INSTITUTO SUPERIOR TÉCNICO

ADI – Aprendizagem e Decisão Inteligente

MEIC 2016-2017 – 2º Semestre
Grupo 27 – Alameda

Pedro Duarte – 78328

Pedro Guerreiro - 78264

Homework 2 – Markov Decision Problems

Exercise 1

a)

X = {

(1,1), (1,2), (1,3), (1,4),
(2,1), (2,2), (2,3), (2,4),
(3,1), (3,2), (3,3), (3,4),
(4,1), (4,2), (4,3), (4,4)}

A = { up, down, left, right, stay }

NOTA: (1,2) significa “lobo na posição 1, lebre na posição 2”, etc.

b)

$P_{up} = P_{down} =$

```
[[ 0.12  0.04  0.04  0.    0.    0.    0.    0.    0.48  0.16  0.16  0.    0.    0.    0.    0.  ]
 [ 0.04  0.12  0.    0.04  0.    0.    0.    0.    0.16  0.48  0.    0.16  0.    0.    0.    0.  ]
 [ 0.04  0.    0.12  0.04  0.    0.    0.    0.    0.16  0.    0.48  0.16  0.    0.    0.    0.  ]
 [ 0.    0.04  0.04  0.12  0.    0.    0.    0.    0.    0.16  0.16  0.48  0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.    0.12  0.04  0.04  0.    0.    0.    0.    0.    0.48  0.16  0.16  0.  ]
 [ 0.    0.    0.    0.    0.04  0.12  0.    0.04  0.    0.    0.    0.    0.16  0.48  0.    0.16]
 [ 0.    0.    0.    0.    0.04  0.    0.12  0.04  0.    0.    0.    0.    0.16  0.    0.48  0.16]
 [ 0.    0.    0.    0.    0.    0.04  0.04  0.12  0.    0.    0.    0.    0.    0.16  0.16  0.48]
 [ 0.48  0.16  0.16  0.    0.    0.    0.    0.    0.12  0.04  0.04  0.    0.    0.    0.    0.  ]
 [ 0.16  0.48  0.    0.16  0.    0.    0.    0.    0.04  0.12  0.    0.04  0.    0.    0.    0.  ]
 [ 0.16  0.    0.48  0.16  0.    0.    0.    0.    0.04  0.    0.12  0.04  0.    0.    0.    0.  ]
 [ 0.    0.16  0.16  0.48  0.    0.    0.    0.    0.    0.04  0.04  0.12  0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.    0.48  0.16  0.16  0.    0.    0.    0.    0.    0.12  0.04  0.04  0.  ]
 [ 0.    0.    0.    0.    0.16  0.48  0.    0.16  0.    0.    0.    0.    0.04  0.12  0.    0.04]
 [ 0.    0.    0.    0.    0.16  0.    0.48  0.16  0.    0.    0.    0.    0.04  0.    0.12  0.04]
 [ 0.    0.    0.    0.    0.    0.16  0.16  0.48  0.    0.    0.    0.    0.    0.04  0.04  0.12]]
```

$P_{left} = P_{right} =$

```
[[ 0.12  0.04  0.04  0.    0.48  0.16  0.16  0.    0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.04  0.12  0.    0.04  0.16  0.48  0.    0.16  0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.04  0.    0.12  0.04  0.16  0.    0.48  0.16  0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.    0.04  0.04  0.12  0.    0.16  0.16  0.48  0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.48  0.16  0.16  0.    0.12  0.04  0.04  0.    0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.16  0.48  0.    0.16  0.04  0.12  0.    0.04  0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.16  0.    0.48  0.16  0.04  0.    0.12  0.04  0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.    0.16  0.16  0.48  0.    0.04  0.04  0.12  0.    0.    0.    0.    0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.12  0.04  0.04  0.    0.48  0.16  0.16  0.  ]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.04  0.12  0.    0.04  0.16  0.48  0.    0.16]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.04  0.    0.12  0.04  0.16  0.    0.48  0.16]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.    0.04  0.04  0.12  0.    0.16  0.16  0.48]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.48  0.16  0.16  0.    0.12  0.04  0.04  0.  ]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.16  0.48  0.    0.16  0.04  0.12  0.    0.04]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.16  0.    0.48  0.16  0.04  0.    0.12  0.04]
 [ 0.    0.    0.    0.    0.    0.    0.    0.    0.    0.16  0.16  0.48  0.    0.04  0.04  0.12]]
```

