

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \lambda \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$$

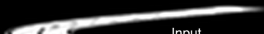
(The New Action Value = The Old Value) + The Learning Rate × (The New Information − the Old Information)



advanced RL

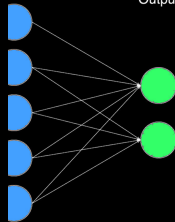
- continuous controll
- curiosity
- imagination

Michal CHOVANEC

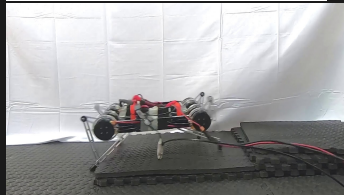
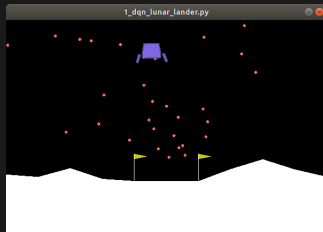
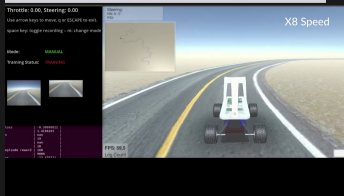


Hidden

Output



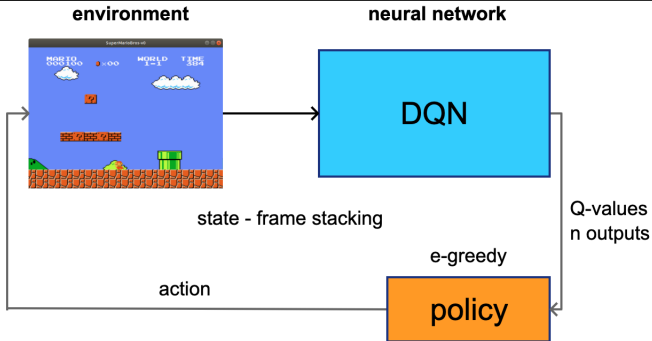
continuous actions space



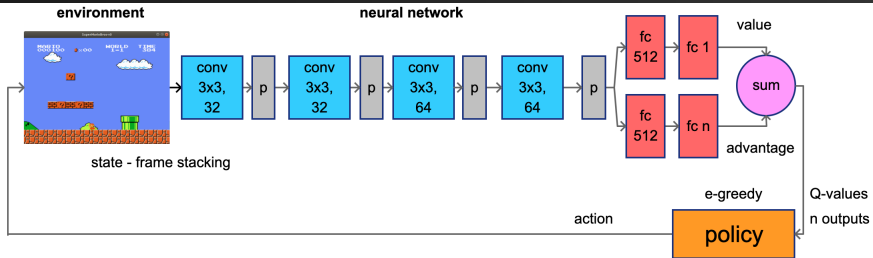
- discrete actions space
 - Deep Q-network, DQN
 - Dueling DQN
 - Rainbow DQN
- continuous actions space
 - Actor Critic
 - Advantage Actor Critic
 - Proximal policy optimization
 - Soft Actor critic
 - Deep deterministic policy gradient
 - D4PG, SDDPG

f.e. SDDPG - sampled DDPG, based on Wasserstein loss : Optimal transport, Cédric Villani, 600+ pages

