0.00 | 0.04 | 0.16 | 0.48 | 0.00 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 02 | 0.04 | 0.00 | 0.12 | 0.04 | 0.16 | 0.00 | 0.48 | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 03 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.16 | 0.16 | 0.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 10 | 0.48 | 0.16 | 0.16 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 11 | 0.16 | 0.48 | 0.00 | 0.16 | 0.04 | 0.12 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 12 | 0.16 | 0.00 | 0.48 | 0.16 | 0.04 | 0.00 | 0.12 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 13 | 0.00 | 0.16 | 0.16 | 0.48 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 20 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.48 | 0.16 | 0.16 | 0.00 |
| 21 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.12 | 0.00 | 0.04 | 0.16 | 0.48 | 0.00 | 0.16 |
| 22 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.12 | 0.04 | 0.16 | 0.00 | 0.48 | 0.16 |
| 23 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.16 | 0.16 | 0.48 |
| 30 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 | 0.48 | 0.16 | 0.16 | 0.00 | 0.12 | 0.04 | 0.04 | 0.00 |
| 31 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.12 | 0.00 | 0.04 | 0.16 | 0.48 | 0.00 | 0.16 | 0.04 | 0.12 | 0.00 | 0.04 |
| 32 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.12 | 0.04 | 0.16 | 0.00 | 0.48 | 0.16 | 0.04 | 0.00 | 0.12 | 0.04 |
| 33 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.12 | 0.00 | 0.16 | 0.16 | 0.48 | 0.00 | 0.04 | 0.04 | 0.12 |

Table 4 - Transition probabilities for action Stay

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 00 | 01 | 02 | 03 | 10 | 11 | 12 | 13 | 20 | 21 | 22 | 23 | 30 | 31 | 32 | 33 |
| 00 | 0.6 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 01 | 0.2 | 0.6 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 02 | 0.2 | 0.0 | 0.6 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 03 | 0.0 | 0.2 | 0.2 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 10 | 0.0 | 0.0 | 0.0 | 0.0 | 0.6 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 11 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.6 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 12 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.6 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 13 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 20 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.6 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 21 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.6 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| 22 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.6 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| 23 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 |
| 30 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.6 | 0.2 | 0.2 | 0.0 |
| 31 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.6 | 0.0 | 0.2 |
| 32 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.6 | 0.2 |
| 33 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.6 |

Our cost function between a state and an action are 0 if after performing action a successfully, the wolf and the hare stay on the same position, and 1.0 otherwise.

Table 5 – Normalized Matrix with the cost of the actions for every state.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| States \ actions | L | R | U | D | S |
| 00 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 |
| 01 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| 02 | 1.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| 03 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 10 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| 11 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 |
| 12 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 13 | 1.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| 20 | 1.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| 21 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 22 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 |
| 23 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| 30 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 31 | 1.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| 32 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| 33 | 1.0 | 1.0 | 1.0 | 1.0 | 0.0 |

**(c)**

Cost-to-Go function:

with :

|  |  |
| --- | --- |
|  |  |
| 00 | 0.06232667 |
| 01 | 0.06298052 |
| 02 | 0.06164706 |
| 03 | 0.06304575 |
| 10 | 0.06298052 |
| 11 | 0.06232667 |
| 12 | 0.06304575 |
| 13 | 0.06164706 |
| 20 | 0.06164706 |
| 21 | 0.06304575 |
| 22 | 0.06232667 |
| 23 | 0.06298052 |
| 30 | 0.06304575 |
| 31 | 0.06164706 |
| 32 | 0.06298052 |
| 33 | 0.06232667 |