$Q(S_4, R)$

Lets start at S_1.

policy of going right

$$Q(S_1, R) = (0.85)[R(S_1, R, S_2) + (0.9)\underbrace{\cancel{(0.6)}}_{\max(-0.8,\,0.6)}]$$

$$\cancel{+0.05[R} + 0.10[R(S_1, R, S_1) + (0.9)\underbrace{[}_{\max(0,0)}$$

$$= (0.85)(R(S_1, R, S_2) + 0.54) -$$

$$+ 0.1(R(S_1, R, S_1)] -$$

$\cancel{Q(S_3, L)} =$

$$Q(S_2, R) = (0.85)(\cancel{3+2}\,0.6 + (0.9)(3.2$$

$$= 0.72$$

$$Q(S_3, R) = (0.85)(2 + (0.9)(2)$$

$$\cancel{+(0.1)(}$$

$$= 3.23$$

$\cancel{Q(S_2, L) =}$

$$Q(S_2, L) = (0.85)(-0.8 - 1)$$
$$= (0.85)(-1.8) = -1.53$$

$$Q(S_3, L) = (0.85)(-1.53 - (0.9)(0.72)]$$
$$= -0.94$$

Q(S, A)	L	R
S_1	0	0
S_2	-1.53	0.72
S_3	-0.94	3.23
S_4	0	0

Thus, the optimal policy is trying to go~~ing~~ right at each position.

ie. $S_1 \xrightarrow{R} S_2 \xrightarrow{R} S_3 \xrightarrow{R} S_4$.