Reinforcement Learning: Assignment #8

Due on Sunday, July 2, 2023

Group 4 - Abu El Komboz, Tareq 3405686 | **Jain, Likhit 3678905** | **Wurm, Marcel 3695946**

Task 1

n-step and eligibility traces (4P)

(a) The nonplanning method looks particularly poor in the figure below; a method using n-step bootstrapping would do better. Do you think one of the n-step bootstrapping methods could do as well as the Dyna-Q method? Explain why or why not. (2P)

> Dyna-Q should perform better as it similarly uses something like ϵ-greedy to find the first path to the goal but during search already optimises the path to the currently farthest reached state and should thus have a pretty good path when the goal is reached for the first time.
> Although it could be that the Dyna-Q algorithm takes longer to calculate because of the additional planning steps if n-step bootstrapping can quickly find the path from the planned states on reaching the goal backwards to the start.

(b) Give the λ-return G_t^λ as recursive function in terms of the first reward R_{t+1} and itself one step later G_{t+1}^λ. (2P)

$$G_t^\lambda = (1 - \lambda) \left[\lambda^0 G_{t:t+1} + \sum_{n=2}^{\infty} \lambda^{n-1} G_{t+1:t+n} \right]$$
$$= (1 - \lambda) \left[R_{t+1} + \gamma V(S_{t+1}) \right] + G_{t+1}^\lambda$$

Task 2

n-step Sarsa on the FrozenLake (5P)

(a) Implement n-step Sarsa and evaluate it on the 8 × 8 Frozen Lake environment. Evaluate the performance for different choices of n and α. Visualize your results (plot the performance over α for different choices of n, similar to lecture 8 slide 12).

> TODO