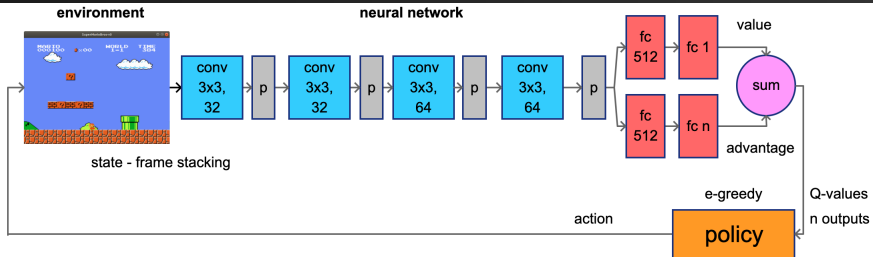
from DQN to DDPG



from DQN to DDPG

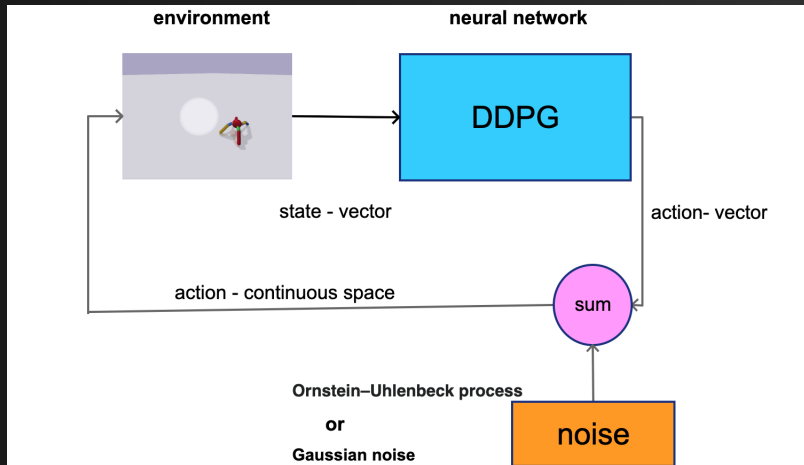


from DQN to DDPG

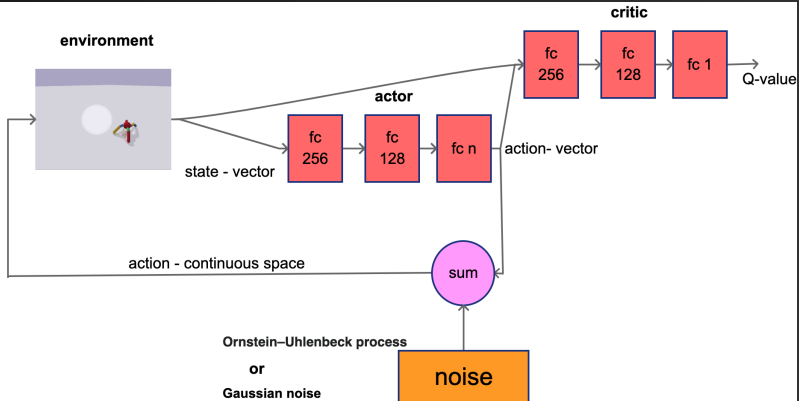


$$\mathcal{L}(\theta) = \left(R + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right)^2$$

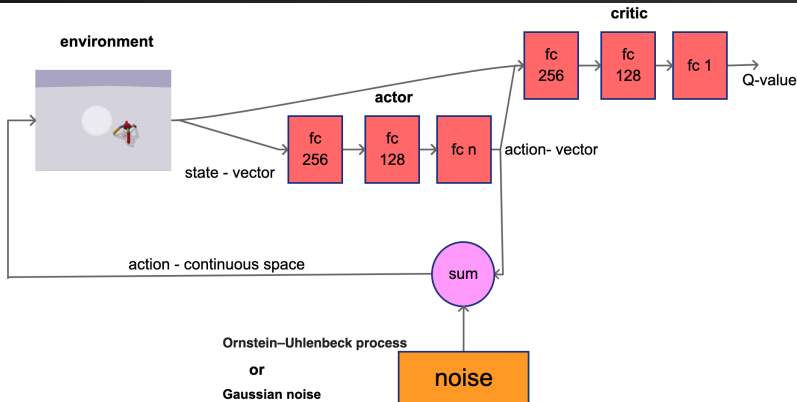
from DQN to DDPG



from DQN to DDPG



from DQN to DDPG



critic loss:

$$\mathcal{L}_c(\theta) = (R + \gamma Q(s', a'; \theta^-, \phi^-) - Q(s, a; \theta, \phi))^2$$

actor loss :

$$\mathcal{L}_a(\phi) = -Q(s, a; \theta, \phi)$$

wise Wizard's DDPG spell chart

- **neurons count** on 1st layer = 10x state vector size
- **neurons count** on 2nd layer = 0.5x neurons on 1st layer
- **weight init** for hidden layers : use Xavier
- **weight init** actor output : use uniform $\langle -0.3, 0.3 \rangle$
- **weight init** critic output : use uniform $\langle -0.003, 0.003 \rangle$
- **gaussian noise** : linear decay variance, from 1 to 0.3, for 1M steps
- use **soft** target network update, $\tau = 0.001$
- actor learning rate $\eta_a = 0.0001$
- critic learning rate $\eta_c = 0.0002$

wise Wizard's magic staff

- fully connected nets (robotic envs) **train on CPU** - AMD Ryzen
- convolutional nets (visual inputs envs) **train on GPU** - NVIDIA GTX1080+
- use fast CPU - envs are slow
- 32GB of RAM is enough
- for small visual envs (Atari, DOOM, Nec) - GTX1080ti



curiosity in RL

$$Q'(s, a) = R + \beta C(s, s', a) + \gamma \max_{a'} Q(s', a')$$

