| Chapters | Main points |
|---|---|
| RL1-2-Lightweight-Cases<br>RL2-1-MDP-DP<br>RL2-2 Model-Free | UCB<br>MDP<br>Q-learning<br>Contraction Mapping |
| RL3-1 Value-Based<br>RL3-2 Value-Based-Distributional | DQN, DDQN<br>C51<br>QR-DQN |
| RL4-1 Policy-Gradient<br>RL4-2-Actor-Critic<br>RL4-3-TRPO-PPO<br>RL4-4-DDPG | PPO<br>TD3<br>SAC<br>DDPG |
| RL5-2-Exploration-RL | RND |
| RL5-1-AlphaZero-Muzero | AlphaZero<br>MuZero |
| RL5-4-Value-Based-Coop-MARL<br>RL5-5-Policy-Based-Coop-MARL | Non-stationarity<br>CTDE, IGM |
| RL5-3-from-Demonstration | DQfD |