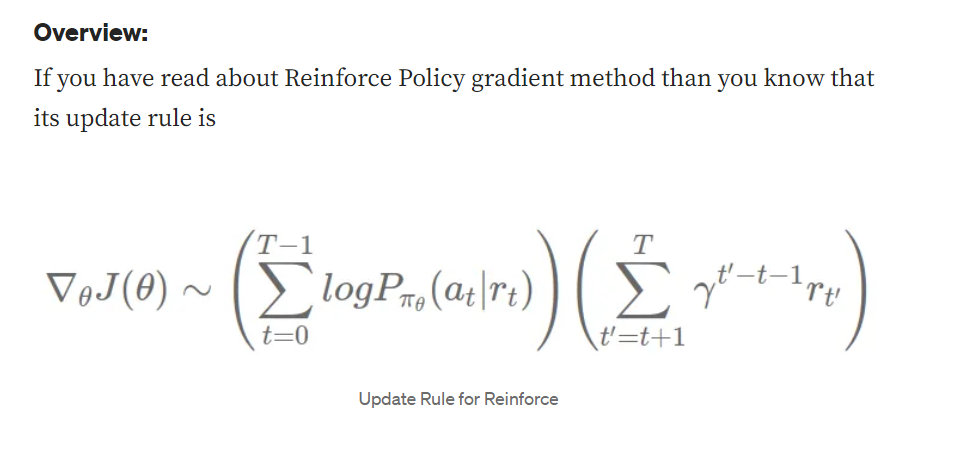
**Actor Critic (Reinforcement Learning)**

****

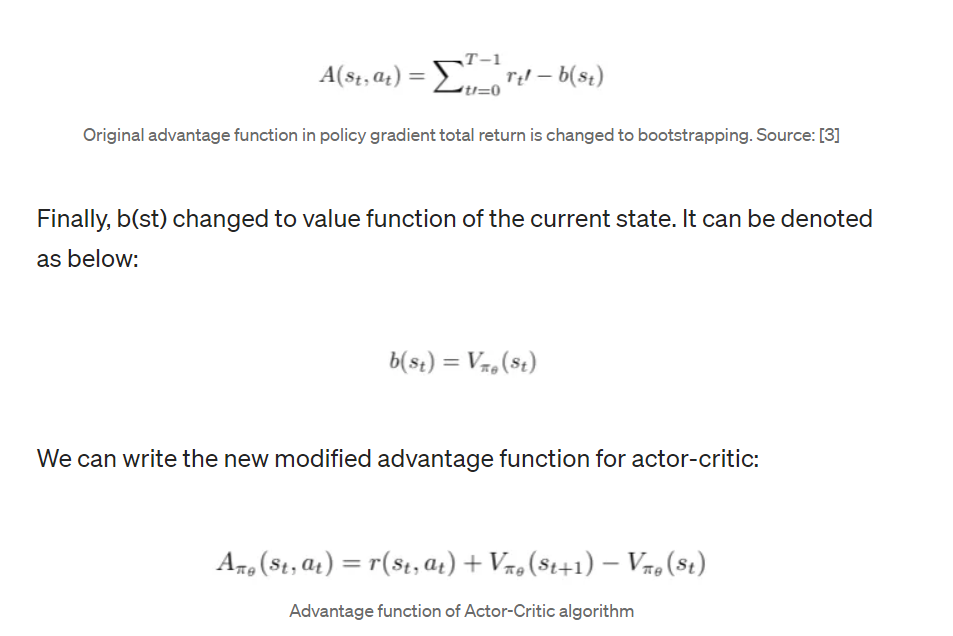
**Actor-critic**

In the Actor-Critic method, the policy is referred to as the actor that proposes a set of possible actions given a state, and the estimated value function is referred to as the critic, which evaluates actions taken by the actor based on the given policy.

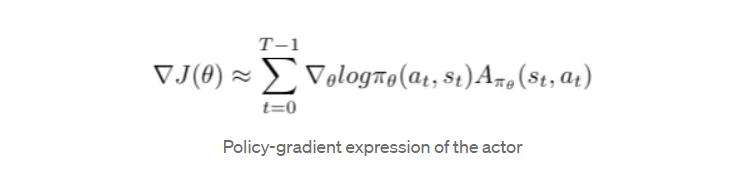
**In a simple term, Actor-Critic is a Temporal Difference (TD) version of Policy gradient.** It has two networks: Actor and Critic. The actor decided which action should be taken and critic inform the actor how good was the action and how it should adjust. The learning of the actor is based on policy gradient approach. In comparison, critics evaluate the action produced by the actor by computing the value function.

This type of architecture is in Generative Adversarial Network(GAN) where both discriminator and generator participate in a game. The generator generates the fake images and discriminator evaluate how good is the fake image generated with its representation of the real image. Over time Generator can create fake images which cannot be distinguishable for the discriminator. Similarly, Actor and Critic are participating in the game, but both of them are improving over time, unlike GAN.

Actor-critic is similar to a policy gradient algorithm called REINFORCE with baseline. Reinforce is the MONTE-CARLO learning that indicates that total return is sampled from the full trajectory. But in actor-critic, we use bootstrap. So the main changes in the advantage function.

****

Alternatively, advantage function is called as TD error as shown in the Actor-Critic framework. As mentioned above, the learning of the actor is based on policy-gradient. The policy gradient expression of the actor as shown below:



# Pseudocode of Actor-Critic algorithm

1. Sample {s\_t, a\_t}using the policy πθ from the actor-network.
2. Evaluate the advantage function A\_t. It can be called as TD error δt. In Actor-critic algorithm, advantage function is produced by the critic-network.

