Artificial Neural Networks (Gerstner). Solutions for week 12

Deep Reinforcement Learning

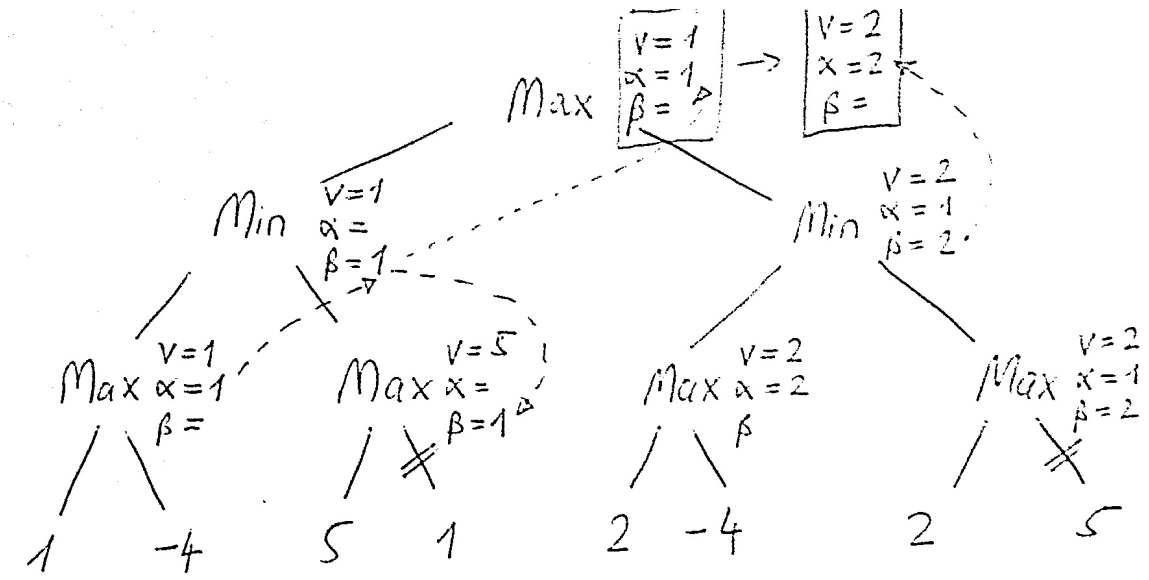
Exercise 1. Minimax with $\alpha - \beta$ pruning

Take the same decision tree as presented in class at the blackboard but replace the values at the bottom from left to right by 1, -4, 5, 1, 2, -4, 2, 5

- Determine the optimal action in the root node. Is it 'left' or 'right'?
- Indicate the pruning points.

Solution:

Minimax with $\alpha - \beta$-pruning will not be asked at the exam.

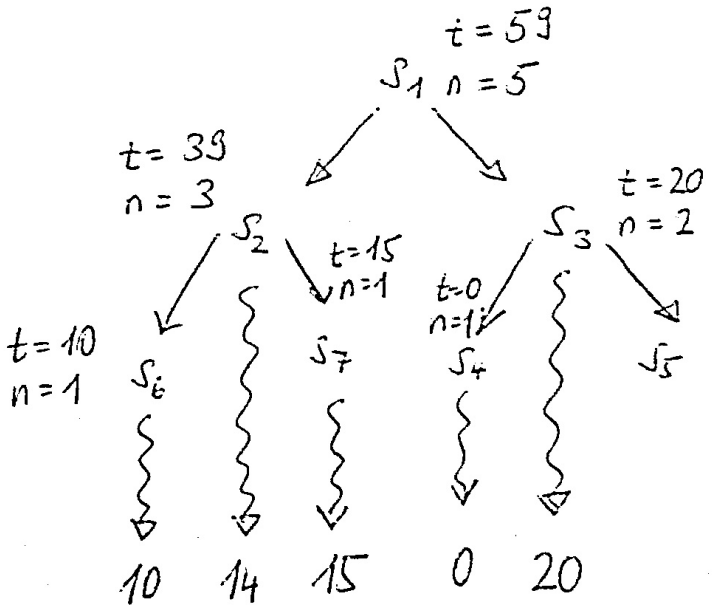


A nice worked example can be found at https://www.youtube.com/watch?v=zp3VMe0Jpf8.

Exercise 2. UCT

Run 5 iterations of UCT with two actions (left, right) in each node and the following order of returns from the simulations 14, 20, 0, 10, 15. Proceed as in the blackboard presentation in class and take the left node first if both children have the same UCB1 value.

Solution:


t = 59
S1 n = 5
t = 39
n = 3
S2
t = 15
n = 1
S3
t = 20
n = 2
t = 0
n = 1
t = 10
n = 1
S6
S7
S4
S5
10
14
15
0
20