

Fecha: 16/12/2024

Evaluación: Parcial

Materia: Inteligencia Artificial

Turno: Nocturno

Duración: 2 hs

Uso de Calculadora: SI

Uso de Material: NO

Puntaje Máximo: 40 Puntos

Página 1 de 1

Ejercicio 1 - 10 puntos

1. Explique como funciona MCMC, el método Monte Carlo por el cual a partir de un conjunto de episodios podemos modelar un MDP. (Si le facilita, puede explicarlo a partir de un ejemplo)
2. Cuando hablamos de $TD(0)$ o $TD(\infty)$ ¿qué significa el ese parámetro que le pasamos a TD? Mencione ambos casos particulares $TD(0)$ y $TD(\infty)$. Máximo 5 renglones.
3. ¿Cómo se relaciona Q-Learning con $TD(0)$? Máximo 2 renglones.

Ejercicio 2 - 10 puntos

1. Defina qué es un juego de suma 0. Máximo 5 renglones. Aclaración: no se pide explicar qué es un juego
2. Si el oponente va a jugar de forma óptima, ¿por qué nos convendría usar minimax, en vez de expectimax?
3. Explique qué problema de minimax/expectimax resuelve la función de evaluación, e indique cual es la condición necesaria para ser considerada una buena función de evaluación
4. Cuente cómo fue el proceso de encontrar una buena función de evaluación en su obligatorio

Ejercicio 3 - 5 puntos

1. ¿Qué problema intenta afrontar el artículo sobre *Stochastic Q-learning* que se hace referencia en la letra del obligatorio?
2. ¿Qué mecanismo propone el artículo para solucionar esto?

Ejercicio 4 - 15 puntos

Considere el siguiente juego alternado de dos jugadores. Hay 9 cajas:

- C1: {C1.1, C1.2, C1.3, C1.4}
 - C1.1: {6, 9}
 - C1.2: {18, 9, -1}
 - C1.3: {10, 58008}
 - C1.4: {3, -14}
- C2: {-42, 2 147 483 647}
- C3: {C3.1, C3.2}
 - C3.1: {-4, 20}
 - C3.2: {-1, 9001}

El jugador A (agente) elige una caja entre C1, C2, y C3.

Si elige C1, se tira un dado cargado para decidir a que caja ir, con $\frac{1}{2}$ de probabilidad de que salga 1, $\frac{1}{4}$ de probabilidad de que salga 2, y $\frac{1}{8}$ de probabilidad de que salga tanto 3 como 4. Luego el jugador O (oponente) elige un valor del conjunto resultante.

Si elige C2, el jugador O elige un valor del conjunto resultante.

Si elige C3, se tira una moneda cargada, con $\frac{1}{3}$ de probabilidad de que salga 1, y $\frac{2}{3}$ de probabilidad de que salga 2. Luego el jugador O elige un valor del conjunto resultante.

Se pide:

1. Indicar qué algoritmo tiene más sentido aplicar en esta situación entre: minimax, expectimax, expectiminimax
2. Dibujar el árbol del juego
3. Calcular los valores del juego para el algoritmo elegido, indicando la acción que debe tomar el agente A en cada estado