

IIC3675: Tarea 4

Bruno Cerda Mardini

a)

Ambos métodos obtienen resultados similares, por lo menos en la práctica, sin embargo, SARSA claramente logra obtener valores menores de largos de promedios, pero yo no diría que es una mejora extremadamente relevante.

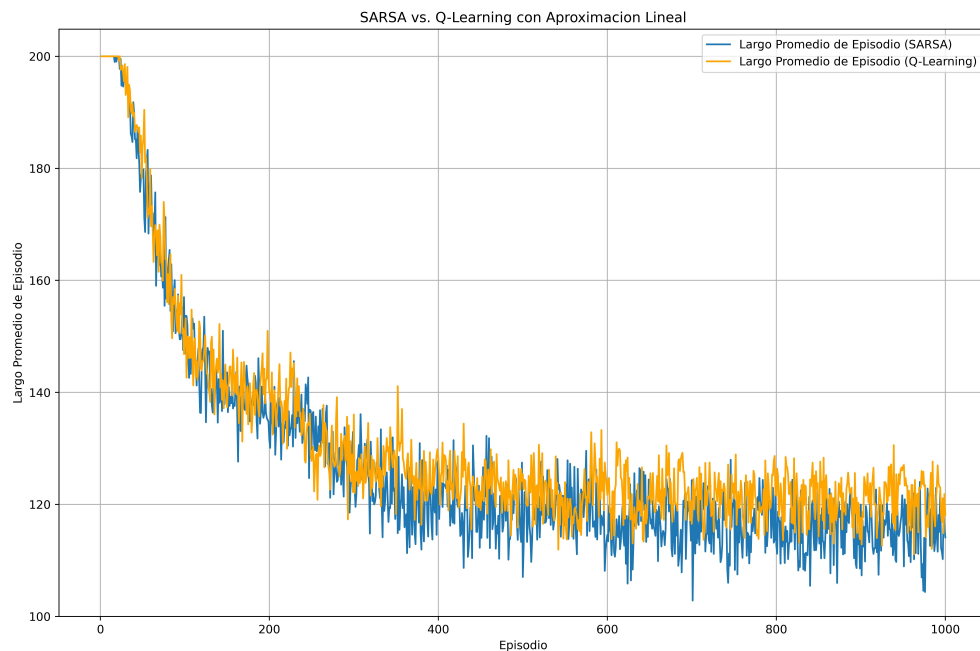


Figure 1: Largo promedio de episodios: SARSA vs Q-learning con aproximación lineal.

b)

hola

c)

La figura 3 muestra el comportamiento de Actor-Critic con aproximación lineal en el environment de MountainCarContinuous-v0. Se puede observar que si está ocurriendo aprendizaje, principalmente debido a que el largo promedio de los episodios va disminuyendo.

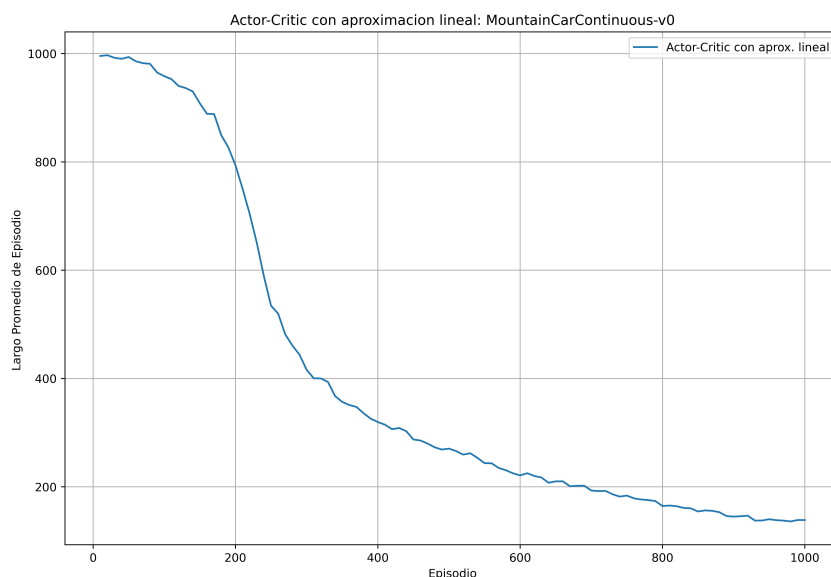


Figure 2: Largo promedio de episodios: Actor-Critic con aproximación lineal.

d)

- Hay mucha variabilidad como para poder decir que es fácil o difícil para DDPG, ya que con parámetros incorrectos simplemente no se aprende nada, pero con parámetros correctos se logra aprender una estrategia para solucionar el problema. Pero si tuviera que elegir, yo diría que es más difícil que fácil, principalmente debido a que DDPG es muy sensible a la elección de hiperparámetros.
- Partí con los hiperparámetros default que tiene la librería, no tuve dificultades mayores ya que el primer parámetro que utilicé logró que hubiera aprendizaje, y encontré que el resto no era tan relevante.
- En la clase vimos resultados de DDPG con un tipo de ruido llamado Ornstein Uhlenbeck, este fue el primer parámetro que implementé y DDPG pudo aprender a resolver el problema. Sí me hace sentido que este sea el más relevante ya que en este entorno está el problema de Sparse Rewards, y este ruido ayuda a que el modelo pueda explorar de mejor manera y lograr conseguir rewards.
- Mi configuración final de hiperparámetros fueron todos los parámetros que vienen por default en la librería, con excepción de `action_noise`, en donde utilicé `OrnsteinUhlenbeckActionNoise` con `media = 0` y `sigma = 0.75`. También probé con `sigma = 0.5`, y también funcionaba, pero también intenté con `sigma = 0.25` y ahí el modelo volvía a fallar. Para contrarrestar esto también intenté con `media = 2` y `sigma 0.25` pero de todas formas fallaba. Debido a lo anterior, decidí quedarme con `media=0` y `sigma=0.75`.
- (Todavía no sé, hay que esperar los resultados)