

# Assignment 4 in Artificial Intelligence

## Spring 2018

Daniel Aaron Salwerowicz

April 30, 2018

### Task 1

**a**

Initial R and Q matrices are as follows:

$$R = \begin{bmatrix} 0 & 0 & -\text{inf} & 0 & -\text{inf} & -\text{inf} \\ 0 & 0 & -10 & -\text{inf} & 0 & -\text{inf} \\ -\text{inf} & 0 & -10 & -\text{inf} & -\text{inf} & 100 \\ 0 & -\text{inf} & -\text{inf} & 0 & 0 & -\text{inf} \\ -\text{inf} & 0 & -\text{inf} & 0 & 0 & 100 \\ -\text{inf} & -\text{inf} & -10 & -\text{inf} & 0 & 100 \end{bmatrix} \quad Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

They are initialized like that since Q-matrix is always set to zero at the start and R-matrix represents enviromental values. As I have interpreted it there's no penalty for staying in the same room or for moving between rooms that were not specified in text. Therefore only moves that end up in room 3 or 6 have values other than zero associated with them.  $-\text{inf}$  is used to show that there is no connection between rooms ensuring that algorithm will never pick them.

Formula for calculating values in Q-matrix given learning factor  $\alpha$  is 1:

$$Q(\text{State}, \text{Action}) = R(\text{State}, \text{Action}) + \gamma \cdot \text{Max}[Q(\text{Next S}, \text{All A})]$$

**b**

$$\begin{aligned} Q(4, 5) &= R(4, 5) + \gamma \cdot \text{Max}[Q(5, 1), Q(5, 4), Q(5, 5), Q(5, 6)] \\ &= 0 + 0.9 \cdot \text{Max}[0, 0, 0, 0] \\ &= 0 + 0.9 \cdot 0 \\ Q(4, 5) &= \underline{\underline{0}} \end{aligned}$$

**c**

$$\begin{aligned}Q(5, 6) &= R(5, 6) + \gamma \cdot \text{Max}[Q(6, 3), Q(6, 5), Q(6, 6)] \\&= 100 + 0.9 \cdot \text{Max}[0, 0, 0] \\&= 100 + 0.9 \cdot 0 \\Q(5, 6) &= \underline{\underline{100}}\end{aligned}$$

**d**

$$\begin{aligned}Q(4, 5) &= R(4, 5) + \gamma \cdot \text{Max}[Q(5, 1), Q(5, 4), Q(5, 5), Q(5, 6)] \\&= 0 + 0.9 \cdot \text{Max}[0, 0, 0, 100] \\&= 0 + 0.9 \cdot 100 \\Q(4, 5) &= \underline{\underline{90}}\end{aligned}$$

**e**

$$\begin{aligned}Q(1, 4) &= R(1, 4) + \gamma \cdot \text{Max}[Q(4, 1), Q(4, 4), Q(4, 5)] \\&= 0 + 0.9 \cdot \text{Max}[0, 0, 90] \\&= 0 + 0.9 \cdot 90 \\Q(1, 4) &= \underline{\underline{81}}\end{aligned}$$

**f**

Q matrix will look like this after going through all of the steps in b-e.

$$Q = \begin{bmatrix} 0 & 0 & 0 & 81 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 90 & 0 \\ 0 & 0 & 0 & 0 & 0 & 100 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Using this we can clearly see that the robot would take following path to the docking station:

$$1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 6$$

Since:

$$Q(1, 4) + Q(4, 5) + Q(5, 6) = 81 + 90 + 100 = 171$$

Which is the highest possible reward using values defined in Q matrix.

## Task 3

a

Going by the algorithm I will choose the splits so that  $S$  will be reduced as much as possible while keeping clean splits. Since I can have maximum of 4 regions I will only make three splits.

First split is going to be  $x > 3$ , since there are 4 red values in region defined by  $x \leq 3$  and no blue values. So  $(4, 0)$ .

Second split will be  $y > 2$ , since there are 3 red values in region defined by  $y \leq 2$  and no blue values. So  $(2, 0)$ .

Third and last split will be  $y > 5.5$ , since there are 5 blue values in region  $y \leq 5.5$  and no red values. So  $(0, 5)$ .

The remaining region will be majorly blue with 4 blue and 2 red values. So  $(2, 4)$ .

Those splits are pretty much justified by the algorithm itself, since I tried to follow it as well as I could and understood it.

Resulting tree and regions can be found in figures 1 and 2.

