



MACHINE LEARNING II

APRENDIZAJE POR REFUERZO

ENTORNO ACROBOT

Juan Miguel Ramos Pugnaire

Andrés Canalejo Oliva

Pablo Pérez González-Alberto

Pablo Cardenal Real de Asúa

12 de mayo de 2023

Índice

1. Introducción	3
2. Objetivos	4
3. Métodos seleccionados	4
4. Deep Q-Learning	4
5. Double Deep Q-Learning	4
6. Algoritmo REINFORCE	4

1. Introducción

El aprendizaje por refuerzo (RL por sus siglas en inglés) se basa en la idea de que un agente aprende a tomar decisiones interactuando con un entorno. De esta forma, el agente aprende a maximizar una recompensa basada en sus acciones, lo que le permite aprender a realizar tareas complejas a través de la exploración del entorno y la adaptación a los cambios en él.

En el presente trabajo, se utilizará RL para entrenar a un agente que controle el movimiento del robot *Acrobot*, utilizando el paquete `gym` de OpenAI.

El entorno consiste en un robot de dos brazos que puede girar alrededor de su base. El robot se encuentra en un estado inicial colgando hacia abajo y debe alcanzar una posición objetivo que se define como el momento en que el extremo superior del segundo brazo del robot alcanza una altura específica. En la siguiente figura se representa la posición inicial del agente y la altura que debe alcanzar para obtener la recompensa:

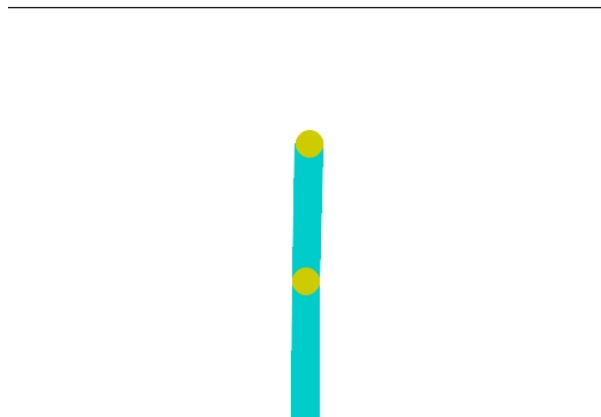


Figura 1: Posición inicial del robot controlado por el agente

Veamos en profundidad el espacio de acciones y el entorno del agente

- 2. Objetivos**
- 3. Métodos seleccionados**
- 4. Deep Q-Learning**
- 5. Double Deep Q-Learning**
- 6. Algoritmo REINFORCE**