

1.

1 / 1 punto

Está utilizando el aprendizaje por refuerzo para controlar un robot de cuatro patas. La posición del robot sería su \_\_\_\_\_.

- ☐ devolver
- ☐ acción
- ☐ premio
- ☒ estado

✓ **Correcto**  
¡Excelente!

2.

1 / 1 punto

Estás controlando un rover de Marte. Estarás muy, muy feliz si llega al estado 1 (descubrimiento científico significativo), un poco feliz si llega al estado 2 (pequeño descubrimiento científico) e infeliz si llega al estado 3 (el vehículo está permanentemente dañado). Para reflejar esto, elija una función de recompensa para que:

- ☒  $R(1) > R(2) > R(3)$ , donde  $R(1)$  y  $R(2)$  son positivos y  $R(3)$  es negativo.
- ☐  $R(1) > R(2) > R(3)$ , donde  $R(1)$ ,  $R(2)$  y  $R(3)$  son positivos.
- ☐  $R(1) < R(2) < R(3)$ , donde  $R(1)$  y  $R(2)$  son negativos y  $R(3)$  es positivo.
- ☐  $R(1) > R(2) > R(3)$ , donde  $R(1)$ ,  $R(2)$  y  $R(3)$  son negativos.

✓ **Correcto**  
¡Buen trabajo!

3.

1 / 1 punto

Está utilizando el aprendizaje por refuerzo para volar un helicóptero. Usando un factor de descuento de 0.75, su helicóptero comienza en algún estado y recibe recompensas -100 en el primer paso, -100 en el segundo paso y 1000 en el tercer y último paso (donde ha llegado a un estado terminal). ¿Cuál es el retorno?

- ☐  $-100 - 0.25 \cdot 100 + 0.25^2 \cdot 1000$
- ☐  $-0.75 \cdot 100 - 0.75^2 \cdot 100 + 0.75^3 \cdot 1000$
- ☐  $-0.25 \cdot 100 - 0.25^2 \cdot 100 + 0.25^3 \cdot 1000$
- ☒  $-100 - 0.75 \cdot 100 + 0.75^2 \cdot 1000$

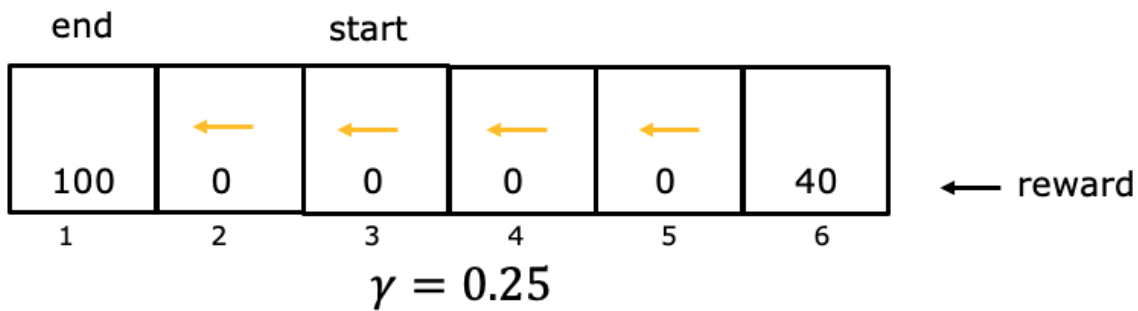
✓ **Correcto**

¡Impresionante!

4.

1 / 1 punto

Dadas las recompensas y acciones a continuación, calcule el rendimiento del estado 3 con un factor de descuento de  $\gamma = 0.25$ .



☐ 0.39

☒ 6.25

☐ 25

☐ 0



Correcto

Si a partir del estado 3, las recompensas están en los estados 3, 2 y 1. El retorno es  $0 + (0.25) \times 0 + (0.25)^2 \times 100 = 6.25$ .