

Fundamentos de los Sistemas Inteligentes

Práctica 3

Q-Learning es uno de los algoritmos más utilizados en aprendizaje por refuerzo. En el código base entregado se dispone de un agente situado en un entorno con forma de malla bidimensional rectangular. Las acciones posibles a realizar serían mover el agente arriba, a la derecha, abajo y a la izquierda (con las obvias restricciones si se encuentra en los bordes). Tal y como está el código, el agente realiza tan solo exploración hasta llegar al estado objetivo con recompensa.

Tareas:

- Ejecutar y calcular la cantidad promedio de acciones que realiza el agente antes de llegar al objetivo.
- Modifica la política del agente para que alterne entre exploración y explotación. Implementa la opción greedy y ϵ -greedy (con distintos valores para el ϵ).
- Calcular y comparar el promedio de acciones que realiza el agente antes de llegar al objetivo con estas nuevas políticas.