Assignment 9

Machine Learning, Summer term 2014, Norman Hendrich

To be discussed in exercise groups on June 23-25

Exercise 9.1 (Example reinforcement learning tasks, 2+2 points) Devise and describe two example tasks of your own that might be solved by the reinforcement learning approach. For both of your examples, identify the states and actions, and sketch a suitable rewards-function. Also estimate the size of the state-space: how many dimensions do you need to represent the (s, a) state-action space, and how many states are there in total? Try to make your examples as different from each other as possible, using different contexts and application areas.

Exercise 9.2 (Grid-world V-function, 1+2 points) The maximum value of V-function for the grid-world example (slides 60/66) is given as 24.4, to one decimal place.

- What is the optimal policy for this grid-world?

- For this optimal policy, derive the expression for the expected reward of the best state symbolically, then compute it to three decimal places (using $\gamma = 0.9$).

Exercise 9.3 (Bellman Equation for the Q-function, 3 points) Derive the Bellman equation for the action-value function $Q^\pi(s, a)$. The equation must give the action value $Q^\pi(s, a)$ in terms of the action values $Q^\pi(s', a')$ of possible successors to the state-action pair (s, a).