

Deterministic optimal policy for DYNA-Q+ AS after 12000 iterations.

