Problem

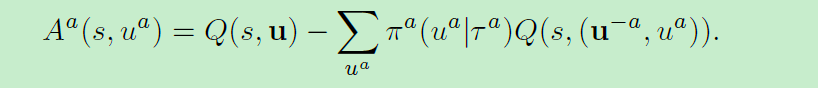
Multi-agent credit assignment.

Method

Actor-Critic

Don’t use TD error. (don’t take into account agent’s contribution)

Use counterfactual baseline. (take into account agent’s contribution)



(Advantage) Contribution of ua = Q of ua - the expectation of Q of ua