≡ Implementation

Implementation: MC Control: GLIE

The pseudocode for (first-visit) GLIE MC control can be found below. (*Feel free to implement either the first-visit or every-visit MC method. In the game of Blackjack, both the first-visit and every-visit methods return identical results.*)

GLIE MC Control

Input: positive integer $num\_episodes$
Output: policy π ($\approx \pi_*$ if $num\_episodes$ is large enough)
Initialize $Q(s, a) = 0$ for all $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$
Initialize $N(s, a) = 0$ for all $s \in \mathcal{S}, a \in \mathcal{A}(s)$
for $i \leftarrow 1$ **to** $num\_episodes$ **do**
    $\epsilon \leftarrow \frac{1}{i}$
    $\pi \leftarrow \epsilon\text{-greedy}(Q)$
    Generate an episode $S_0, A_0, R_1, \ldots, S_T$ using π
    for $t \leftarrow 0$ **to** $T - 1$ **do**
        if (S_t, A_t) *is a first visit (with return G_t)* **then**
            $N(S_t, A_t) \leftarrow N(S_t, A_t) + 1$
            $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \frac{1}{N(S_t, A_t)}(G_t - Q(S_t, A_t))$
    end
end
return π

Please use the next concept to complete **Part 3: MC Control: GLIE** of `Monte_Carlo.ipynb`. Remember to save your work!

If you'd like to reference the pseudocode while working on the notebook, you are encouraged to open **this sheet** in a new window.

Feel free to check your solution by looking at the corresponding section in `Monte_Carlo_Solution.ipynb`.

Implementation