

General Policy Iteration

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

ϵ -schedule

0.1

0.03

0.01

0.003

0.001

0.0

Policy
Improvement

