

# Assignment: A7

---

Due: 6 December 2012

You are to explore the use of Reinforcement Learning as the basis for solving the cart pole problem. The method is used to generate a policy to keep the pole upright (0 degrees) for as long as possible. The state is a 4-tuple:  $[x, x_{\dot{}}, \theta, \theta_{\dot{}}]$ , where  $x$  is the  $x$  position of the cart,  $x_{\dot{}}$  is the velocity of the cart,  $\theta$  is the angle of the pole (the vector  $[0;1]$  is considered 0 degrees, and rotating toward  $[1;0]$  is a negative rotation down to a min of  $-\pi/2$ , and rotating toward  $[-1;0]$  is a positive rotation to a max of  $\pi/2$ ), and  $\theta_{\dot{}}$  is the angular velocity. Several aspects of this approach deserve careful attention:

- **Reward Function**: explore possible reward functions:
  1. The one given in the CS5350\_CP code
  2. Others of your choosing
- **Setup Issues**:
  1. Determine how the time step size impacts the results
  2. Determine how the number of episodes impacts the result
- **Policy Methods**:
  1. Compare greedy to epsilon-greedy
  2. Determine the impact of exploiting state symmetry
- **Data Management**:
  1. Describe how you set up the experiments to explore these issues

In addition, the results need to be presented in a strong statistical framework; this means computing statistics (e.g., mean, variance) over several trials (how many?), and showing confidence intervals.

Finally, the analysis and interpretation are the essential parts of the report; use these to present your findings, understanding and remaining problems.

In this assignment, the major goal is to explore the use of reinforcement learning algorithms.