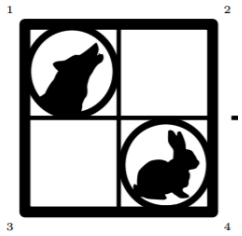
João Rodrigues 78672
Luís Morais 78416

Homework 2

a)



State Space:

['1W1H', '1W2H', '1W3H', '1W4H', '2W1H', '2W2H', '2W3H', '2W4H', '3W1H', '3W2H', '3W3H', '3W4H', '4W1H', '4W2H', '4W3H', '4W4H']

Em que 1W2H significa que o **Wolf** está presente na posição 1 e o **Hare** na posição 2.

Action Space:
['U', 'D', 'R', 'L', 'S']

U = Up
D = Down
R = Right
L = Left
S = Stay

Em que as ações dizem respeito aos movimentos do **Wolf**, pois é o único que considerámos ser um agente que executa ações, sendo que o **Hare** apenas age de forma aleatória.

b) Nesta alínea criámos uma matriz de probabilidades para cada ação possível do **Wolf**

Matriz de Probabilidades para Ação U (UP):

```
Probability Matrix for action UP:
[[ 0.6  0.2  0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.2  0.   0.6  0.2  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. ]
 [ 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.   0.2  0.6  0.   0.2  0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2
  0.2  0.6]]
```

Matriz de Probabilidades para Ação D (DOWN):

```
Probability Matrix for action DOWN:
[[ 0.6  0.2  0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.2  0.   0.6  0.2  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. ]
 [ 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.   0.2  0.6  0.   0.2  0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2
  0.2  0.6]]
```

Matriz de Probabilidades para Ação L (LEFT):

```
Probability Matrix for action LEFT:
[[ 0.6  0.2  0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.2  0.   0.6  0.2  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. ]
 [ 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.   0.2  0.6  0.   0.2  0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2
  0.2  0.6]]
```

João Rodrigues 78672
Luís Morais 78416

Matriz de Probabilidades para Ação R (RIGHT):

```
Probability Matrix for action Right:
[[ 0.6  0.2  0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.2  0.   0.6  0.2  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. ]
 [ 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.   0.2  0.6  0.   0.2  0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2
  0.2  0.6]]
```

Matriz de Probabilidades para Ação S (STAY):

```
Probability Matrix for action STAY:
[[ 0.6  0.2  0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.   0.2  0.   0.   0.   0.   0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.2  0.   0.6  0.2  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. ]
 [ 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.   0.2  0.6  0.   0.2  0.   0.   0.
  0. ]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6  0.2  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.   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.   0.2  0.6  0.
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2  0.   0.6
  0.2]
 [ 0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.   0.2
  0.2  0.6]]
```

João Rodrigues 78672
Luís Morais 78416

Matriz de Custos:

```
Cost Function for the MDP:
 [[ 1.   1.   1.   1.   0. ]
 [ 1.   1.   0.   0.   0.5]
 [ 0.   0.   1.   1.   0.5]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 1.   1.   0.   0.   0.5]
 [ 1.   1.   1.   1.   0. ]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 0.   0.   1.   1.   0.5]
 [ 0.   0.   1.   1.   0.5]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 1.   1.   1.   1.   0. ]
 [ 1.   1.   0.   0.   0.5]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 0.   0.   1.   1.   0.5]
 [ 1.   1.   0.   0.   0.5]
 [ 1.   1.   1.   1.   0. ]]
```

Explicação:

Nos estados em que o **Wolf** e o **Hare** estão na mesma posição é dado custo 0 para a ação **Stay** e custo 1 às restantes visto que o **Hare** tem maior probabilidade de falhar ao tentar mover-se, ficando na mesma posição.

É sempre dado custo zero a qualquer ação que possibilite ao **Wolf** ficar na mesma posição que o **Hare**, visto que é esse o objetivo do **Wolf.**

Nos estados em que o **Wolf** está ao lado do **Hare,** é dado custo 0 para as ações **Left e Right**, custo 0.5 à ação **Stay** e custo 1 às restantes.

Nos estados em que o **Wolf** está em cima ou em baixo do **Hare,** é dado custo 0 para as ações **Down e Up**, custo 0.5 a **Stay** e custo 1 às restantes.

Nos estados em que o **Wolf** está numa diagonal em relação ao **Hare**, é dado custo 1 à Ação **Stay** e 0.5 às restantes.

c)

Usando a fórmula: $$\boldsymbol{J}^{\pi} = (\mathbf{I} - \gamma \mathbf{P}_{\pi})^{-1} \boldsymbol{c}_{\pi}$$

Obtemos a seguinte Cost-to-go function dada a política " in which the wolf always goes up":

```
The costo-to-go funtion associated with the given policy:
 [ 63.2716239   63.57566674  61.11261256  62.0400968   63.57566674
  63.2716239   62.0400968   61.11261256  61.11261256  62.0400968
  63.2716239   63.57566674  62.0400968   61.11261256  63.57566674
  63.2716239 ]
```