

Reinforcement Learning

Taller #3: Algoritmos de aprendizaje por refuerzo tabulares.

En este taller se trabajará sobre la modificación del juego de escaleras y serpientes del taller anterior.

1. Considere el MDP que usted formuló para modelar este problema en el taller anterior. Escriba un módulo de Python que permita la interacción de un agente con el ambiente para el MDP formulado. Su implementación debe:
 - Dada una acción en un estado, retornar la recompensa y el estado resultante en esa acción. Considere el caso especial del estado terminal.
 - Ejecutar una política arbitraria.
2. Usando su módulo del MDP, implemente el algoritmo de control de Montecarlo off-policy para encontrar una política óptima.
3. Usando su módulo del MDP, implemente el algoritmo SARSA on-policy para encontrar una política óptima.
4. Usando su módulo del MDP, implemente el algoritmo Q-Learning para encontrar una política óptima.
5. Ejecute 1000 partidas con cada una de las políticas encontradas y compare los resultados obtenidos en cuanto a:
 - Porcentaje de partidas ganadas.
 - Número de pasos promedio en los episodios.