

## Work Sheet 6

### Reinforcement Learning – Chapter 4

#### Exercise 1

##### Dynamic Programming

This exercise relates to the Gambler's Problem (Example 4.3) from Chapter 4 of [1]. It examines different approaches for the calculation of the optimal policy based on dynamic programming.

- a) Implement the Iterative Policy Evaluation and Improvement algorithms to find the optimal policy for the Gambler's Problem using the parameters from Example 4.3

Can you reproduce the plots from Figure 4.3?

- b) Implement the Value Iteration algorithm to find the optimal policy for the Gambler's Problem using the parameters from Example 4.3

Again, can you reproduce the plots from Figure 4.3?

- c) Use your implementations from 1 a) resp. 1 b) and experiment with the following parameters:

- Coin probabilities
- Values for  $\gamma$  in the algorithms
- Different rewards. Try, for example,  $r \in \{0.1, 1.0\}$  or  $r \in \{-0.1, 1.0\}$  instead of  $r \in \{0, 1\}$ .

Can you plausibly explain the calculated state values and policies?

## References

- [1] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2018.