Prof. Dr. Frank Deinzer
Technical University of Applied Sciences
Würzburg-Schweinfurt

Work Sheet 4

Reinforcement Leaning – Chapter 3

Exercise 1

Grid-World Environment.

a) Implement an *environment* and the corresponding *states*, *actions* and *rewards* for the gridworld example 3.5 in [1].

From an architectural perspective, your implementation must fit into the flowchart shown in Fig. 1:

- The environment has complete knowledge about the state of the gridworld. Especially about where you are.
- The environment will later get only an action A as input. In this example: one out of four possible movements.
- The environment returns as a result a state S and a reward R. In the simplest case, the state contains the coordinates where you are currently located *after executing the action*. The reward is a scalar.
- In the implementation it will be useful to be able to query at the environment which actions are executable in a certain state.

b) Extend the environment so that other gridworld examples from the book work with the implementation. Other relevant gridworld examples are those from example 3.8, 4.1, 6.5 and 6.6.

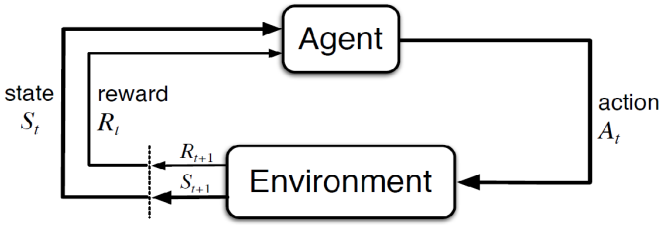

Fig. 1: Agent-Environment interaction

Exercise 2

Black-Jack Environment.

Implement an environment and the corresponding states, actions and rewards for our Black-Jack project. You can start with a simple scenario without special rules (like doubling down), casino rules, card counting, etc.

Make sure that your implementation is tested properly. The easiest way to do this is to allow you to play as a human against the environment.

Once the basic functionality is functional and tested, you can add more specific rules.

References

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