Reinforcement Learning

Taxi Learning

Luccas Rojas 201923052 31/03/2023

Para el caso del taxi genere una matriz de estados que era una tripla, donde i,j era la posición en x y y del taxi, mientras que k era si el taxi tenía o no pasajero. Debido a que el algoritmo no convergía decidí usar sólo el punto azul como lugar para recoger pasajeros y el punto amarillo para dejar los pasajeros recogido en el punto azul. El algoritmo sin embargo seguía sin converger, esto se daba debido a que la política óptima no era sólo una, esto se debía a que varias casillas tenían un valor muy similar al ir en una dirección u otra, por lo que de forma aleatoria la política podía alternar entre cualquiera de dichas políticas, debido a esto tuve que frenar el algoritmo en las 300 iteraciones, el resultado se presenta a continuación:

Calendario

Descripción generada automáticamente

Los parámetros utilizados para obtener el resultado fueron: Un épsilon de 0.7, un factor de aprendizaje de 0.2, una recompensa por movimiento de -0.1 y un factor de descuento de 0.6, estos valores fueron ajustados para obtener una política de forma rápida gracias a la exploración que sucedía en el 30% de los casos. Se observa como el agente es capaz de aprender el camino para recoger y dejar a un pasajero dado un punto de inicio al azar sobre el ambiente.