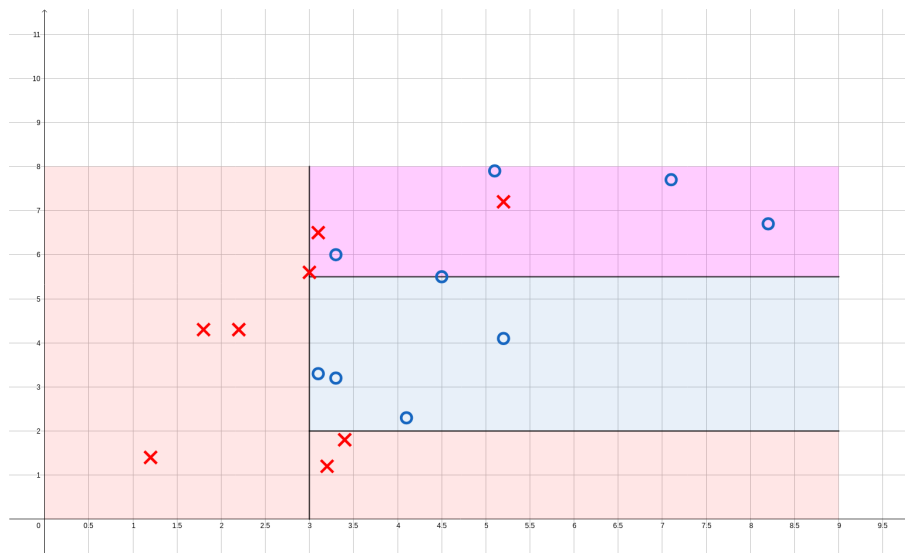


Figure 1: Regions split using the CART algorithm

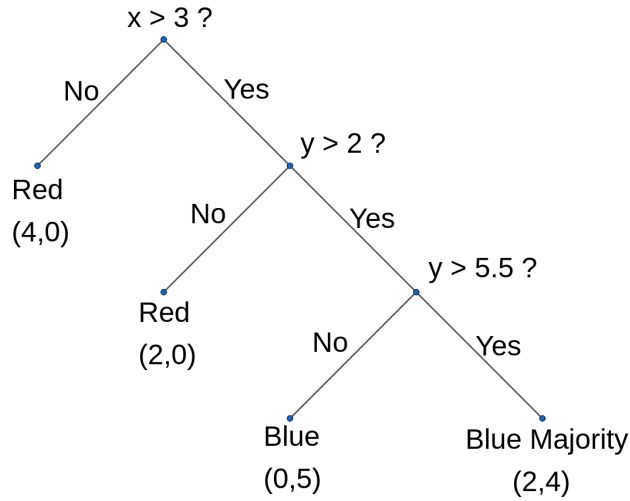


Figure 2: Regression tree for the region above

**b**

Plotting the test values  $(6.5, 5.5)$  and  $(4.1, 6.5)$  we can clearly see that the first point lands in the blue region, and is therefore defined as blue, whereas the second point falls in majorly blue region therefore there is  $66.\bar{6}\%$  bias that it will be defined as blue and  $33.\bar{3}\%$  bias that it will be defined as red. Therefore in my opinion and if my work is correct both of these points will be defined as blue. Figure 3 shows those new points plotted in the graph.

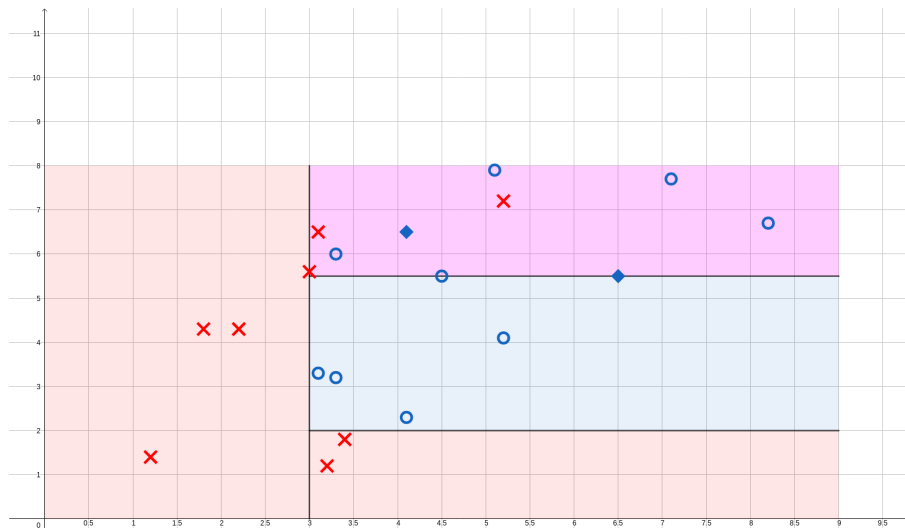


Figure 3: Test values plotted into the graph split by regions