Actions: {Left, Right, Up, Down, Nothing}

Immediate rewards, in this example depend only on states, not actions.

Example: $R((0,0), Right) = R((0,0), left) =$
$R((0,0), Up) = R((0,0), Down) = 0$
$R((0,1), Right) = R((0,1), left) =$
$R((0,1), Up) = R((0,1), Down) = -5$

(0,0)	(0,1)	(0,2)
0	-5	+10
(1,0)	(1,1)	(1,2)
0	0	0
(2,0)	(2,1)	(2,2)
0	0	0

State Space an rewards.

The transition probabilities for every state and action are given in the description.

$\gamma = 0.9$

Initial Policy π^0

State	0,0	0,1	0,2	1,0	1,1	1,2	2,0	2,1	2,2
π^0(state)	Right	Right	Right	Right	Right	Right	Right	Right	Right

Evaluating Policy π^0

S	Initial values $V^0(S)$	$V^1(S)$
(0,0)	0	$R((0,0), \pi^0[0,0]) + \gamma[T([0,0], \pi^0([0,0]), [0,1]) \times V^0([0,1]) + T([0,0], \pi^0([0,0]), [1,0]) \times V^0([1,0])]$ $= 0 + 0.9 \times [0.9 \times 0 + 0.1 \times 0] = 0$
(0,1)	0	$R([0,1], \pi^0[0,1]) + \gamma \times [T([0,1], \pi^0[0,1], [0,2]) V^0([0,2]) + T([0,1], \pi^0[0,1], [0,0]) V^0([0,0]) + T([0,1], \pi^0[0,1], [1,1]) V^0([1,1])]$ $= -5 + 0.9 \times [0.9 \times 0 + 0.05 \times 0 + 0.05 \times 0]$ $= -5$
(0,2)	0	$R([0,2], \pi^0([0,2])) + \gamma \times [T([0,2], \pi^0([0,2]), [0,2]) V^0([0,2])]$ $= +10 + 0.9 \times [1 \times 0] = +10$
(1,0)	0	. . . . .