Monte Carlo Prediction: 100.000 episodes

Discount = 1

Text

Description automatically generated

Discount = 0.9

Text

Description automatically generated

Temporal Difference Learning: 100.000 episodes

Discount = 1

Text

Description automatically generated

Discount = 0.9

Text

Description automatically generated

Monte Carlo Control: 100.000 episodes

Discount = 1

Background pattern

Description automatically generated

Discount = 0.9

Background pattern

Description automatically generated

SARSA: 100.000 episodes

Discount = 1

Background pattern

Description automatically generated

Discount = 0.9

Background pattern

Description automatically generated

Q-learning: 100.000 episodes

Discount = 1

Background pattern

Description automatically generated

Discount = 0.9

Background pattern

Description automatically generated

Double Q-leaning: 100.000 episodes (Q1 + Q2)

Discount = 1

Background pattern

Description automatically generated

Discount = 0.9

Background pattern

Description automatically generated