Homework 2
Due September 29, Friday

1.Consider a model with $S = \{s_1, s_2, s_3\}$, $A_{s_1} = \{a_{1,1}, a_{1,2}\}$ and $A_{s_2} = \{a_{2,1}, a_{2,2}\}$, and $A_{s_3} = \{a_{3,1}\}$; $r(s_1, a_{1,1}) = r(s_1, a_{1,2}) = 0$, $r(s_2, a_{2,1}) = 3$, $r(s_2, a_{2,2}) = 4$ and $r_1(s_3, a_{3,1}) = 4$ and $p(s_1|s_1, a_{1,1}) = p(s_2|s_1, a_{1,1}) = 1/2$, $p(s_1|s_1, a_{1,2}) = 2/3$, $p(s_3|s_1, a_{1,2}) = 1/3$, $p(s_1|s_2, a_{2,1}) = 1$, $p(s_3|s_2, a_{2,2}) = 1$ and $p(s_1|s_3, a_{3,1}) = 1$. Suppose that the objective is to maximize the infinite horizon discounted reward with discount factor $\alpha = 0.8$. Use value iteration to solve the problem. Please submit all the outputs.

2. Show that $\mathcal{L}$ is a contraction mapping.

3. Suppose you have solved a discounted Markov decision process (with finite state space and finite action set) under maximization and have computed v_λ^* and an optimal policy d^∞ for which $v_\lambda^{d^\infty} = v_\lambda^*$. Suppose a new action a' becomes available in state s'. How can you determine whether d^∞ is still optimal without resolving the problem?