$P_{stay} =$

```
[[ 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]]
```

$C =$

```
[[ 1.   1.   1.   1.   0. ]
 [ 1.   1.   0.   0.   1. ]
 [ 0.   0.   1.   1.   1. ]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 1.   1.   0.   0.   0. ]
 [ 1.   1.   1.   1.   0. ]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 0.   0.   1.   1.   1. ]
 [ 0.   0.   1.   1.   0. ]
 [ 0.5  0.5  0.5  0.5  1. ]
 [ 1.   1.   1.   1.   0. ]
 [ 1.   1.   0.   0.   1. ]
 [ 0.5  0.5  0.5  0.5  0. ]
 [ 0.   0.   1.   1.   1. ]
 [ 1.   1.   0.   0.   1. ]
 [ 1.   1.   1.   1.   0. ]]
```

c)

```
discount * P:
[[ 0.12  0.04  0.04  0.    0.    0.    0.    0.    0.48  0.16  0.16  0.    0.    0.    0.    0.  ]
 [ 0.04  0.12  0.    0.04  0.    0.    0.    0.    0.16  0.48  0.    0.16  0.    0.    0.    0.  ]
 [ 0.04  0.    0.12  0.04  0.    0.    0.    0.    0.16  0.    0.48  0.16  0.    0.    0.    0.  ]
 [ 0.    0.04  0.04  0.12  0.    0.    0.    0.    0.    0.16  0.16  0.48  0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.    0.12  0.04  0.04  0.    0.    0.    0.    0.    0.48  0.16  0.16  0.  ]
 [ 0.    0.    0.    0.    0.04  0.12  0.    0.04  0.    0.    0.    0.    0.16  0.48  0.    0.16]
 [ 0.    0.    0.    0.    0.04  0.    0.12  0.04  0.    0.    0.    0.    0.16  0.    0.48  0.16]
 [ 0.    0.    0.    0.    0.    0.04  0.04  0.12  0.    0.    0.    0.    0.    0.16  0.16  0.48]
 [ 0.48  0.16  0.16  0.    0.    0.    0.    0.    0.12  0.04  0.04  0.    0.    0.    0.    0.  ]
 [ 0.16  0.48  0.    0.16  0.    0.    0.    0.    0.04  0.12  0.    0.04  0.    0.    0.    0.  ]
 [ 0.16  0.    0.48  0.16  0.    0.    0.    0.    0.04  0.    0.12  0.04  0.    0.    0.    0.  ]
 [ 0.    0.16  0.16  0.48  0.    0.    0.    0.    0.    0.04  0.04  0.12  0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.    0.48  0.16  0.16  0.    0.    0.    0.    0.    0.12  0.04  0.04  0.  ]
 [ 0.    0.    0.    0.    0.16  0.48  0.    0.16  0.    0.    0.    0.    0.04  0.12  0.    0.04]
 [ 0.    0.    0.    0.    0.16  0.    0.48  0.16  0.    0.    0.    0.    0.04  0.    0.12  0.04]
 [ 0.    0.    0.    0.    0.    0.16  0.16  0.48  0.    0.    0.    0.    0.    0.04  0.04  0.12]]

I - discount * P:
[[ 0.88 -0.04 -0.04  0.    0.    0.    0.    0.   -0.48 -0.16 -0.16  0.    0.    0.    0.    0.  ]
 [-0.04  0.88  0.   -0.04  0.    0.    0.    0.   -0.16 -0.48  0.   -0.16  0.    0.    0.    0.  ]
 [-0.04  0.    0.88 -0.04  0.    0.    0.    0.   -0.16  0.   -0.48 -0.16  0.    0.    0.    0.  ]
 [ 0.   -0.04 -0.04  0.88  0.    0.    0.    0.    0.   -0.16 -0.16 -0.48  0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.    0.88 -0.04 -0.04  0.    0.    0.    0.    0.   -0.48 -0.16 -0.16  0.  ]
 [ 0.    0.    0.    0.   -0.04  0.88  0.   -0.04  0.    0.    0.    0.   -0.16 -0.48  0.   -0.16]
 [ 0.    0.    0.    0.   -0.04  0.    0.88 -0.04  0.    0.    0.    0.   -0.16  0.   -0.48 -0.16]
