Setting the Value of α

Remember that the `forgetful_mean` function is closely related to the **Evaluation** step in constant-α MC control. You can find the associated pseudocode below.

Evaluation

Generate an episode $S_0, A_0, R_1, \ldots, S_T$ using π.
For $t \leftarrow 0$ to $T-1$:
    If (S_t, A_t) is a first visit (with return G_t):
        $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(G_t - Q(S_t, A_t))$

Before moving on to the next concept, use the above coding environment to verify the following facts about about how to set the value of α when implementing constant-α MC control.

- You should always set the value for α to a number greater than zero and less than (or equal to) one.
    - If $\alpha = 0$, then the action-value function estimate is never updated by the agent.
    - If $\alpha = 1$, then the final value estimate for each state-action pair is always equal to the last return that was experienced by the agent (after visiting the pair).
- Smaller values for α encourage the agent to consider a longer history of returns when calculating the action-value function estimate. Increasing the value of α ensures that the agent focuses more on the most recently sampled returns.

Note that it is also possible to verify the above facts by slightly rewriting the update step as follows:

$$Q(S_t, A_t) \leftarrow (1-\alpha)Q(S_t, A_t) + \alpha G_t$$