


# RL基础-10-Actor-Critic Method

[第10课-Actor-Critic方法（最简单的Actor-Critic \(QAC\)）\\_哔哩哔哩\\_bilibili](#)

最后一节课；最流行的方法之一；  
把value的方法引入到policy gradient中；

## 引入

- actor, critic, 分别对应于policy update和policy evaluation/value-estimate；

image.png

主要内容

image.png

## AC基础

之前讲到计算 $q_t$ 的值有两种方法：

- MC计算，REINFORCE方法，或者叫MC policy gradient；
- TD方法，叫做actor-critic方法；

## 基础算法

image.png