CAP4630 -- Problem Set on Markov Decision Processes

Question: Consider the following MDP described by three states (A, B, and C) and two actions (k and l) with transition probabilities and rewards (in parentheses) as shown in the tables below:

action k	to A	to B	to C
from A	0.3 (+2)	0.7 (+2)	
from B		1 (+2)	
from C		1 (+2)	

action l	to A	to B	to C
from A	1 (+3)		
from B	0.4 (+3)		0.6 (+3)
from C			1 (-5)

(a) Assume a 0.8 discount. Apply the Value Iteration algorithm to calculate the missing values in the value iteration table below. For each value you calculate, you must show in parentheses the action that produced such value.

Iteration	A	B	C
V_0	0	0	0
V_1	3(l)	3(l)	2(k)
V_2	5.4(l)		4.4(k)
V_3	7.32(l)	6.84(l)	

(b) Describe in plain English the policy that these results show.