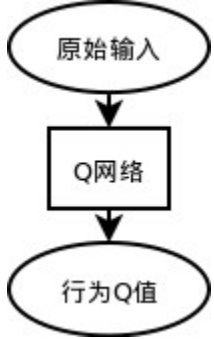
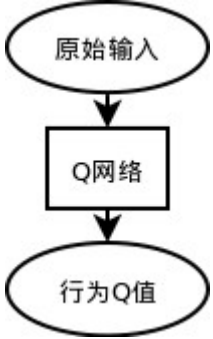
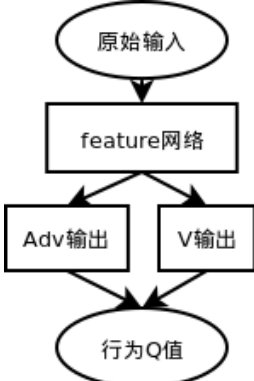
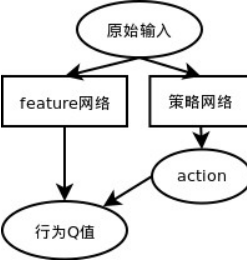
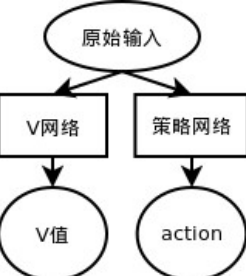


1) 策略对比：

DQN	Double DQN	Duel DQN	DDPG	A3C
离策略	离策略	离策略	离策略	离策略，异步训练
				
需要	需要	需要	需要	不需要
greedy-e	greedy-e	greedy-e	ou 过程	随机策略采样
最小化 target Q 和 Q 的误差	最小化 target Q 和 Q 的误差	最小化 target Q 和 Q 的误差	最大化 action 和 action 对应的 Q 值的联合，最小化 Q 值和 target Q 值的差	最大化 action 和 V 值的联合，最小化 V 值和 target V 值的差

2) 策略对比：

