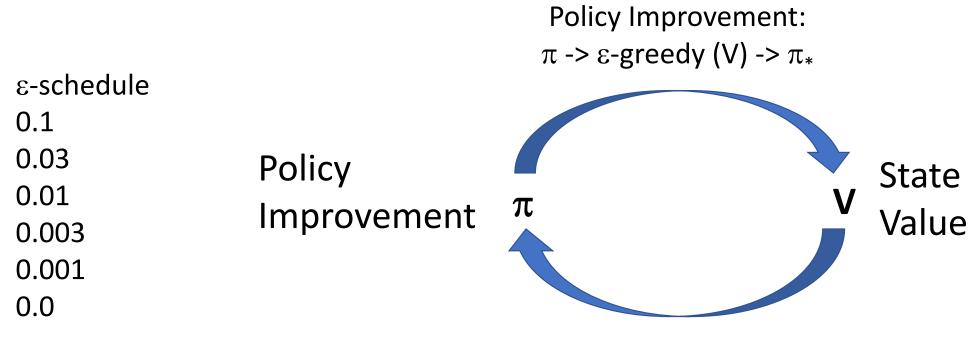
Geneal Policy Iteration

Iterative algorithm alternates state value policy evaluation and action value policy improvement:



Evaluation: $V \rightarrow v_{\pi} \rightarrow V_{*}$