

Tarea 3: Aprendizaje Por Refuerzos

Problema

Consideraremos el problema de una nave espacial que quiere aterrizar en una luna. La nave tiene tres propulsores, uno principal apuntando hacia abajo, y uno apuntando a cada lado. Este problema es conocido como el problema del aterrizaje lunar, o Lunar Lander.

En cada instante, la nave puede realizar una acción: activar un propulsor, o no hacer nada.

Se propone implementar una solución a dicho problema utilizando un algoritmo de Q-Learning, haciendo uso de un ambiente de open-ai gymnasium.

El cuaderno de python adjunto¹ contiene la descripción del problema, el ambiente y ejemplos de agentes sin aprendizaje.

La entrega deberá ser realizada en una copia de este cuaderno, donde deberán:

1. Programar las funciones de la clase `AgenteRL`, manteniendo cualquier función adicional necesaria dentro de la clase.
2. Analizar los resultados de una ejecución de mil episodios con el agente programado. Agregar uno más bloques de texto nuevos discutiendo los resultados obtenidos.

Criterio de Evaluación

Dado que la dificultad del problema a resolver es elevada, consideraremos una implementación exitosa si las ejecuciones sin exploración (de explotación de lo aprendido) supera el 1%.

Se tomará especial atención a la evidencia de un entendimiento del algoritmo de Q-Learning, su aplicación a este problema y los resultados obtenidos en la documentación. También tendrá gran importancia la evidencia de cómo llegaron a los parámetros empleados.

Sugerencias

Realizar tareas de control sobre tantas variables, y discretizadas, puede requerir un espacio de estados grande. Esto afecta la cantidad de episodios necesarios para el entrenamiento.

Cuando se está iterando sobre la implementación del algoritmo, es importante mantener las iteraciones cortas, basándose en indicadores más tempranos de mejora. Por ejemplo, se puede utilizar la evolución positiva de la recompensa promedio, sin tener que esperar a ver si una implementación soluciona la tarea luego de el entrenamiento completo.

Entregables

El informe a entregar debe ser una copia del cuaderno provisto, donde se incluirá:

- el código escrito para resolver el problema
- un informe con las pruebas realizadas y los resultados obtenidos

Fecha límite de entrega

¹<https://gitlab.fing.edu.uy/aprendaut2/AA24/-/blob/main/lab3/tarea.ipynb>

Miércoles 16 de Octubre (inclusive).