

**Homework 6**

1. Consider the gardener's problem. Compute the optimal policy that maximizes the expected discounted reward with a discount factor of  $\alpha = 0.8$  using the LP method.
2. Consider a model with  $S = \{s_1, s_2\}$ , set of actions in state  $s_1$  as  $A_{s_1} = \{a_{1,1}, a_{1,2}\}$  and set of actions in state  $s_2$  as  $A_{s_2} = \{a_{2,1}\}$ . We have  $r(s_1, a_{1,1}) = 1$ ,  $r(s_1, a_{1,2}) = 0$  and  $r(s_2, a_{2,1}) = 2$ , and  $p(s_1|s_1, a_{1,1}) = 1$ ,  $p(s_2|s_1, a_{1,2}) = 1$ , and  $p(s_2|s_2, a_{2,1}) = 1$ . Suppose that you want to maximize the discounted infinite horizon expected reward.
  - a. Use value iteration to compute the optimal policy when the discount factor  $\alpha = 0.5$ .
  - b. Use value iteration to compute the optimal policy when the discount factor  $\alpha = 0.7$ .
  - c. Now use the LP method to solve parts (a) and (b).