



Policy improvement: Estimate V^π
Iterative policy evaluation

Policy evaluation: Generate $\pi' \geq \pi$
Greedy policy improvement

