

**Homework 8**  
**Due November 3, Friday**

- 1.** Consider a model with  $S = \{s_1, s_2, s_3\}$ ,  $A_{s_1} = \{a_{1,1}, a_{1,2}\}$  and  $A_{s_2} = \{a_{2,1}\}$ , and  $A_{s_3} = \{a_{3,1}\}$ ;  $r(s_1, a_{1,1}) = r(s_1, a_{1,2}) = 0$ ,  $r(s_2, a_{2,1}) = 3$ , and  $r(s_3, a_{3,1}) = 4$  and  $p(s_1|s_1, a_{1,1}) = p(s_2|s_1, a_{1,1}) = 1/2$ ,  $p(s_1|s_1, a_{1,2}) = 2/3$ ,  $p(s_3|s_1, a_{1,2}) = 1/3$ ,  $p(s_1|s_2, a_{2,1}) = 1$ , and  $p_1(s_1|s_3, a_{3,1}) = 1$ .

  - a.** Is this model unichain? Justify your answer.
  - b.** Use value iteration to solve this problem.
  - c.** Use policy iteration to solve this problem.
  - d.** Using this example show that the policy iteration may fail to find a bias optimal policy (i.e. a maximal gain policy which has greater bias than any other maximal gain policy).