

Please enter input here:15 12 8 6 q 11

Q-values after 10000 iterations:

State 11 has optimal Q-values:

↑ 19.90

→ 19.90

↓ 0.68

← 0.68