RL methods:

1. **Value Based**: mapping between an action and a value (optimal value). Sample efficient and steady.

- Q value, Bellmann equation

- Deep Q networks, Double Dueling Q Networks

2. **Policy Based**: find optimal policy directly without the Q value. They are better for continuous and stochastic environments, have a faster convergence

- Policy Gradients

- REINFORCE

3. **Actor-Critics:**

Both value- and policy-based. The actor decides which action to take; the critic tells the actor how good its action was and how it should adjust.

- A2C (Advantage Actor-Critic): (Improved version: synchronous, with multiple agents)

- Q(s, a) = V(s) + A(s, a)

(V(s) = state value function, A(s, a) = Advantage)

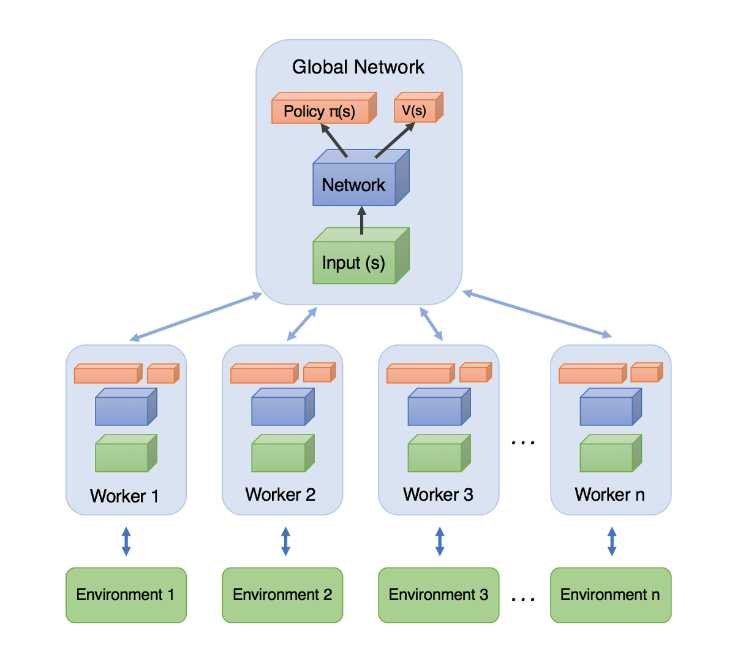
- A3C (Asynchronous Advantage Actor-Critic):

- released by DeepMind in 2016. Simple, robust, fast, achieving higher scores.

- multiple agents with their own weights, interacting with a different copy of

the environment in parallel

- drawback: some agents will be using older version of parameters



***A self-learning finite element extraction system based on reinforcement learning***

*Jie Pan1, Jingwei Huang2 , Yunli Wang3, Gengdong Cheng4 and Yong Zeng*

1. Use A2C to generate high-quality quadrilateral elements while maintaining the remaining geometry’s quality for the continuous generation of good quality elements.

2. Take the good quality elements generated from the A2C method as samples to train an FNN for fast generation of high-quality meshes. Relieves the human algorithm designers.

Smart designing of the smart element extraction system:

A smart system can make adaptive decisions corresponding to various environmental situations. The smartness of the element extraction system is two-fold:

1) how various complex boundary shapes can be processed by three element extraction rules. These three rules alone are sufficient to process any environmental situation.

2) among all boundary edges of a geometric domain, the system is able to smartly pick an optimal edge and its corresponding rule to generate a quadrilateral.

Questions from paper: