S		-5
		+5

S \ a	↑	→	↓	←
0,0	0	0	0	0
0,1	0	0	0	0
1,0	0	0	0	0
1,1	0	0	0	0

$\alpha = 0.5$; $\gamma = 0.5$

-1 între

1. Primul episod: (0,0) - (0,1) - (0,2)

~~(1,1) → (1,2)~~ (0,0) → (0,1)

$$Q((0,0),\rightarrow) = Q((0,0),\rightarrow) + \alpha \cdot (-1 + \gamma \max_{a'} Q((0,1),a') - Q((0,0),\rightarrow))$$

$$Q((0,0),\rightarrow) = 0 + 0.5(-1 + 0.5 \cdot 0 - 0) = 0 + 0.5 \cdot (-1) = -0.5$$

(0,1) → (0,2)

$$Q((0,1),\rightarrow) = Q((0,1),\rightarrow) + \alpha(-5 + \gamma \max_{a'} Q((0,2),a') - Q((0,1),\rightarrow)$$

$$= 0 + 0.5(-5 + 0.5 \cdot 0 - 0) = -2.5$$

2. Episodul 2: (0,0) → (0,1) → (1,1) → (1,2)

Tabelul arată așa acum

S \ a	↑	→	↓	←
0,0	0	-0.5	0	0
0,1	0	-2.5	0	0
1,0	0	0	0	0
1,1	0	0	0	0

$(0,0) \rightarrow (0,1)$

$Q((0,0),\rightarrow) = Q((0,0),\rightarrow) + \alpha \cdot (-1 + \gamma \max_{a'} Q((0,1),a') - Q((0,0),\rightarrow))$

$= -0.5 + 0.5 \cdot (-1 + 0 - (-0.5)) = -0.5 + 0.5 \cdot (-0.5) = -0.75$

$(0,1) \rightarrow (1,1)$

$Q((0,1),\downarrow) = Q((0,1),\downarrow) + \alpha \cdot (0 + \gamma \max_{a'} Q((1,1),a') - Q((0,1),\downarrow))$

$Q((0,1),\downarrow) = 0 + 0.5(0 + 0.5 \cdot 0 - 0) = -0.5$

$(1,1) \rightarrow (1,2)$

$Q((1,1),\rightarrow) = Q((1,1),\rightarrow) + \alpha \cdot (5 + \gamma \max_{a'} Q((1,1),a') - Q((1,1),\rightarrow))$

$Q((1,1),\rightarrow) = 0 + \alpha(5 + 0 - 0) = 0.5 \cdot 5 = 2.5$

Tabel final

S \ a	↑	→	↓	←
0,0	0	-0.75	0	0
0,1	0	-2.5	-0.5	0
1,0	0	0	0	0
1,1	0	2.5	0	0