Advanced AI Techniques

Prof. Dr. Burgard, Prof. Dr. Nebel, Dr. Kersting
M. Ragni, A. Rottmann
University of Freiburg
Department of Computer Science
WS 2006/2007

Exercise Sheet 5

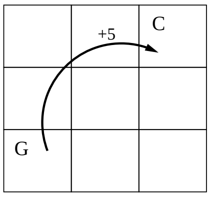
Due: Tuesday, 5. December 2006

Exercise 5.1

In the gridworld example rewards are positive for goals, negative for moving out of the world, and zero the rest of the time. Are the signs of these rewards important or only the intervals between them? Prove that adding a constant C to all the rewards adds a constant K to the values of all states, and thus does not affect the relative values of any states under any policies.

Exercise 5.2

Consider the following gridworld.



- At each cell the agent chooses an action $a \in \{north, east, south, west\}$ with equal probability.
- Actions that would take the agent off the grid will result in no movement and produces a reward of -1.
- All other actions produce a reward of 0 except at cell G. When the agent is in cell G he automatically goes to cell C regardless of it's choice of action and gets a reward of $+5$.

Determine the state-value function for equiprobable random policy of each cell with $\gamma = 0.9$.

Exercise 5.3

Determine the optimal policy for the gridworld from Exercise 5.2 using policy iteration.