Name

Please be neat and show all your work so that I can give you partial credit. GOOD LUCK.

Question 1
Question 2
Question 3

Total
1. (15) a. Show that $|r(s,a)| \leq M$ implies that $\|v_\lambda^*\| \leq M(1 - \lambda)^{-1}$ where $v_\lambda^*$ is the optimal expected discounted total reward with discount factor $\lambda$.

(15) b. Consider a model with $T = \{1, 2\}$, $S = \{s_1, s_2\}$, $A_{s_1} = \{a_{1,1}, a_{1,2}\}$ and $A_{s_2} = \{a_{2,1}, a_{2,2}\}$, $r_1(s_1, a_{1,1}) = 5$, $r_1(s_1, a_{1,2}) = 10$, $r_1(s_2, a_{2,1}) = -1$, and $r_1(s_2, a_{2,2}) = 2$ and $p_1(s_1|s_1, a_{1,1}) = p_1(s_1|s_1, a_{1,2}) = 0.5$, $p_1(s_1|s_2, a_{1,2}) = 1$, $p_1(s_1|s_2, a_{2,1}) = 0.8$, $p_1(s_2|s_2, a_{2,1}) = 0.2$, $p_1(s_1|s_2, a_{2,2}) = 0.1$ and $p_1(s_2|s_2, a_{2,2}) = 0.9$. Compute $r_1(s_1, d_1(s_1))$ and $p_1(j|s, d_1(s))$ for the randomized decision rule $d_1$ which in state $s_1$ chooses action $a_{1,1}$ with probability $q$ and action $a_{1,2}$ with probability $1 - q$, and in state $s_2$ chooses action $a_{2,1}$ with probability $1$.
2. (40) Each quarter the marketing manager of a retail store divides customers into two classes based on their purchase behavior in the previous quarter. Denote the classes as $L$ for low and $H$ for high. The manager wishes to determine to which classes of customers he should send quarterly catalogs. The cost of sending a catalog is $15$ per customer and the expected purchase depends on the customer’s class and the manager’s action. If a customer is in class $L$ and receives a catalog, then the expected purchase in the current quarter is $20$ and if a class $L$ customer does not receive a catalog his expected purchase is $10$. If a customer is in class $H$ and receives a catalog, then his expected purchase is $50$, and if a class $H$ customer does not receive a catalog his expected purchase is $25$. The decision whether or not to send a catalog to a customer also affects the customer’s classification in the subsequent quarter. If a customer is class $L$ at the start of the present quarter, then the probability he is in class $L$ at the subsequent quarter is 0.3 if he receives a catalog and 0.5 if he does not. If a customer is class $H$ in the current period, then the probability that he remains in class $H$ in subsequent period is 0.8 if he receives a catalog and 0.4 if he does not. The objective of the store manager is to maximize his total expected discounted infinite horizon reward. Formulate this problem as an infinite horizon discounted Markov Decision Process problem with discount rate $\lambda$. Write down the optimality equations. Does there exist a Markovian stationary deterministic policy which is optimal? Justify your answer.
1. (15) a. Show that the sum of two superadditive functions is superadditive.

(15) b. Is the product of two superadditive functions is superadditive? Explain your answer.