

a.	-2	-1	0	1	2
$V_{opt}(s)$	0	0	0	0	0
	0		0		

$$V_{opt}(-2) = 0$$

$$V_{opt}(-1) = \max_{a \in \{-1, +1\}} Q_{opt}(-1, a)$$

$$\begin{aligned} Q_{opt}(-1, -1) &= \sum_{s'} T(-1, -1, s') [Rewards(-1, -1, s') + \gamma V_{opt}(s')] \\ &= T(-1, -1, -2) [Rewards(-1, -1, -2) + (\cdot) V_{opt}(-2)] + \\ &\quad T(-1, -1, 0) [Rewards(-1, -1, 0) + (\cdot) V_{opt}(0)] \\ &= 0.8 [20 + V_{opt}(-2)] + 0.2 [-5 + V_{opt}(0)] \\ &= 0.8 [20 + 0] + 0.2 [-5 + V_{opt}(0)] \end{aligned}$$

$$\begin{aligned} Q_{opt}(-1, +1) &= \sum_{s'} T(-1, +1, s') [Rewards(-1, +1, s') + \gamma V_{opt}(s')] \\ &= T(-1, +1, 0) [Rewards(-1, +1, 0) + (\cdot) V_{opt}(0)] + \\ &\quad T(-1, +1, -2) [Rewards(-1, +1, -2) + (\cdot) V_{opt}(-2)] \\ &= 0.7 [-5 + V_{opt}(0)] + 0.3 [20 + V_{opt}(-2)] \\ &= 0.7 [-5 + V_{opt}(0)] + 0.3 [20 + 0] \end{aligned}$$















