Chapters	Main points
RL1-2-Lightweight-Cases.pdf RL2-1-MDP-DP.pdf RL2-2 Model-Free.pdf	UCB MDP Q-learning ε-greedy Contraction mapping
RL3-1 Value-Based.pdf RL3-2 Value-Based-Distributional.pdf	DQN, DDQN QR-DQN TD, MC
RL4-1 Policy-Gradient.pdf RL4-2-Actor-Critic.pdf RL4-3-TRPO-PPO.pdf RL4-4-DDPG.pdf	PPO TD3 SAC DDPG
RL5-2-Exploration-RL.pdf	RND
RL5-1-AlphaZero-Muzero.pdf	AlphaZero MuZero
RL5-3-from-Demonstration.pdf	DQfD
RL5-4-Value-Based-Coop-MARL.pdf RL5-5-Policy-Based-Coop-MARL.pdf	Non-stationarity CTDE IGM VDN QMIX MAPPO/HAPPO