

Evaluating Policy π^0 , Second Iteration

S	$V^1(S)$	$V^2(S)$
(0,0)	0	$R([0,0], \pi^0([0,0])) + \gamma [T([0,0], \pi^0([0,0]), [0,1]) \times V^1([0,1]) + T([0,0], \pi^0([0,0]), [1,0]) \times V^1([1,0])]$ $= 0 + 0.9 \times [0.9 \times (-5) + 0.1 \times 0] = -4.05$
(0,1)	-5	$R([0,1], \pi^0([0,1])) + \gamma [T([0,1], \pi^0([0,1]), [0,2]) \times V^1([0,2]) + T([0,1], \pi^0([0,1]), [0,0]) \times V^1([0,0]) + T([0,1], \pi^0([0,1]), [1,1]) \times V^1([1,1])]$ $= -5 + 0.9 \times [0.9 \times 10 + 0.05 \times 0 + 0.05 \times 0]$ $= 3.1$
(0,2)	+10	$+10 + 0.9 \times [1 \times 10] = +19$
(1,0)	0	$0 + 0.9 \times [0.9 \times 0 + 0.05 \times 0 + 0.05 \times 0] = 0$
(1,1)	0	$0 + 0.9 \times [0.9 \times 0 + \frac{0.1}{3}(-5) + \frac{0.1}{3} \times 0 + \frac{0.1}{3} \times 0] = -0.15$
(1,2)	0	$0 + 0.9 \times [1 \times 0] = 0$
(2,0)	0	$0 + 0.9 \times [0.9 \times 0 + 0.1 \times 0] = 0$
(2,1)	0	$0 + 0.9 \times [0.9 \times 0 + 0.05 \times 0 + 0.05 \times 0] = 0$
(2,2)	0	$0 + 0.9 \times [1 \times 0] = 0$

Final Values of Policy π^0

State	0,0	0,1	0,2	1,0	1,1	1,2	2,0	2,1	2,2
$V(\text{State})$	-4.05	3.1	19	0	-0.15	0	0	0	0