TPC2

1.

1. The dragon and the princess do not move, the knight can move to six different positions, thus, there are 6 different states:

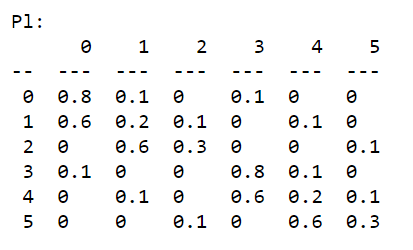
X = {1, 2, 3, 4, 5, 6}

The action space for the MDP is the knight move to LEFT, RIGHT, UP and DOWN:

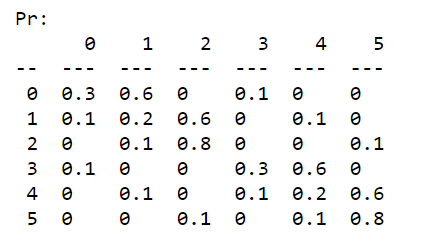
A = {LEFT, RIGHT, UP, DOWN}

1. The transition probabilities for each action are:

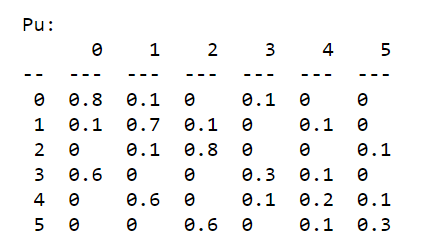
Probability for LEFT action:



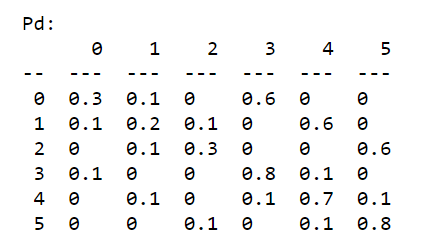
Probability for RIGHT action:



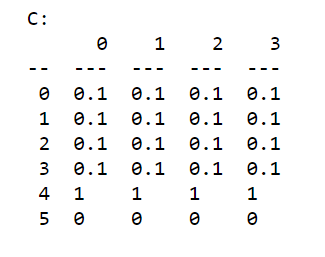
Probability for UP action:



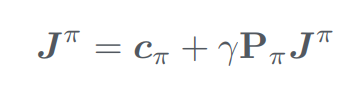
Probability for DOWN action:



For the cost function we assume that the cost for the states where there’s no dragon or princess is 0.1, when reaching the state where there is a dragon, the knight receives a penalty, of cost 1, and no cost at all at the state where princess is. Resulting in:

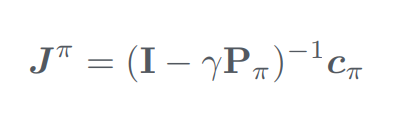


1. To compute the cost-to-go function associated with the policy which the knight always do the action UP, we solve the linear system:



Where .

Solving the equation:



Resulting:



Grupo nº 37

João Silvestre, nº 80996

João Pestana, nº 79515