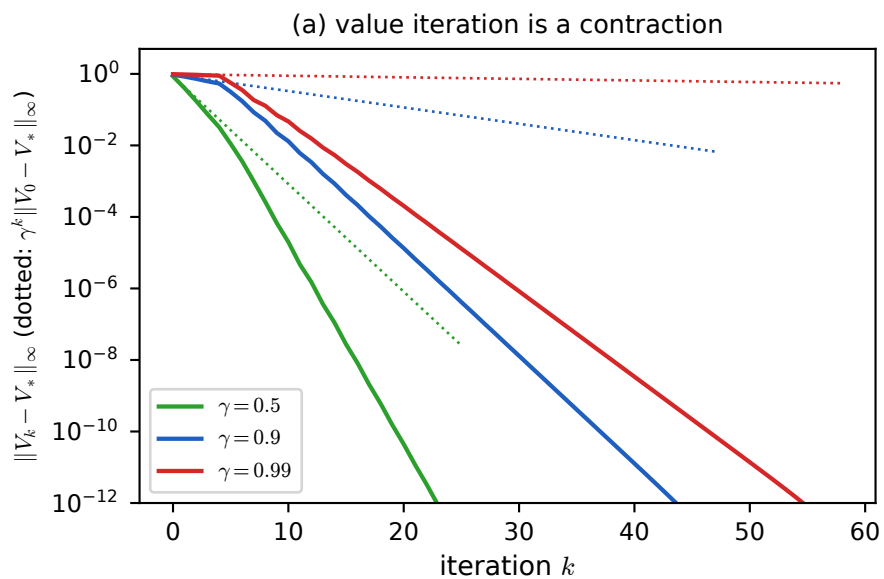
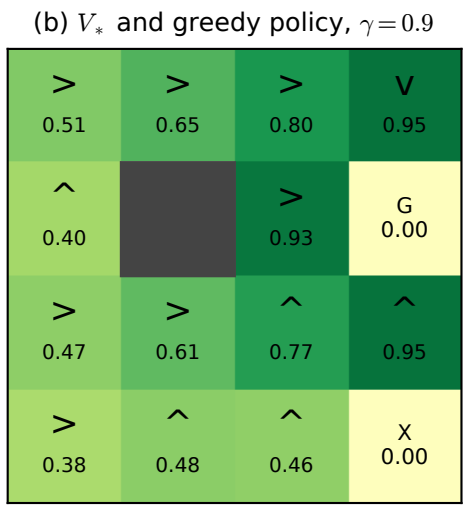

(a) value iteration is a contraction
$\|V_k - V_*\|_\infty$ (dotted: $\gamma^k \|V_0 - V_*\|_\infty$)
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
0
10
20
30
40
50
60
iteration k
$\gamma = 0.5$
$\gamma = 0.9$
$\gamma = 0.99$


(b) V_* and greedy policy, $\gamma = 0.9$
> 0.51
> 0.65
> 0.80
v 0.95
^ 0.40
> 0.93
G 0.00
> 0.47
> 0.61
^ 0.77
^ 0.95
> 0.38
^ 0.48
^ 0.46
X 0.00