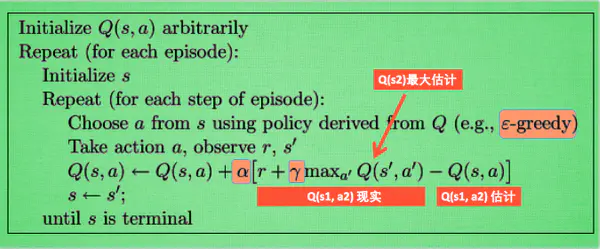
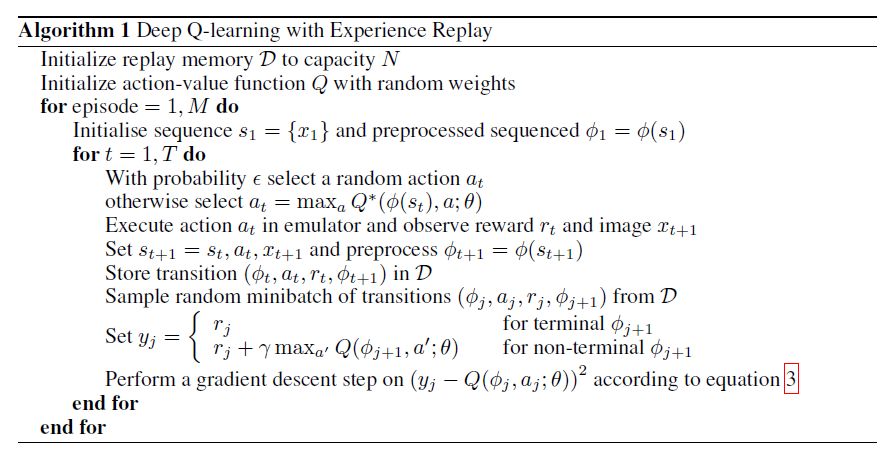
Value-based:

初始算法：Q-learning (TD control)



DQN:



Policy-based:

Policy-Gradient:

Actor Critic: