

Deterministic optimal policy for DYNA-Q+ AS 0.5 after 6000 iterations.

