

CGIS 4100
TDL exercise

	1	2	3	4
3	.67	.70	.80	1
2	.50		.55	-1
1	.43	.32	.45	.20

} Populated with current \hat{V}^π estimates

1. observe $(3,3) \rightarrow (3,4)$; $\gamma=1$, $\alpha=.1$ (given)

$$\begin{aligned} \hat{V}^\pi(3,3) &\leftarrow (1-\alpha)\hat{V}^\pi(3,3) + \alpha \{ R((3,3) \rightarrow (3,4)) + \gamma \hat{V}^\pi(3,4) \} \\ &= (.9)(.8) + (.1) \{ -.1 + 1 \} \\ &= .81 \end{aligned}$$

2. $(2,3) \rightarrow (2,4)$

$$\hat{V}^\pi(2,3) \leftarrow \overbrace{(.9)(.55)}^{\text{weighted current est.}} + \underbrace{(.1)}_{\text{learning rate}} \{ -.1 + (-1) \}$$

$$= 0.385$$

3. $(2,3) \rightarrow (3,3)$

$$\hat{V}^\pi(2,3) \leftarrow (.9)(.385) + (.1) \{ -.1 + .8 \}$$

$$\approx 0.417$$