Yiping Guo
S3177105
31 May 2019

# Report on how stochasticity affect the number of iterations required and the resulting policy

Given a large enough iteration number, value iteration on deterministic environments is converge to the optimal values in deterministic environments. In a stochastic environment, though there is uncertainty about the actions effect, since we are adding the true reward during each iteration, value iteration is increasingly closer to its real optimal value at each next iteration. In experiment with Frozenlake, the agent program need more than 10 times the number of iterations in stochastic world to get a policy to reach goal state.

Given a large enough iteration number, policy iteration is guaranteed to converge to the optimal policy in deterministic environments. In a stochastic environment, with the iteration number increases, policy iteration gets a better policy closer to the true optimal policy than last iteration.

Differences between value iteration and policy iteration methods in general is **value iteration** keeps improving the value function estimate at each iteration until the value-function converges while **policy iteration** re-define the policy at each step and compute the value according to this new policy until the policy converges. Policy iteration is also guaranteed to converge to the optimal policy and it often takes less iterations to converge than value iteration.