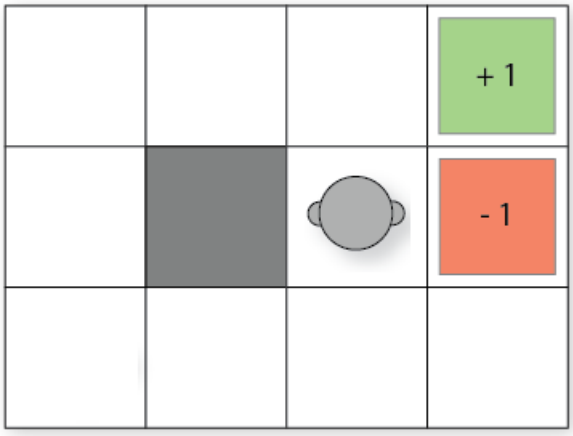
Problem Set VIII: Monte-Carlo Tree Search

Aim The purpose of this workshop is to help you get a better understanding of Monte-Carlo Tree Search for solving MDPs in an online manner.

Tasks

In this workshop, you will consider the example from the lectures of the agent that moves in a 2D grid world.



1. The agent is at cell (2,1), in which 2 is the x-coordinate and 1 the y-coordinate (both start from 0). It samples the following 10 iterations of MCTS, where E (East) goes right and W (West) goes left:

Iteration	**Trace**	
1	N	$simulate = 0$
2	E	$simulate = 0$
3	$N \rightarrow succ$	$simulate = 0$
4	$N \rightarrow slip(E)$	$simulate = -1$
5	$E \rightarrow succ$	$simulate = -1$
6	$N \rightarrow succ \rightarrow E$	$simulate = 1$
7	$E \rightarrow slip(N)$	$simulate = 0$
8	$E \rightarrow slip(N) \rightarrow E$	$simulate = 1$
9	$N \rightarrow succ \rightarrow E \rightarrow succ$	$simulate = 1$
10	W	$simulate = 1$

   Here, $N \rightarrow succ$ means that we select N then select that the outcome of N was successful, and $N \rightarrow slip(E)$ means that we select N and the outcome was unsuccessful: the agent went east.

   Draw the MCTS tree for this. Label the lines on the tree with the actions & outcomes and label the nodes with the backpropagated information.

2. Based on your tree, calculate which action should be returned.

3. Based on your tree, which of action, North, South, East, or West, will be the chosen one by a UCT? Show your work. Assume that $C_p = \frac{1}{2}$.