

COMPUTACIÓN INTELIGENTE

Reinforcement Learning

Joan Grau 172578

Ignacio Estaún 173987

¿Qué funciona?

Las preguntas obligatorias Q1-Q7

¿Qué problemas encontraron?

Guión de la práctica poco claro.

Breve respuesta a aquellas preguntas que la requieran

Question 3

- a) Ponemos un factor de descuento bajo de forma que prefiere el camino más corto. Ruido bajo para que coja el camino cercano a las penalizaciones.
- b) El mismo que el anterior, pero subiendo el ruido, preferirá los caminos largos de las penalizaciones.
- c) Para que prefiera el camino largo, subimos el factor de descuento. Ruido bajo para que se acerque a las penalizaciones.
- d) De nuevo el mismo que el anterior, pero para evitar el camino cercano a las penalizaciones, subimos el ruido.
- e) Para que no haga nada, ponemos un factor de descuento lo más grande posible.

Question 4

GetQValue: valor Q del estado que le pasemos y su acción

computeActionFromQValues: computar la próxima acción teniendo en cuenta el mayor valor

computeValueFromQValues: computar el valor máximo Q del siguiente estado

getAction: devuelve la siguiente acción con mejor Q

update: actualizamos el valor de los estados

Question 5

modificamos getAction para que escoja entre la siguiente acción con mayor Q o una random si la probabilidad es igual a un valor igual a epsilon

Question 6

No llegamos a ningún valor que permitiese llegar al estado final, debido al algoritmo usado.