aaBhatthana
03/22/2022

Artificial Intelligence HW2

a) $V_{k+1}(s) = \max_a \sum T(s,a,s')[R(s,a,s') + \gamma V_k(s')]$

States	0	2	3	4	5
V_0	0	0	0	0	0
V_1	0	(1)(2) = 2	(1)(3) = 3	(1)(4) = 4	(1)(5) = 5
V_2	(1/3)(0) = 0	0 + 2 = 0	(1)(3) = 3	(1)(4) = 4	(1)(5) = 5
V_3	(1/3)(2) = 2/3	11/3	3	4	5
V_4	2/3	11/3	3	4	5

b)

States	0	2	3	4	5
π^*	Move	Move	Stop	Stop	Stop

c) By decreasing γ to 0.1, it would result in a Move, Stop, Stop, Stop, Stop since it would have ~~favor~~ more short-term rewards rather than longterm rewards.

d)

States	0	2	3	4	5
π_0	Move	Stop	Move	Stop	Move
$V_0^{\pi_0}$	0	0	0	0	0
$V_1^{\pi_0}$	0	2	2	4	4
$V_2^{\pi_0}$	2	2	2	4	4
$V_3^{\pi_0}$	2	2	2	4	4
π_1	Move	Stop	Stop	Stop	Stop
$V_0^{\pi_1}$	0	0	0	0	0
$V_1^{\pi_1}$	0	2	3	4	5
$V_2^{\pi_2}$	2	2	3	4	5
$V_3^{\pi_3}$	2	2	3	4	5
π_2	Move	Move	Stop	Stop	Stop

I am a bit confused on this