

## Introduction to Multi-Agent-Programming

B. Nebel, A. Kleiner  
C. Dornhege, D. Zhang  
Winter Semester 2010/2011

University of Freiburg  
Department of Computer Science

### Exercise Sheet 10

**Due: January 18th, 2011**

#### **Exercise 10.1** (Reinforcement Learning - Problem Formulation (2pt))

Consider a game of writing a string with nine chars, each char can have three possible values "X", "O", or "-". Each time, an agent can choose a place to put "X" in the string. Suppose the putting action is stochastic, and there is an oracle, which tells the agent the resulting string after each round, and three possible awards  $+1$ ,  $-1$ , or  $0$  after some rounds when the game is terminated. Otherwise rewards are  $0$ . Formulate the description to a Reinforcement-Learning problem in a formal form. Describe explicitly the state space, the actions, the transitions and the rewards.

#### **Exercise 10.2** (Policy-Iteration (3pt))

Consider the following grid world, where the numbers are rewards associated with the cells. The numbered cells are terminal states, unnumbered cells have a  $0$  reward. An agent starts at the left-bottom corner. It can perform four possible actions: North, South, East and West. With probability  $0.6$  the agent reaches the intended state, with probability  $0.2$  it moves to the right of the intended direction, and  $0.2$  to left (If the movement is not possible it does not move). Compute the optimal policy using Policy-Iteration.

