UNIVERSITÄT BONN
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN
COMPUTER SCIENCE VI AUTONOMOUS INTELLIGENT SYSTEMS
Prof. Dr. Sven Behnke
Endenicher Allee 19a

Robot Learning

Assignment 4

Due Tuesday, May 19th, before class.

4.1) Consider the following grid world:

X	X	X	X	X	X	X	X	X	X	X	X	X	X
X	X	X	X	X									G
X													G
X													G
X													G
X													G
X	X										X	X	X
X	X								X	X	X	X	X
X	X							X	X	X	X	X	X
X	X	X						X	X	X	X	X	X
X	X	X						X	X	X	X	X	X
X	X	X	X					X	X	X	X	X	X
X	X	X	X					X	X	X	X	X	X
X	X	X	X				X	X	X	X	X	X	X
X	X	X	X				X	X	X	X	X	X	X
X	X	X					X	X	X	X	X	X	X
X	X	X					X	X	X	X	X	X	X
X	X	X	S	S	S	S	X	X	X	X	X	X	X

The agent always starts in one of the cells marked with S.
Its initial velocity is (0,1), i.e. it would move one cell upward.
Actions are to increase or decrease its velocity components by one or to leave it unchanged.
Both velocity components are restricted to be nonnegative and their sum must be at least 1 and cannot exceed 5.
The rewards are -1 for each step that the agent stays on the track, and -10 if the agent tries to drive off the track (enters a cell marked with X or leaves the grid).
The episode ends when the agent tries to drive off the track or when it crosses the finish line marked with G.

Compute V(s) for each grid cell using Monte Carlo Policy Evaluation for 1000 episodes of a policy that uniformly chooses one of the available actions.

5 Points

4.2) Compute the optimal policy using on-policy Monte Carlo control. Visualize the resulting trajectories for all start states S.

10 Points

4.3) When executing the velocity, the agent moves randomly one cell to far in either the horizontal or the vertical direction. Again, compute the optimal policy using on-policy Monte Carlo control. Visualize the resulting trajectories for all start states S.

5 points