Monte Carlo methods can be used to evaluate the performance of a policy in an episodic system by….

There are 2 variants of the Monte Carlo algorithm, first visit and every visit.

With first visit algorithms… update v(s) (expected value of the return) with every step of G (Expected return from step t = reward + gamma\*reward + gamma^2 etc.) calculated unless it is repeated. Advantage is it shows statistical independence

Every step algorithms.. update v(s) on every step regardless so some states will have multiple updates