 [ 0.    0.    0.    0.    0.   -0.04 -0.04  0.88  0.    0.    0.    0.    0.   -0.16 -0.16 -0.48]
 [-0.48 -0.16 -0.16  0.    0.    0.    0.    0.    0.88 -0.04 -0.04  0.    0.    0.    0.    0.  ]
 [-0.16 -0.48  0.   -0.16  0.    0.    0.    0.   -0.04  0.88  0.   -0.04  0.    0.    0.    0.  ]
 [-0.16  0.   -0.48 -0.16  0.    0.    0.    0.   -0.04  0.    0.88 -0.04  0.    0.    0.    0.  ]
 [ 0.   -0.16 -0.16 -0.48  0.    0.    0.    0.    0.   -0.04 -0.04  0.88  0.    0.    0.    0.  ]
 [ 0.    0.    0.    0.   -0.48 -0.16 -0.16  0.    0.    0.    0.    0.    0.88 -0.04 -0.04  0.  ]
 [ 0.    0.    0.    0.   -0.16 -0.48  0.   -0.16  0.    0.    0.    0.   -0.04  0.88  0.   -0.04]
 [ 0.    0.    0.    0.   -0.16  0.   -0.48 -0.16  0.    0.    0.    0.   -0.04  0.    0.88 -0.04]
 [ 0.    0.    0.    0.    0.   -0.16 -0.16 -0.48  0.    0.    0.    0.    0.   -0.04 -0.04  0.88]]

(I - discount * P)^-1
[[ 13.65  12.31  12.31  12.05   0.     0.     0.     0.    12.9   12.38  12.38  12.03   0.     0.     0.     0.  ]
 [ 12.31  13.65  12.05  12.31   0.     0.     0.     0.    12.38  12.9   12.03  12.38   0.     0.     0.     0.  ]
 [ 12.31  12.05  13.65  12.31   0.     0.     0.     0.    12.38  12.03  12.9   12.38   0.     0.     0.     0.  ]
 [ 12.05  12.31  12.31  13.65   0.     0.     0.     0.    12.03  12.38  12.38  12.9    0.     0.     0.     0.  ]
 [  0.     0.     0.     0.    13.65  12.31  12.31  12.05   0.     0.     0.     0.    12.9   12.38  12.38  12.03]
 [  0.     0.     0.     0.    12.31  13.65  12.05  12.31   0.     0.     0.     0.    12.38  12.9   12.03  12.38]
 [  0.     0.     0.     0.    12.31  12.05  13.65  12.31   0.     0.     0.     0.    12.38  12.03  12.9   12.38]
 [  0.     0.     0.     0.    12.05  12.31  12.31  13.65   0.     0.     0.     0.    12.03  12.38  12.38  12.9 ]
 [ 12.9   12.38  12.38  12.03   0.     0.     0.     0.    13.65  12.31  12.31  12.05   0.     0.     0.     0.  ]
 [ 12.38  12.9   12.03  12.38   0.     0.     0.     0.    12.31  13.65  12.05  12.31   0.     0.     0.     0.  ]
 [ 12.38  12.03  12.9   12.38   0.     0.     0.     0.    12.31  12.05  13.65  12.31   0.     0.     0.     0.  ]
 [ 12.03  12.38  12.38  12.9    0.     0.     0.     0.    12.05  12.31  12.31  13.65   0.     0.     0.     0.  ]
 [  0.     0.     0.     0.    12.9   12.38  12.38  12.03   0.     0.     0.     0.    13.65  12.31  12.31  12.05]
 [  0.     0.     0.     0.    12.38  12.9   12.03  12.38   0.     0.     0.     0.    12.31  13.65  12.05  12.31]
 [  0.     0.     0.     0.    12.38  12.03  12.9   12.38   0.     0.     0.     0.    12.31  12.05  13.65  12.31]
 [  0.     0.     0.     0.    12.03  12.38  12.38  12.9    0.     0.     0.     0.    12.05  12.31  12.31  13.65]]
```

J =

```
[ 62.580312  62.972622  61.803924  62.643142  62.972622  62.580312  62.643142  61.803924  61.803924  62.643142  62.580312  62.97
2622  62.643142  61.803924  62.972622  62.580312]
```