

Universidad del Valle de Guatemala
Facultad de ingeniería



Laboratorio 7

Cayetano Molina 20211
Estefanía Elvira 20725
Priscilla Gonzalez 20689

Guatemala 24 de septiembre del 2024

Task 1

Explique la diferencia entre los métodos de aprendizaje de refuerzo basado en valores y en políticas. ¿Por qué los métodos de gradiente de políticas son especialmente útiles para entornos con espacios de acción continua?

Los métodos de aprendizaje por refuerzo basados en valores tratan de aprender una función de valor la cuál decide o ver qué tan bueno es hacer una acción estando en cierto estado. Se puede encontrar el Q-Learning y SARSA entre uno de los modelos más comunes. Ahora bien, el método basado en política, como su nombre lo indica, aprenden directamente a través de una política, ya sea estocástica o determinística.

Por otro lado, al hablar de acción continua se entiende que es un número finito de acciones en cada uno de los estados. Por lo que en este caso, los métodos basados en valores deberían de pasar por cada una de las acciones haciendo así ineficiente el proceso causando así un infinito en un espacio continuo. En este caso, es mejor utilizar métodos de aprendizaje por refuerzo basados en políticas, ya que las políticas pueden ser fácilmente ajustadas a lo que se requiere para que así elijan o tomen mejores acciones. Además, son eficientes explorando espacios continuos y fluidos.