

Policy Iteration Calculations

Default State

$$\lambda = 0.9$$

$V_0 \rightarrow 0$	$V_1 \rightarrow 0$	$V_2 \rightarrow 0$	scribble
$V_4 \rightarrow 0$	$V_5 \rightarrow 0$	scribble	$V_7 \rightarrow 0$
$V_8 \rightarrow 0$	$V_9 \rightarrow 0$	$V_{10} \rightarrow 0$	$V_{11} \rightarrow 0$
$V_{12} \rightarrow 0$	$V_{13} \rightarrow 0$	scribble	$V_{15} \rightarrow 0$

First Iteration $V_i = r(i) + \lambda V(S')$

$$V_0 = -0.04 + 0.9[0.8(0) + 0.1(0) + 0.1(0)]$$

$$= \boxed{-0.04}$$

$$V_1 = -0.04 + 0.9[0.8(0) + 0.1(0) + 0.1(0)]$$

$$= \boxed{-0.04}$$

$$V_2 = -0.04 + 0.9[0.8(0) + 0.1(-0.04) + 0.1(0)]$$

$$= -0.04 + 0.9(-0.004) = \boxed{-0.0436}$$

$$V_4 = \boxed{-1000} \quad V_5 = \boxed{1000} \quad 1000 + 0.9(0)$$

$$V_7 = -0.04 + 0.9[0.8(0) + 0.1(0) + 0.1(0)]$$

$$= \boxed{-0.04}$$

$$V_8 = -0.04 + 0.9[0.8(0) + 0.1(-1000) + 0.1(0)]$$

$$= -0.04 + 0.9[-100] = \boxed{-90.04}$$

$$V_9 = -0.04 + 0.9[0.8(1000) + 0.1(-90.04) + 0.1(0)]$$

$$= -0.04 + 0.9[790.996] = \boxed{711.8564}$$

$$V_{10} = -0.04 + 0.9[0.8(0) + 0.1(0) + 0.1(0)]$$

$$= \boxed{-0.04}$$

$$V_{11} = -0.04 + 0.9[0.8(-0.04) + 0.1(-0.04) + 0.1(0)]$$

$$= -0.04 + 0.9(-0.036) = \boxed{-0.0724}$$

$$V_{12} = -0.04 + 0.9[0.8(0) + 0.1(-90.04) + 0.1(0)]$$

$$= -0.04 + 0.9[-9.004] = \boxed{-8.1436}$$

$$V_{13} = -0.04 + 0.9[0.8(-8.1436) + 0.1(711.8564) + 0.1(0)]$$

$$= -0.04 + 0.9[64.67076] = \boxed{58.163684}$$

$$V_{15} = -0.04 + 0.9[0.8(0) + 0.1(-0.0724) + 0.1(0)]$$

$$= -0.04 + 0.9(-0.00724) = \boxed{-0.046516}$$

First Iteration State

\rightarrow $V_0 = -0.04$	\rightarrow $V_1 = -0.04$	\downarrow $V_2 = -0.0436$	
$V_4 = -1000$	$V_5 = 1000$		\leftarrow $V_7 = -0.04$
\rightarrow $V_8 = -90.04$	\uparrow $V_9 = 711.8564$	\rightarrow $V_{10} = -0.04$	\leftarrow $V_{11} = -0.0724$
\rightarrow $V_{12} = -8.1436$	\leftarrow $V_{13} = 58.1637$		\leftarrow $V_{15} = -0.0465$

Second Iteration

$$V_0 = -0.04 + 0.9[0.8(-0.04) + 0.1(-0.04) + 0.1(-1000)]$$

$$= -0.04 + 0.9[-100.036] = \boxed{-90.0724}$$

$$V_1 = -0.04 + 0.9[0.8(-0.0436) + 0.1(-0.04) + 0.1(1000)]$$

$$= -0.04 + 0.9[99.96112] = \boxed{89.925008}$$

$$V_2 = -0.04 + 0.9[0.8(-0.0436) + 0.1(89.925008) + 0.1(-0.0436)]$$

$$= -0.04 + 0.9[8.9532608] = \boxed{8.01793472}$$

$$V_4 = -1000 + 0.9(-1000) = \boxed{-1900}$$

$$V_5 = 1000 + 0.9(1000) = \boxed{1900}$$

$$V_7 = -0.04 + 0.9[0.8(-0.04) + 0.1(-0.04) + 0.1(-0.0724)]$$

$$= -0.04 + 0.9[-0.04324] = \boxed{-0.078916}$$

$$V_8 = -0.04 + 0.9[0.8(711.8564) + 0.1(-1900) + 0.1(-8.1436)]$$

$$= -0.04 + 0.9[378.67076] = \boxed{340.763684}$$

$$V_9 = -0.04 + 0.9[0.8(1900) + 0.1(340.763684) + 0.1(-0.04)]$$

$$= -0.04 + 0.9[1554.0723684] = \boxed{1398.62513156}$$

$$V_{10} = -0.04 + 0.9[0.8(-0.0724) + 0.1(-0.04) + 0.1(-0.04)]$$

$$= -0.04 + 0.9[-0.06592] = \boxed{-0.099328}$$

~~$$V_{11} = -0.04 + 0.9[0.8(-0.04) + 0.1(-0.04) + 0.1(-0.04)]$$~~

$$V_{11} = -0.04 + 0.9[0.8(-0.099328) + 0.1(-0.078916) + 0.1(-0.0465)]$$

$$= -0.04 + 0.9[-0.0920056] = \boxed{-0.12280504}$$

$$V_{12} = -0.04 + 0.9[0.8(58.163684) + 0.1(340.763684) + 0.1(-8.1436)]$$

$$= -0.04 + 0.9[79.7929556] = \boxed{71.77366004}$$

$$V_{13} = -0.04 + 0.9 [0.8(71.77366004)^{V_{12}} + 0.1(711.8564)^{V_9} + 0.1(58.163684)^{V_{13}}]$$

$$= -0.04 + 0.9 [134.429936432] = \boxed{120.938842789}$$

$$V_{15} = -0.04 + 0.9 [0.8(-0.046516)^{V_{15}} + 0.1(-0.12280504)^{V_{11}} + 0.1(0.046516)^{V_{15}}]$$

$$= -0.04 + 0.9 [-0.054144904] = \boxed{-0.0887304136}$$

Second Iteration State

\rightarrow $V_0 = -90.0724$	\rightarrow $V_1 = 87.9250$ $V_1 = 14.92$	\downarrow $V_2 = 8.0179$	~~~~~
$V_4 = -1900$	$V_5 = 1900$	~~~~~	\leftarrow $V_7 = -0.0789$
\rightarrow $V_8 = 340.7637$	\uparrow $V_9 = 1398.6251$	\rightarrow $V_{10} = -0.0993$	\leftarrow $V_{11} = -0.1228$
\rightarrow $V_{12} = 71.7737$	\leftarrow $V_{13} = 120.9388$	~~~~~	\leftarrow $V_{15} = -0.0887$