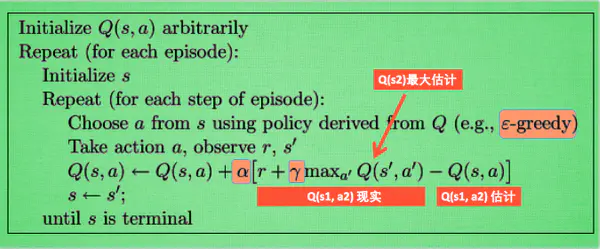
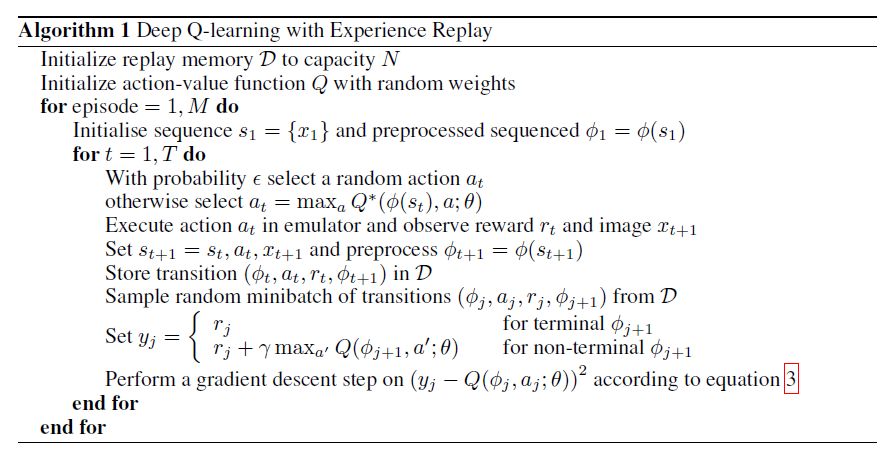
Value-based:

初始算法：Q-learning (TD control)



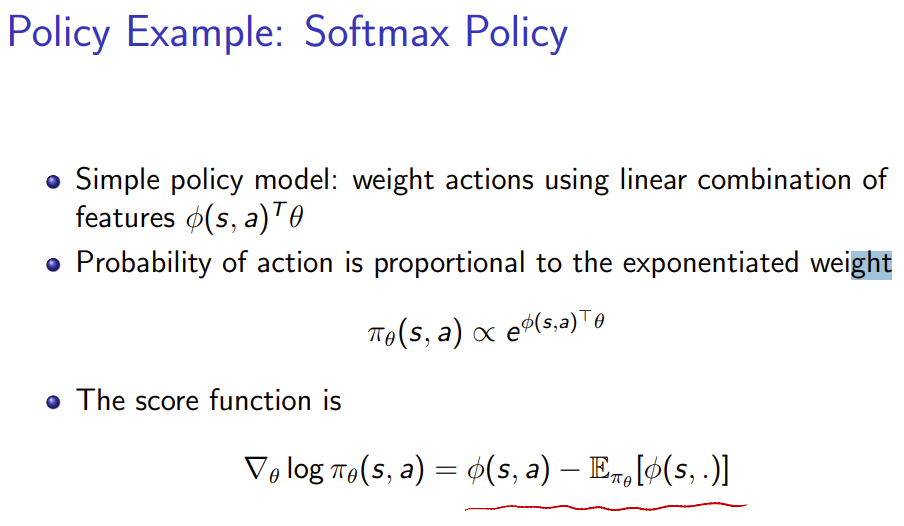
DQN:

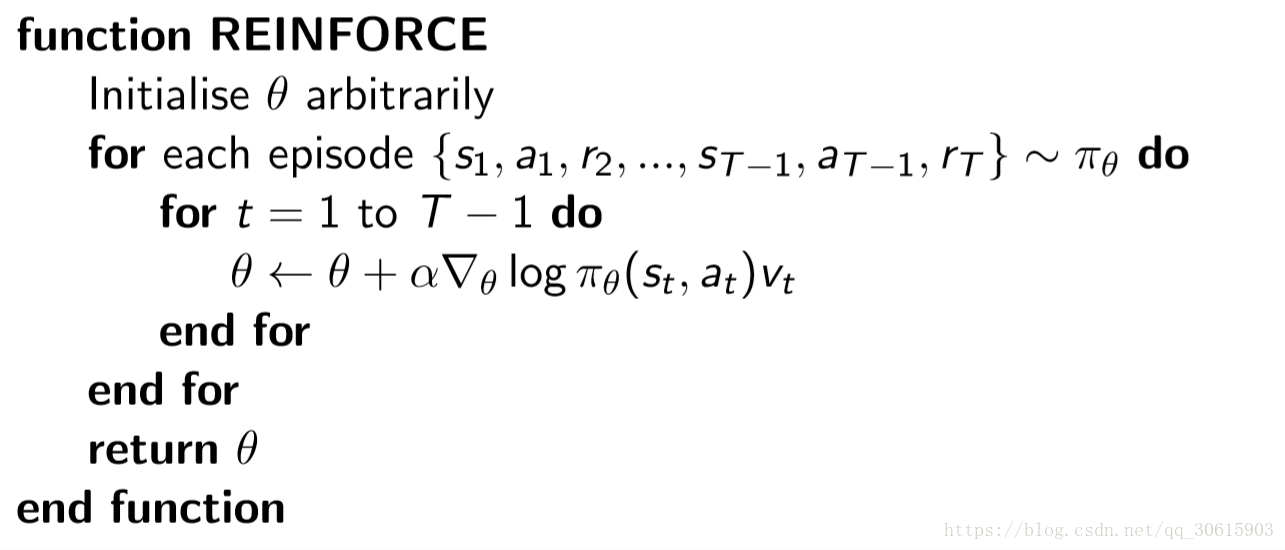
https://zhuanlan.zhihu.com/p/21421729?refer=intelligentunit



Policy-based:

Policy-Gradient:





<https://blog.csdn.net/qq_30615903/article/details/80747380>

Actor Critic:

（基于TD control）https://blog.csdn.net/qq\_30615903/article/details/80774384

