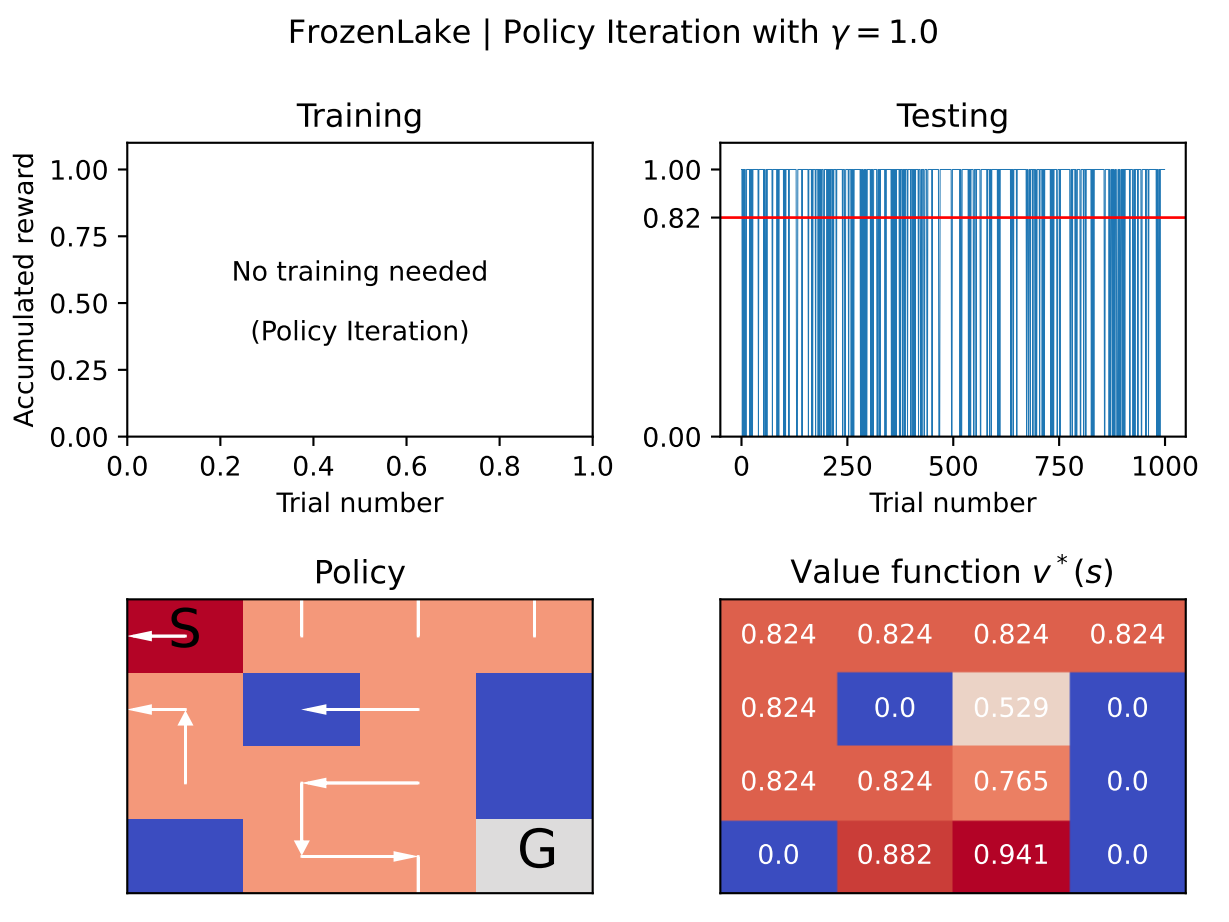
FrozenLake | Policy Iteration with $\gamma = 1.0$
Training
Accumulated reward
1.00
0.75
0.50
0.25
0.00
No training needed
(Policy Iteration)
0.0
0.2
0.4
0.6
0.8
1.0
Trial number
Testing
1.00
0.82
0.00
0
250
500
750
1000
Trial number
Policy
S
G
Value function $v^*(s)$
0.824 0.824 0.824 0.824
0.824 0.0 0.529 0.0
0.824 0.824 0.765 0.0
0.0 0.882 0.941 0.0