Q3 Report

1. 'X' vector is solution by LP.

2. 'A*X' vector is AX matrix multiplication which is compared for equality with vector 'Alpha' in LP.

3. 'R' is reward vector. Here reward is for each action. So reward is counted with considering probability. (35 is my team number hence X=35) example, for cell (0, 1) - up action, reward is (3.5) 5. answer is the value to be maximized in LP. After solving LP, answer is 7.697 which is optimal utility of start state. In Value Iteration, utility of start state is 7.7 (discount factor = 1) The difference is less than delta as expected. Hence Answers of LP and VI match.