

The optimum values and policy for each gamma (for both deterministic and stochastic environment) are highlighted in the table below.

Observation: Gamma = 1 leads to an infinite loop when environment is deterministic. Gamma = 1 leads to an optimum policy in a non-deterministic (stochastic) environment.

For deterministic, the number of iterations stays the same for $0.5 \leq \gamma < 1.0$ at $N = 28$. But for $\gamma < 0.50$, it is finite and less than $N = 28$.

For stochastic, the iteration count increases with gamma (tested till gamma = 1).

Note: Not all values and their iteration may be depicted in the values in the below table. They depict the change itself though but putting every iteration would have been difficult.

Deterministic environment:

Gamma	Values	Policy	Iteration
0.5	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 1.] [0. 0. 1. 0.]]	[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]	28
	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0.5 1.] [0. 0.5 1. 0.]]	[0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 2 0]	
	[[0. 0. 0. 0.] [0. 0. 0.25 0.] [0. 0.25 0.5 1.] [0. 0.5 1. 0.]]	[0 0 0 0 0 0 0 0 0 0 0 1 1 0 2 2 0]	
	[[0. 0. 0.125 0.0625] [0. 0. 0.25 0.] [0.125 0.25 0.5 1.] [0. 0.5 1. 0.]]	[0 0 0 0 0 0 1 0 0 1 1 1 0 2 2 0]	
	[[0. 0.0625 0.125 0.0625] [0.0625 0. 0.25 0.] [0.125 0.25 0.5 1.] [0. 0.5 1. 0.]]	[0 0 1 0 0 0 1 0 2 1 1 1 0 2 2 0]	
	[[0.03125 0.0625 0.125 0.0625] [0.0625 0. 0.25 0.] [0.125 0.25 0.5 1.] [0. 0.5 1. 0.]]	[0 2 1 0 1 0 1 0 2 1 1 1 0 2 2 0]	
	[[0.03125 0.0625 0.125 0.0625] [0.0625 0. 0.25 0.] [0.125 0.25 0.5 1.] [0. 0.5 1. 0.]]	[1 2 1 0 1 0 1 0 2 1 1 1 0 2 2 0]	

0.8	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 1.] [0. 0. 1. 0.]]	[000000000000000000]	28
	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0.8 1.] [0. 0.8 1. 0.]]	[00000000000010020]	
	[[0. 0. 0. 0.] [0. 0. 0.64 0.] [0. 0.64 0.8 1.] [0. 0.8 1. 0.]]	[00000000000110220]	
	[[0. 0. 0.512 0.4096] [0. 0. 0.64 0.] [0.512 0.64 0.8 1.] [0. 0.8 1. 0.]]	[0000001001110220]	
	[[0. 0.4096 0.512 0.4096] [0.4096 0. 0.64 0.] [0.512 0.64 0.8 1.] [0. 0.8 1. 0.]]	[0010001021110220]	
	[[0.32768 0.4096 0.512 0.4096] [0.4096 0. 0.64 0.] [0.512 0.64 0.8 1.] [0. 0.8 1. 0.]]	[0210101021110220]	
	[[0.32768 0.4096 0.512 0.4096] [0.4096 0. 0.64 0.] [0.512 0.64 0.8 1.] [0. 0.8 1. 0.]]	[1210101021110220]	
1.0	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 0.]]	[000000000000000000]	∞ Goes into an infinite loop unless iteration is limited.
	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 1.] [0. 0. 1. 0.]]	[00000000000010020]	
	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 1.] [0. 0. 1. 0.]]	[00000000000110210]	
	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 1. 1.] [0. 0. 1. 0.]]	[00000000000210020]	
	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 1.] [0. 0. 0. 0.]]	[0000001002100210] The policies repeat themselves on a loop.	

