MONTE CARLO

Aim:

Aim to analyze the stability and convergence properties of Monte Carlo methods in reinforcement learning.

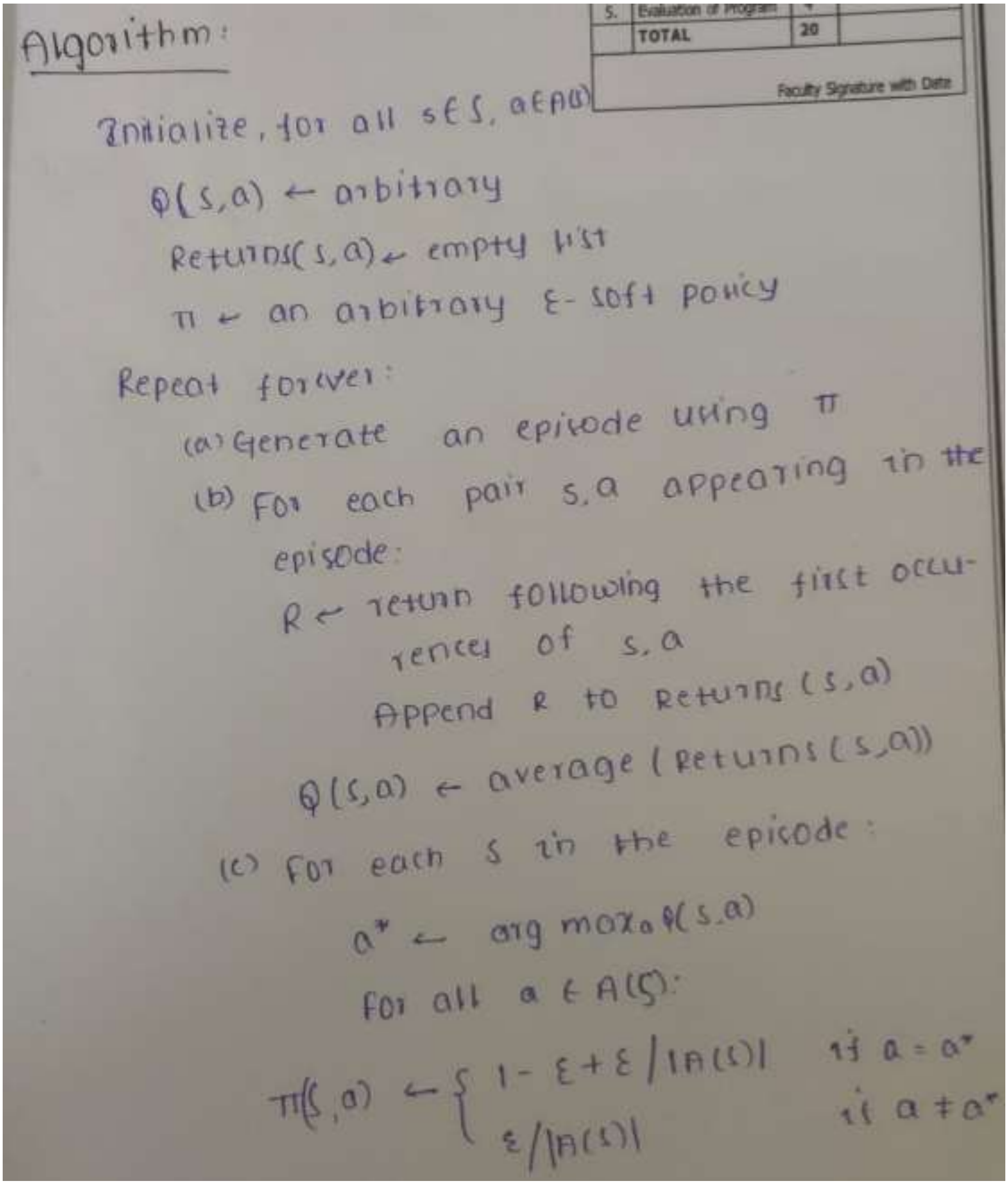
Description:

Monte Carlo

The agent collects experience by interacting with the environment then uses this experiences to estimate the value of states or state-action pairs by averaging the returns obtained from different trajectories or episodes.

A Monte Carlo's stimulation is a model used to project the probability of a variety of outcomes the potential for random variables is present. Monte carlo simulations helps to explain the impact of risk & uncertaininty in predictions & forecasting models.

Algorithm.



5.	Evaluation of Program		
	TOTAL	20	

Faculty Signature with Date

Algorithm:

```
Initialize, for all $s \in S$, $a \in A(s)$
    $Q(s,a) \leftarrow$ arbitrary
    $Returns(s,a) \leftarrow$ empty list
    $\pi \leftarrow$ an arbitrary $\epsilon$-soft policy
Repeat forever:
    (a) Generate an episode using $\pi$
    (b) For each pair s,a appearing in the episode:
        $R \leftarrow$ return following the first occurences of s,a
        Append R to Returns(s,a)
        $Q(s,a) \leftarrow$ average(Returns(s,a))
    (c) For each s in the episode:
        $a^* \leftarrow \arg\max_a Q(s,a)$
        For all $a \in A(s)$:
        $\pi(s,a) \leftarrow \begin{cases} 1-\epsilon+\epsilon/|A(s)| & \text{if } a=a^* \\ \epsilon/|A(s)| & \text{if } a \neq a^* \end{cases}$
```

Result:

Successfully Implemented the Monte carlo Algorithm.