Reinforcement Learning

Value Function

discrete action-space

Discrete state-space
(model based)

Dynamic Programming
(model based)

Bellman Equation

Bootstrapping

continuous state-space
(model free)

DQN

Double Q-Learning

delayed Q-learning

Duelling networks

prioritized replay

adding noise

multi-step Q-learning

Discrete state-space (model free)

Temporal Difference Learning

Monte-Carlo estimates

SARSA

Expected SARSA

TD(lambda)

TD(n)

Watkins Q-Learning

Q(sigma)

Actor Critic

A2C, A3C

DDPG - Family

deterministic policy

Dueling DDPG

Prioritized DDPG

TD3

SAC
(stochastic policy with sigma as output)

Policy Gradient

continuous action-space

REINFORCE
(Vanilla Policy Gradient)

TRPO
(some degree off-policy)

PPO
(some degree off-policy)
(simplification of TRPO)

on-policy

off-policy

Importance Sampling

Experience Replay