Stochastic environment:

Gamma	Values	Policy	Iteration
0.5	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 0.33333333] [0. 0. 0. 0.]]	[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] 2 times	34
	[[0. 0. 0.00463677 0.00090616] [0. 0. 0.02372844 0.] [0. 0. 0.1378927 0.42755005] [0. 0. 0.39999143 0.]]	[0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0] 6 times	
	[[0. 0.00075732 0.00507735 0.00115582] [0. 0. 0.02533476 0.] [0. 0.03870766 0.14729074 0.42941593] [0. 0.08575326 0.42941593 0.]]	[0 1 2 3 0 0 0 0 0 1 2 1 0 1 2 0] 6 times	
	[1.34316415e-03 0.00000000e+00 2.58156229e-02 0.00000000e+00]	[1 2 2 3 0 0 0 0 1 1 1 1 0 2 2 0] 6 times	
	[[2.14334705e-04 8.50194330e-04 5.15474966e-03 1.17110101e-03] [1.34316415e-03 0.00000000e+00 2.58156229e-02 0.00000000e+00] [8.00182899e-03 4.18565910e-02 1.50225376e-01 4.29995534e-01] [0.00000000e+00 9.41671919e-02 4.29995534e-01 0.00000000e+00]]	[2 3 2 3 0 0 0 0 3 1 1 1 0 2 2 0] 7 times	
	[[4.10808185e-04 1.07524577e-03 5.35701150e-03 1.28320225e-03] [1.67161224e-03 0.00000000e+00 2.59316327e-02 0.00000000e+00] [8.63762246e-03 4.21742507e-02 1.50361331e-01 4.30059700e-01] [0.00000000e+00 9.43910361e-02 4.30059700e-01 0.00000000e+00]]	[1 3 2 3 0 0 0 0 3 1 1 1 0 2 2 0] 7 times	
0.8	[[0. 0. 0. 0.] [0. 0. 0. 0.] [0. 0. 0. 0.33333333] [0. 0. 0. 0.]]	[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0] 2 times	48
	[[0. 0. 0.02828108 0.0101021] [0. 0. 0.07902366 0.] [0. 0. 0.2683217 0.55200699] [0. 0. 0.45453383 0.]]	[0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0] 8 times	
	[[0. 0.01191161 0.04530546 0.02504508] [0. 0. 0.10115333 0.] [0. 0.14489444 0.33415233 0.57602968] [0. 0.20941162 0.57602968 0.]]	[0 1 2 3 0 0 0 0 0 1 2 1 0 1 2 0] 12 times	
	[[0.01662442 0.01736148 0.04908253 0.02702558] [0.03066914 0. 0.10999736 0.] [0.06897085 0.19052418 0.36365954 0.58673576] [0. 0.28245626 0.58673576 0.]]	[1 2 2 3 0 0 0 0 1 1 1 1 0 2 2 0] 12 times	
	[[0.02282036 0.02618203 0.0501801 0.02816727] [0.03876929 0. 0.11082404 0.]]	[1 3 2 3 0 0 0 0 3 1 1 1 0 2 2 0]	

	[0.08507851 0.19602848 0.36554483 0.58744724] [0. 0.28480774 0.58744724 0.]]	14 times	
1.0	Too many, my editor was not able to print and keep all of it.	378
	[[0.98097585 0.73857834 0.49692524 0.49657118] [0.98375773 0. 0.49773146 0.] [0.98798739 0.99318196 0.996451 0.99817884] [0. 0.99551893 0.99817884 0.]]	[0 3 0 3 0 0 0 0 3 1 1 1 0 2 2 0]	
	[[0.98097585 0.8417415 0.70383197 0.70263833] [0.98375773 0. 0.56670037 0.] [0.98798739 0.99318196 0.996451 0.99817884] [0. 0.99551893 0.99817884 0.]]	[0 3 3 3 0 0 0 0 3 1 1 1 0 2 2 0]	
	[[0.99186477 0.98651186 0.98271931 0.98077437] [0.99305437 0. 0.66037463 0.] [0.99486309 0.99708442 0.99848235 0.99922122] [0. 0.99808378 0.99922122 0.]]	[0 3 3 3 0 0 0 0 3 1 1 1 0 2 2 0]	